DISSECTING THE MOLECULAR MECHANISM UNDERLINING THE INTERACTIONS WITH THE AUTOIMMUNE REGULATOR PROTEIN

Bader Alabdelaly

Submitted for the degree of Doctor of Philosophy

The University of Sheffield

Faculty of Medicine, Dentistry and Health

Department of Infection, Immunity and Cardiovascular Disease

March 2020

ABSTRACT

Autoimmune Polyendocrine Syndrome type 1 (APS-1) is an inherited autosomal disorder. The most common clinical features of the disease include primary adrenal insufficiency (Addison's disease), hypoparathyroidism (HP), and chronic mucocutaneous candidiasis (CMC).

APS-1 is caused by mutations in the autosomal regulator gene (AIRE). AIRE facilitates central tolerance by promoting the expression of peripheral tissue antigens (PTAs) in medullary thymic epithelial cells (mTECs), leading to the deletion of self-reactive thymocytes. It interacts with several proteins to play its role in transcription elongation. We therefore used the plant homeodomain 1 (PHD1) domain of the Autoimmune regulator gene (AIRE) to detect proteins interacting with it using a yeast-two-hybrid (Y2H) system. The latter allowed us to identify six potential partners of AIRE: Proliferating cell nuclear antigen (PCNA), ATPase Na+/K+ Transporting Subunit Beta 1 (ATP1B1), NPC intracellular cholesterol transporter 2 (NPC2), Phosphoglucomutase-1 (PGM1), RING finger-2 (RNF2), and Syntaxin-8 (STX8). PCNA is a protein that is involved in transcription and had previously been identified by another group as interacting with AIRE, which validated our screening.

The second part of the study involved characterising the effect of APS-1 mutations on AIRE, mainly those involving its caspase activation and recruitment (CARD) and SAND (named after **S**p100, **A**IRE-1, **N**ucP41/75, **D**EAF-1) domains. Our aim was to establish a correlation between AIRE-protein interaction and APS-1 mutations. First, we showed that AIRE-CARD interacts with Bromodomain-containing protein 4 (BRD4)- Death-associated protein 6 (DAXX)- Protein inhibitor of activated STAT1 (PIAS1) via co-IP and have demonstrated that these interactions are not affected by the APS-1 T16M, L29P, L93R mutations.

We then analysed the G228W mutation in the SAND domain. This is the only known APS-1 mutation with a dominant negative mode of inheritance. We therefore measured the effect of G228W mutation on the interaction of SAND with its partner, the nuclear phosphoprotein 63 (P63). The SAND-P252L

mutation was used as a control as it is transmitted under a recessive mode of inheritance. We showed that the SAND mutant, G228W, has stronger interaction with P63 than the wild-type(wt)SAND, suggesting that SAND G228W is strong competitor of the wt-SAND for P63, and thereby affects AIRE function. In contrast, the mutant SAND-P252L and the wt-SAND both seem to interact with P63 to the same degree.

To further characterise these interactions, we measured the effect of SAND-G228W on P63 interaction using the mammalian-2-hybrid system. The interaction of wt-SAND showed a significantly higher transcription activity than mutant G228W, suggesting that the mutated form was hampering AIRE function, and hence, reducing its transcriptional activation. This result is in keeping with the theory of dominant negative inheritance effect, by which the strong affinity of G228W protein to P63 impairs the formation of the functional complexes required for transactivation. This justifies the mild phenotype that we can see in APS-1 patients bearing the G228W mutant, as it competes but does not completely prevent the wild type from interacting with P63. Moreover, when the functional activity of the interaction of p63 with AIRE was examined, a significantly higher interaction was observed between AIRE and p63 than between P63 and the SAND domain on its own, suggesting that other domains of AIRE have a role in strengthening the interaction with the protein p63. Hence, our data fit well with the unique dominant negative mode of inheritance in APS-1.

In conclusion, by understanding how AIRE operates as a multiprotein complex, we will be able to further appreciate the relationship between AIRE gene malfunction and the breakdown of self-tolerance, which promises to help unravel the pathogenesis not only of APS-1 but also of other types of autoimmune disease.

ACKNOWLEDGMENTS

I would like to thank Dr. Rachid for his constant encouragement, support, patience and advice. His impact has allowed me to develop as a researcher and his contribution to my project and thesis was significant.

I would like to specifically thank my tutor, Dr. Andy Streets for his help, advice and for providing his expertise for most of the experiments. Many thanks also go to Dr. Tom Lovewell for his technical expertise and assistance. I would like to thank my co-supervisor Martin Nicklin for his generous ideas and support.

I would also like to thank my Mother, Father, and brothers for their continuing encouragement and support throughout my education and career. I would like to thank my family for their love, support and companionship, particularly my wife withoutI would never have completed this journey. I wish to express gratitude to my extended family who continued to inspire me significantly. I also extend my gratitude to my friends in Kuwait whose enduring support was extraordinary.

Thank you also to every member of the Rachid Lab, past and present, for offering endless support. I would also like to take this opportunity to thank Yvonne Stephenson, Mark Ariaans, and Jon Kilby who have managed to order me lab materials. To all the students of the IICD, thank you for all your help and support.

My final acknowledgements go to my funding bodies, a scholarship funded by the Civil Service Commission and managed by the Ministry of Health, Kuwait.

"All life is an experiment. The more experiments you make the better."

Ralph Waldo Emerson

LIST OF CONTENTS

Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Aba	Aureobasidin A
AD	Activation domain
-Ade	Adenine deficient
AD	Addison's disease
AF	Adrenocortical failure
AIRE	Autoimmune regulator
ALPS	Autoimmune lymphoproliferative syndrome
Amp	Ampicillin
APC	Antigen presenting cell
APS-1	Autoimmune polyendoicrnopathy syndrome Type I
ATF7ip	Activating Transciption Factor 7- interacting protein
ATP1B1	ATPase Sodium/potassium transporting subunit beta 1
BCR	B cell receptor
BRD4	Bromodomain-containing protein 4
CARD	Caspase recruitment domain
CBP	Cyclic AMP response element-binding protein
CD40L	p450
CHRNA1	Acetylcholine receptor gene
CLP	Common lymphoid progenitor

CMC	Chronic mucocutaneous candidiasis
Co-IP	Coimmunoprecipitation
COS-1	African green monkey kidney cells
cTEC	Cortical thymic epithelial cell
CTLA-4	Ctotoxic T-lymphocyte-associated protein 4
CYP17A1	Steroid 17α-monooxygenase
CYP21A1	Steroid 21-monooxygenase
DAMPS	Danger-associated pattern
DAXX	Death domain-associated protein 6
DBD	DNA-binding domain
DC	Dendritic cells
-DDO	Leucine and tryptophan deficient
Direct-IP	Direct immunoprecepitation
DMEM	Dulbecco's Minimum Essential Medium
DNA	Deoxyribonucleic acid
DNA-PK	DNA-dependent protein kinase
FEZF2	FEZ Family Zinc Finger 2
FOxN1	Forkhead box N1
FOxP3	Forkhead box P3
$Foxp3^{sf}$	Scurfy mutation of the FoxP3 gene
G-1-P	Glucose 1-phosphate

G-6-P	Glucose 6-phosphate
GAD65	Glutamic acid decarboxylase 65
H3K4	Histone 3 lysine 4
H /K ATPase	Gastric hydrogen potassium ATPase
HIPK2	Homeodomain-interacting protein kinase 2
-His	Histidine deficient
HLA	Human leukocyte antigen
HP	Hypoparathyroidism
HRP	Horseradish peroxidase
IFN-α	Human type I interferons(alpha)
IFN-ω	Human type I interferons (omega)
IL-17	Interleukin-17
IL-22	Interleukin-22
IPEX	Immune-dysregulation polyendocrinopathy enteropathy X-linked
IRAK1	Interleukin 1 Receptor Associated Kinase 1
IRBP	Interphotoreceptor retinoid-binding protein
kd	Equilibrium dissociation constant
LAR II	Luciferase assay reagent II
LB	Lysogeny broth
LDB3	LIM-domain binding protein 3
-Leu	Leucine deficient

LiAc	Lithium acetate
LXXL	L is leucine and × is any amino acid
Kan	Kanamycin
M2H	Mammalin-2-hybrid
MBP	Myelin basic protein
MHC II	Major histocompatibility complex class II
mRNA	Messenger RNA
MS	Multiple Sclerosis
mTECS	Medullary thymic epithelial cells
NALP5	NACHT-, LRR- and PYD-containing protein 5
(NH2)	Amine group
NHS ester	N-hydroxysuccinimide ester
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NK	Natural killer
NLS	Nuclear localisation signal
NPC2	Niemann-Pick type C2
P450c17	17α-hydroxylase/17,20-lyase
P450cSCC	Cholesterol side-chain cleavage enzyme
PAMPS	Pathogen associated molecular pattern
PBS	Phosphate-buffered saline
PCNA	Proliferating cell nuclear antigen

PGM1	Phosphoglucomutase 1
PHD	Plant homeodomain
PIAS1	Protein inhibitor of activated STAT 1
PPI	Protein-protein interaction
PRR	Proline rich region
PRR	Pattern recognition receptor
P-TEFb	Positive transcription elongation factor b
RA	Rheumatoid Arthritis
RANK	Receptor activator of nuclear factor K
RelB	Transcription factor RelB
RING	Really interesting new gene
RNA pol 2	RNA polymerase 2
RNF-2	RING finger-2
ROS	Reactive oxygen species
PGM1	Phosphoglucomutase 1
PNGase F	Peptide:N-glycosidase F
-QDO	Adenine, histidine, tryptophan and leucine deficient
SAND	Sp100, AIRE-1, NucP41/75, DEAF-1
SD	Synthetically defined
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SLE	Systemic lupus erythematosus

STX8	Syntaxin 8
TAE	Tris-acetic acid-EDTA
TBS	Tris-buffered saline (TBS)
TE/LiAC	TE buffer/lithium acetate
TCR	T cell receptor
Tconv	Conventional T cels
TFIIB	Transcription factor IIB
TGFβ	Transforming growth factor beta
Th1	T helper type 1
Th2	T helper type 2
TRAF6	Tumor necrosis factor receptor (TNFR)-associated factor 6
Treg	Regulatory T cells
TNF-α	Tumor necrosis factor α
TOP2A	Topoisomerase IIA
-Trp	Tryptophan deficient
X-α-gal	5-Bromo-4-chloro-3-indolyl-α-D-galactopyranoside
XCL1	X-C motif chemokine ligand 1
Y2H	Yeast-2-Hybrid
YPD	Yeast Extract Peptone Dextrose
YPDA	Yeast Extract Peptone Dextrose + Adenine

Chapter 1
Introduction

1. Introduction

1.1 Immunological tolerance

The mammalian immune system has evolved to give protection against a hugely varied range of pathogens and is divided, according to its nature of response, into adaptive and innate. The innate response is the first line of defence that occurs when a foreign entity is encountered (Delves and Roitt, 2000), consisting of cellular components as well as physical and chemical barriers. Initially, mucosal and skin surfaces produce anti-microbial agents and a protective shield. If these barriers are disturbed, however, the cellular components such as basophils, neutrophils, eosinophils, natural killer (NK) cells, macrophages, monocytes and dendritic cells (DCs), are recruited to the affected area to limit the entry of pathogens. The recruitment of these cellular components occurs through the expression of pattern recognition receptors (PRR). These receptors either detect conserved pathogen-associated molecular patterns (PAMPS) or danger-associated molecular patterns (DAMPS), both indicators of cellular damage. PRRs have numerous roles, ranging from their opsonisation of invading particles, to their function in the stimulation of complement pathways, cellular death and inflammation (Akira et al., 2006, Iwasaki and Medzhitov, 2010, Vivier and Medzhitov, 2016). After infection, phagocytes including the neutrophils, macrophages, and monocytes, are the first recruited cellular components. These cells use their PRRs to detect signals that orchestrate with cytokines and chemokines to help attract them to the infected tissues. After entry, neutrophils adhere in aggregates to separate the local site and the pathogen (Lämmermann et al., 2013, Lämmermann, 2016). Furthermore, monocytes are attracted to the infected tissue, where they

mature into either DCs or macrophages. Neutrophils and macrophages share known mechanisms, such as the production of cytokines and the destruction of pathogens, either by phagocytosis or via the production of reactive oxygen species (ROS) or lytic granules (Segal, 2004, Lämmermann, 2016). Triggering of DCs at infection sites leads to the production of cytokines, which, together with those generated by neutrophils and macrophages, functions to minimise the spread of pathogens and produce cells of the adaptive immune system (Moretta, 2002, Chiesa et al., 2005).

Moreover, DC activation encourages their maturation into antigen presenting cells (APCs) that can prime T lymphocytes (T cells); therefore, allowing them to be central correspondents between innate and adaptive immunity. Nevertheless, despite the cells of the innate immune system being protective against numerous foreign substances, they are limited due to the nature of their germline-encoded and non-clonal distribution of receptors. This, in turn, minimises their pathogen recognition and thus limits their ability to protect against an evolving pathogen.

In contrast, the adaptive response is initiated to a specific pathogen that causes an immunological memory. Different cell types that evolve from such a response stem from a pluripotent lymphoid progenitor; namely, B cells and T cells. Each cell retains unique receptors for a particular antigen. Their unique antigen-specificity is a consequence of their generation of highly diverse and clonally distributed receptors: the B cell receptor (BCR) and T cell receptor (TCR) respectively. These antigen receptors, however, are not associated with germline-encoded sequences (Bonilla and Oettgen, 2010) but are produced somatically or through gene shuffling processes that occur during their

activation and development. Even though B and T cells are both stemmed from hematopoietic stem cells in the bone marrow, their later development appears in different locations (Kondo et al., 1997). The development of B cells initiates and continues in the bone marrow, whereas the development of T cell requires common lymphoid progenitor (CLP) cells to leave the bone marrow and travel to the thymus. The remaining sections of this chapter focus on T cells due to the focus of this thesis being on *AIRE* gene expression.

Naive T cells follow a series of events as soon as TCR is expressed on the surface of the cell. Primarily, the antigen recognition in TCR is tested against molecules in the thymus (Goodnow et al., 2005). Sequentially, self-antigens are presented to the major histocompatibility complex class II (MHCII). Only lymphocytes that adhere with low affinity to self-antigens via their receptors escape apoptosis. This process is known as positive selection (Vrisekoop et al., 2014). Positive selection takes place in the cortex and is distinguished by a distinctive subset of APCs known as cortical thymic epithelial cells (cTECs). Throughout positive selection, the vast majority of the thymocytes – around 90% – bind with weak affinity to cTECs, and hence fail to reach TCR signaling requests. They are consequently eliminated by neglect (Klein et al., 2014). Developing T cells that interact with cTECs, and hence reach the TCR signaling requirement, are positively selected. These then travel to the medulla where negative selection events operate. In the medulla, thymocytes come across a unique and more heterogeneous subset of APCs, including the medullary thymic epithelial cells (mTEC) and DCs. Throughout negative selection, thymocytes that bind with high affinity to self-antigens are removed by clonal deletion; a process that includes cell death via apoptosis. It should be noted

that negative selection is crucial in overcoming self-reacting lymphocytes from damaging their own tissue (Žumer et al., 2013). Besides clonal deletion, very high TCR signalling may also result in thymic differentiation into regulatory T cells (Tregs) (Fu et al., 2014). Tregs are discussed in detail in Section 1.3.5.

There are also other molecules that can enhance the perceived TCR signal strength upon thymic selection. For instance, extremely high signalling can be adjusted either by the up-regulation of negative modulators of TCR signalling, such as CD5, (Tarakhovsky et al., 1995, Azzam et al., 1998); (Mandl et al., 2013);(Persaud et al., 2014, Mujal and Krummel, 2015) or through the down-regulation of positive regulators of signalling such as CD4/CD8 (Fu et al., 2014).

1.1.2 Autoimmunity and autoimmune disease

Autoimmune disease was re-defined by Rose and Bona in 1993 to involve the following three classifications: 1) direct evidence that an infection can be reproduced in a healthy individual or animal by the transmission of autoantibodies – the immunoglobulins that are generated in response to an autoantigen – or antigen-specific T cells; 2) indirect evidence that an autoimmune disease can be triggered by immunisation with a target autoantigen or by taking autoantibodies from an autoimmune animal model; and 3) circumstantial indirect evidence of disease where autoantibodies are taken from the major target tissue of the individuals infected by disease (Witebsky et al., 1957).

In clinical terms, autoimmune diseases are divided into two major classifications: organ-specific, and non-organ-specific or systemic autoimmune disease. Organ-specific diseases can be initiated by direct cell damage as with type-1 diabetes, where an autoimmune attack is trafficked towards the β-cells

of the pancreas, leading to a decrease in insulin levels and thus high blood glucose levels (Lernmark and Larsson, 2013). Organ-specific disease can also be triggered by obstructing autoantibodies, as with Myasthenia Gravis, where antibodies hinder the acetylcholine receptors that are situated on the muscles, blocking the binding of acetylcholine and resulting in weakened skeletal muscles (Fridkis-Hareli, 2008, Atassi and Casali, 2008). Hence, cells destruction can occur either directly by T cells or via prolonged autoantibody reactions through T cells (Fujio et al., 2012). Destruction can also occur via both pathways (Betterle et al., 2002, Atassi and Casali, 2008). On the other hand, systemic autoimmune disease denotes when disease is spread across the entire body, such as with systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), and scleroderma. In these cases, the autoimmune response targets many organs and antigens as a result of self-reactive B and self-reactive T cells (Kindt et al., 2007); (Dzhagalov et al., 2013).

1.1.3 Etiology of autoimmune disease

The understanding of autoimmune diseases has evolved over the past century. That said, the mechanism leading to autoimmune disease is extremely complex. Various genetic factors contribute to the consequent vulnerability in most autoimmune diseases. Normally, grouping of numerous diseases and co-association of autoimmune disease is observed in families. Furthermore, the rate of concordance in autoimmune disease is greater in monozygotic twins than in dizygotic twins (Klein and Sato, 2000). Most autoimmune diseases are polygenic, whereby numerous susceptibility genes act synergistically in the generation of autoimmunity. Most autoimmune diseases are also linked to human leukocyte antigens (HLA). The highest correlation is observed with the

MHC II, However, some HLA alleles can enhance protection from autoimmune disease even when other susceptibility genes exist (Davidson and Diamond, 2001). HLA is reviewed in detail in Section 1.2.1.3. Moreover, there are susceptibility genes that function on immune cells and thereby regulate the overall immune reaction such as in the polymorphism of the CTLA-4 gene. This is crucial for down-regulation of T cell activation, which is linked with type-1 diabetes and celiac disease (Holopainen et al., 1999); Marrack et al., 2001).

On the other hand, a small number of autoimmune diseases are caused by mutations in a single gene. Thus far, the most categorised monogenic autoimmune diseases are autoimmune polyendocrine syndrome type I (APS I), autoimmune lymphoproliferative syndrome (ALPS) and immune-dysregulation polyendocrinopathy enteropathy X-linked (IPEX) (Davidson and Diamond, 2001). APS-1 is an autosomal recessive disorder caused by mutations in the autosomal regulator (*AIRE*) gene. ALPS is an autosomal-dominant inherited disease where deficiencies in the Fas protein lead to hyperproliferation of lymphocytes due to its incapability to trigger apoptosis (Rieux-Laucat et al., 1995). IPEX develops due to mutations in the Forkhead/winged helix transcription factor (FOXP3) gene: a gene that is responsible for the regulation of Tregs (Bennett et al., 2001). It is crucial to note though that patients developing an identical disease in this group of monogenic autoimmune diseases are likely to exhibit variations in the spectrum and severity of disease. Thus, it is common that relatives with an identical mutation show differences in clinical manifestations (Davidson and Diamond, 2001; Marrack et al., 2001).

Environmental factors and other genes can also contribute to the aetiology of autoimmune diseases. The trigger of many autoimmune diseases has been

linked to infectious pathogens, such as viruses and bacteria, which might occur via alternative mechanisms; for instance, molecular mimicry and the secretion of sequestered antigens (Sfriso et al., 2010, Costenbader et al., 2012). Molecular mimicry occurs when an epitope resembling that antigen causes cross-reaction. For instance, the cross-reaction between cardiac proteins and streptococcal leads to rheumatic fever and cross-reactivity between the protein 2C of the coxsackie B virus and a peptide from the glutamic acid decarboxylase 65 (GAD65) and results in the initiation of type-1 diabetes (Kukreja and Maclaren, 2000; Cunningham, 2019). Antigens that are secured from immune recognition, as in the eye (immune privileged locations), can be generated throughout organ destruction and can encounter T cells, leading to re-stimulation and a resultant strong autoimmune response (Kindt et al., 2007, Caspi, 2008). Furthermore, stochastic events, somatic mutations, T and B cell receptor recombination, and the level of cell death may explain the variabilities in disease spectrum among monozygotic twins bearing an identical autoimmune defect (Germain, 2001; Mackay, 2005).

There are also non-genetic causes that can lead to autoimmunity, such as the stress of diseases, which can cause lymphopenia – a reduction in the number of lymphocytes. To balance this situation, the immune system switches on a homeostatic proliferation of cells, resulting in skewed lymphocyte repertoire. This would, for instance, support the development of autoreactive T cells (Ernst et al., 1999);(Sprent and Surh, 2002). Added to this, the susceptibility of disease between males and females varies in autoimmune diseases. To be more specific, the frequency of SLE occurrence is approximately 71-90% greater in women than in men. Both physiological and genetic variations also have an

impact on this predisposition. In SLE, for example, X-linked genes such as the Interleukin 1 Receptor Associated Kinase 1 (IRAK1) and CD40 ligand (CD40L) are overexpressed in women as a result of incomplete X inactivation (Tiniakou et al., 2013). Finally, female hormones can also affect the susceptibility of disease, as oestrogen supports the persistence of auto-reactive T cells (Pennell et al., 2012). The danger model is also another recognition model in the immune system, where APC is triggered only via danger signals from infected cells, and hence allows the body to distinguish between normal and harmful cells (Matzinger, 1994); (Matzinger, 2002).

1.1.4 Autoantigens and autoantibodies

The autoantigens acknowledged by the immune system to cause damage to tissues and organs are only identified for a small portion of all autoimmune diseases. This is due to the variations in B and T cell responses, as diseases develop such that the reaction relies on another autoantigen than the targeted one. This is known as epitope spreading (Davidson and Diamond, 2001). In addition, the autoimmune destruction does not always occur at all of the autoantigen sites since both autoantibodies and lymphocytes participate in destruction (Lernmark, 2001). Amongst the autoimmune diseases for which the autoantigen is identified and linked with disease are these that target the neuro-muscular junction, such as in the voltage-gated calcium channels in Isaac's disease. The autoantigen glutamic acid decarboxylase (GAD) is a well-known antigen that is highly correlated with myelin basic protein (MBP) in multiple sclerosis (MS) and fibrinogen in RA and type-1 diabetes (Hayter and Cook, 2012, Lernmark, 2001).

Although autoantibodies are targeted towards the self, they class switch from IgM to IgG. This is because the concentration of antigen changes as the cells differentiate, demonstrating an elevated specificity. They hence experience affinity maturation as with normal antibodies (Plotz, 2003, Omori et al., 2006). Autoantibodies in Grave's disease, for example, affect the thyroid's cell physiology such that they do not react with desmoglein I in *Pemphigus Vulgaris* (PV). This triggers proteases, causing blistering (Davidson and Diamond, 2001). In-spite of the complications of locating the exact autoantibodies to the correct autoantigen, and thereby the disease, autoantibodies have been proven to be a vital tool for the prediction and characterisation of some autoimmune diseases. For example, autoantibodies for type-1 diabetes and thyroiditis can be identified well in advance of tissue destruction occurring. Atrophic gastritis is another disease that may have different causes, yet the autoantibodies' detection of H/K ATPase defines the manifestation as autoimmune gastritis (Lernmark and Larsson, 2013).

1.2 Autoimmune polyendocrine syndromes

Autoimmune polyendocrine syndromes are a group of rare autoimmune diseases with endocrine signs. The syndromes are classified into four groups (APS I-IV) (Eisenbarth and Gottlieb, 2004, Neufeld et al., 1981). Autoimmune polyendocrine syndromes (APS) are also a group of rare autoimmune diseases (Eisenbarth & Gottlieb 2004). The autoimmune polyendocrine syndromes are APS-1, -2, -3 and -4 (Brun, 1982); (Obermayer-Straub and Manns, 1998);(Betterle et al., 2002); Betterle & Zanchetta, 2003). They are associated with each other as well as with other organ-specific autoimmune disorders (Neufeld et al., 1981); (Riley, 1992). APS-1 and APS-2 have a common clinical

phenotype feature yet vary in their occurrence, inheritance, and onset (Eisenbarth and Gottlieb, 2004). APS-2 is a polygenic disease that is a mixture of Addison's disease with autoimmune thyroid disease and/or type-1 diabetes. Individuals do not exhibit candidiasis or hypoparathyroidism, and the onset is at adulthood with a complex aetiology (Bennett et al., 1964); Obermayer-Straub & Manns, 1998; (Michels and Eisenbarth, 2009, Obermayer-Straub and Manns, 1998). In addition, APS-2 and isolated Addison's disease show strong association with the MHC locus (HLA-DR3 and DR4 haplotypes), and both diseases are more common in females (Lebovitz, 2013). APS-3 is a heterogenous disease with autoimmune thyroid disease and at least one other autoimmune disorder, excluding Addison's disease (Neufeld et al., 1981, Obermayer-Straub and Manns, 1998). Finally, APS-4 is a group of all the clinical associations not included in the other APS subtypes, such as the presence of type-1 diabetes with hypogonidism, coeliac disease and chronic gastritis (Betterle & Zanchetta, 2003). Of the four, only APS-1 is a monogenic disorder and is further reviewed in the following section.

1.2.1 Autoimmune Polyendocrine Syndrome type 1 (APS-1) disease

1.2.1.1 Clinical phenotype of APS-1

Autoimmune Polyendocrine Syndrome type-1 (APS-1) is an inherited autosomal disorder caused by mutation of the *AIRE* gene. The most common clinical features of the disease are primary adrenal insufficiency (Addison's disease), hypoparathyroidism (HP), and chronic mucocutaneous candidiasis (CMC). APS-1 seems to occur worldwide but is more dominant in genetically-isolated populations such as in Finns, Sardinians, and Iranian Jews (Bratanic et al., 2015, Perheentupa, 2006). A minimum of two of the clinical manifestations should arise to confirm the diagnosis of APS-1 (Podkrajšek et

al., 2008), and diagnosis of APS-1 via DNA analysis is reinforced via the identification of the mutation. Interestingly, a study was carried out with Italian patients stated that the mean diagnostic delay was ten years from symptom initiation until genetically-confirmed APS-1 (Mazza et al., 2011). In addition, it has been mentioned that the existence of auto-antibodies initiated against interferons of type-1 could be utilised as a molecular diagnostic tool to confirm APS-1 (Meloni et al., 2008).

Less common clinical features include ectodermal dystrophy and endocrine deficiencies. Furthermore, autoimmune respiratory illnesses are correlated with APS-1 disease (De Luca et al., 2008), and ectodermal dystrophy occurs as nail dystrophy, vitiligo, alopecia, and enamel hypoplasia (Kisand et al., 2011). The components of ectodermal dystrophy that arise in APS-1, such as enamel and nail dystrophy, are the result of chronic infections (Collins et al., 2006, Bratland et al., 2013). Intestinal malabsorption is another manifestation that occurs in the gastrointestinal tract (Sonal et al., 2012). Other autoimmune symptoms involve pernicious anaemia, which is more frequent in adults. Finally, it has been noted that oesophageal and oral carcinoma are more prevalent in adults due to chronic inflammation in the mucosa as a result of *Candida* infection rather than lack of AIRE (Rautemaa et al., 2007);(Böckle et al., 2010) . All the above-mentioned clinical features that occur in APS-1 patients are summarised based on their type in *Table 1*.

Disease component	Prevalence	Age of Onset	References
Nonendocrine			
Enamel hypoplasia	71-79%	Childhood to puberty	(Kisand et al., 2011)
Nail Dystrophy	52%		(Collins et al., 2006) (Kisand et al., 2011)
Alopecia	14-38%	Puberty to Adulthood	(Perheentupa, 2006) (Wolff et al., 2007)
Vitiligo	5-14%	Puberty to Adulthood	(Kisand et al., 2011) (Mazza et al., 2011)
Intestinal malabsorption	21%	Childhood to Adulthood	(Podkrajsek et al., 2008)
Autoimmune hepatitis	11-20%	Puberty to Adulthood	(Perheentupa, 2006)
Pernicious Anaemia	10-17%	Adulthood	(Bruserud et al., 2016)
Keratitis	14%	Childhood to Puberty	(Perheentupa, 2006)
Oesophageal and oral carcinoma	11%	Adulthood	(Bockle et al., 2010) (Rautemaa et al., 2007)
Arthralgia with Fever	13%	Childhood	(Podkrajsek et al., 2008)
Endocrine			
Disease component	Prevalence	Age of Onset	References
Hypoparathyroidism	83-95%	Infancy to childhood	(Kisand et al., 2011) (Meloni et al., 2012)
Mucotaneous Candiasis	76-100%	Infancy to childhood	(Perheentupa, 2006) (Kisand et al., 2011) (Meloni et al., 2012)
Gonadal failure	13-36%	Puberty to Adulthood	(Kisand et al., 2011)
Type 1 diabetes	12%	Adulthood	(Proust-Lemoine et al., 2010)
Hypothyroidism	4%	Childhood	(Proust-Lemoine et al., 2010)
Addison disease	61-72%	Childhood	(Meloni et al., 2012) (Perheentupa, 2006)

Table 1. Non-endocrine and endocrine clinical manifestations that occur in APS-1 patients

1.2.1.2 The impact of autoantibodies on APS-1

Deficiencies in self-tolerance for patients diagnosed with APS-1 result in the secretion of antibodies targeted to self-antigens (Eisenbarth and Gottlieb, 2004). Importantly, the identification of autoantibodies in APS-1 is a crucial marker for identifying autoimmune clinical features, and therefore enhances diagnosis. Most autoantibodies are tissue-specific and are directed to intracellular enzymes, such as GAD65, which is located in the pancreatic islet (Klemetti et al., 2000). NACHT-, LRR- and PYD-containing protein 5 (NALP5) is another autoantigen located in the parathyroid, and 49% of APS-1 patients that express hypoparathyroid symptoms exhibit autoantibodies to NALP5 (Alimohammadi et al., 2008). Levels of autoantibodies in APS-1 vary according to different autoimmune diseases. In type 1-diabetes patients, for example, they are low (Eisenbarth and Gottlieb, 2004). On the other hand, APS-1 patients with thymomas have a relatively increased titre of immunoglobulin G autoantibodies, neutralising type-1 interferon-alpha (IFN-α) and interferon-omega (IFN-ω) (Meager et al., 2006;Kisand et al., 2008). In CMC, autoantibodies targeted to the cytokines; interleukin-22(IL-22) and interleukin (IL-17) initiate the disease in APS-1 patients. This is because large amounts of autoantibodies targeted against IL-22 and IL-17 exist in patient serums (Puel et al., 2010, Liang et al., 2006). This result reinforces the fact that patients diagnosed with APS-1 secrete more IL-17 than healthy individuals in response to *Candida* (Ahlgren et al., 2011) Another study discovered dominant and autosomal recessive mutations in the IL-17 gene in CMC individuals not diagnosed with APS-1 (Puel et al., 2011). This highlights the fact that IL-17 is not necessarily linked with APS-1. Another interesting finding is that the secretion of IL-22 cytokine is lower in individuals with CMC compared to normal individuals, even when the former

have been diagnosed with APS-1 (Kisand et al., 2011) Moreover, gonadal deficiency is highly correlated with autoantibodies targeted to the antigens 17α-hydroxylase/17,20-lyase (P450c17) and the cholesterol side-chain cleavage enzyme (P450cSCC) (Peterson and Peltonen, 2005). *Table 2* below presents a summary of the autoantibodies in APS-1 patients and their association with different clinical manifestations.

Clinical Manifestation	Targeted Antigen	Prevalence %	No. of affected patients	Reference
Adrenal Insufficiency	Cytochrome P450 21 Cytochrome P450 SCC	66 52	7/11 47/90	(Peterson & Peltonen, 2005) (Söderberg et al., 2004) (Wolff et al., 2007)
Mucotaneous Candiasis	IL-17 IL-22	47 47	33/70 33/70	Puel et al 2010
Autoimmune diabetes	Glutamate decarboxylase 65 Tyrosine phosphatase like protein IA-2	37 7	33/90 6/90	(Klemetti et al., 2000) (Wolff et al., 2007) (Söderbergh et al., 2004) (Meloni et al., 2012) (Proust-Lemoine et al., 2010)
Hypogonadism	Cytochrome P450 17 Cytochrome P450 SCC	44 43	5/11 12/29	(Söderbergh et al., 2004) (Wolff et al., 2007)
Alopecia	Tyrosine Hydroxylase	40	36/90	(Kisand et al., 2011) (Heino et al., 2000) (Söderbergh et al., 2004) (Meager et al., 2006)
Intestinal dysfunction	Tryptophan Hydroxylase	75	15/20	(Söderbergh et al., 2004) (Pra et al., 2004)
Respiratory Symptoms	Potassium Channel Regulator	7	7/102	(Alimohammadi et al., 2009)
Hypoparathyroidism	NACHTleucine-rich-repeat protein 5	49	36/87	(Alimohammadi et al., 2008)

Table 2. The association of APS-1 autoantibodies to specific disease components

1.2.1.3 HLA/APS-1 association

As previously mentioned, human leukocyte antigen (HLA) genes are by far the most reliable genetic contribution to autoimmune disease, with almost every specific HLA antigen being associated with a certain autoimmune disease (Gough and Simmonds, 2007). Many MHC molecules can bind and hence display different peptides. For instance, the adaptive immune systems of different humans may identify and react against different antigens, albeit from an identical pathogen (van Lith et al., 2010). The MHC consists of three main loci encoding MHC classical I genes (A, B and C) and three types of genes that encode for MHC class II (DR, DP, DQ). There are also a number of MHC II chaperone genes (DOβ, DOα and DM) that are crucial in selection of epitope for attachment of peptide to MHC II (van Lith et al., 2010). Moreover, the number of loci allows for diversity among humans, and thus MHC is extremely polymorphic. Therefore, it is unlikely for unrelated humans to have an identical HLA genotype. The heterogeneity is, therefore, vital for humans and for the health of the species in terms of combating and resisting infections (Sommer, 2005). In-spite of the manner in which HLA diversity is preserved, HLA genes have a substantial effect on influencing immune responses. The antigens that are known to trigger autoimmunity are quite restricted. They are displayed rarely at immune-privileged sites or T cells and are eliminated during thymic development. Specific MHC molecules may also adhere weakly to an autoimmune-associated antigens, resulting in positive selection (Gough and Simmonds, 2007). Furthermore, specific MHC molecules can have a weak ability to generate Tregs through thymic selection, hence reducing the potency for peripheral tolerance mechanisms (Gough and Simmonds, 2007). The T cells in the periphery also interact with a large number of self-antigens, and the

content of the HLA-bound antigen pool presented via APC is essential in preventing autoimmunity. Thus, HLA shown via the thymus delineate the HLA-associated antigen repertoire that the T cells may encounter at the periphery that includes peptides originating from peripheral tissue antigen (PTA) (Bunnell et al., 2002). In terms of the correlation between HLA and APS-1, the phenotypic variation among APS-1 patients is difficult to explain by variations in the allele of the *AIRE* gene, despite the emergence of various APS-1 causing mutations.

Variations in phenotypes of APS-1 Finnish patients who have homozygous mutation as well as phenotypic variations among siblings reflect the existence of other environmental and/or genetic factors that affect the disease phenotype. Indeed, the correlation of the genotype-phenotype can only be considered reliable after long-term monitoring of APS-1 patients and by separating them based on their HLA haplotypes or via any other genetic causes (Gylling et al., 2000). It has also been noted that some APS-1 diseases are associated with certain haplotypes of HLA. For instance, a research group established that HLA DRB1*03 is correlated with Addison's disease, and DRB1*04 -DQB1*0302 is associated with alopecia areata. HLA-DRB1*15DQB1*0602, however, is negatively-associated with type-1 diabetes (Halonen et al., 2002). Another research group established an identical protective effect for type-1 diabetes in the DRB1*0602 haplotype in Finnish APS-1 patients (Gylling et al., 2000). Furthermore, patients that do not have the R257X allele exhibit a low rate of mucocutaneous candidiasis (Halonen et al., 2002). Moreover, patients with a missense mutation of G228W that acts in a dominant manner show an increased rate of thyroiditis (Cetani et al., 2001). Another related finding is that

the APS-1 phenotype in Iranian Jews is milder than in Finnish populations. The former commonly manifest hypoparathyroidism yet are rarely diagnosed with ectodermal dystrophy and mucocutaneous candidiasis (Zlotogora and Shapiro, 1992). Y85C, the most frequent mutation associated with Iranian Jews, does not have an impact on the subcellular localisation of the AIRE protein and hence helps preserve AIRE's transcriptional activity. Hence, as the AIRE protein is only slightly affected, this might justify the milder phenotype exhibited by Iranian Jews (Björses et al., 2000). As with type 1A diabetes, Addison's disease can also be categorised into phases of disease progression. Individuals that are predisposed genetically exhibit autoantibodies against the enzyme 21-hydroxylase and ultimately lose the ability to generate cortisol. As discussed earlier, susceptibility is deliberated via the genes that encode MHC class II, and as with type 1A diabetes, there is a high correlation with the DR3 haplotype (Michels and Eisenbarth, 2009).

1.2.1.4 Mouse model for APS-1

APS-1 is known by the complex interaction of immune cells within the thymus. In order to generate an experimental model of APS-1, the mouse homologue Aire was cloned and characterised (Ramsey, 2002). In fact, experiments on vivo mouse models have significantly enhanced the understanding of the molecular mechanisms that lead to autoimmunity in APS-1 patients (Pereira et al., 2005). Genomic sequencing indicated that the 545 amino acid human AIRE proteins are 71% identical with the 552 amino acid mouse Aire proteins, signifying that both proteins are conserved between the species (Blechschmidt et al., 1999, Mittaz et al., 1999).

It has been reported that the background strain affects the nature of autoimmunity in Aire-deficient mice (Jiang et al., 2005). Numerous Aire-deficient mice on dissimilar genetic backgrounds have been studied, and the six strains identified by five different research groups are summarised in *Table 3* (Anderson et al., 2002); (Ramsey, 2002, Jiang et al., 2005, Kuroda et al., 2005); (Hubert et al., 2009). Two groups generated an Aire knockout (KO) mouse on a C57BL/6 background. Ramsey's group represented a human equivalent mutation, R257X, in this mouse strain to mimic human APS-1 disease (Ramsey, 2002). The insertion of the Neo-cassette in exon 6 resulted in the early termination of Aire polypeptides in the SAND domain. Anderson's group, however, generated Aire KO mice via omitting exon 2 (Anderson et al., 2002). Both mice exhibited normal Aire gene transcript levels, yet sequence analysis confirmed that these transcripts do not result in functional protein due to the frame shift deletions in both Aire copies. Kuroda's group developed another Aire KO mouse by generating a gene-targeting vector via deleting

exons 5 to 12 by inserting a Neo cassette (Kuroda et al., 2005). This led to the deletion of crucial functional Aire domains, including SAND, PHD-1 and PHD-2. Hubert's group generated an Aire KO mouse with a disrupted PHD-1 domain at exon 8, reflecting the 13-bp deletion of human APS-1 patients (Hubert et al 2009). Niki's group had demonstrated that abrogation of Aire in NOD mice modulated the autoimmune phenotypes of NOD mice, suggesting that Aire and/or Aire+ cells in NOD mice are functionally competent (Niki, 2006).

All Aire KO mice had normal morphology in their thymic compartments, suggesting that absence of Aire does not affect thymic morphology. The expression of Aire protein, however, was aborted, either due to lack of expression or a truncated expression due to premature Aire (Anderson et al., 2002);(Ramsey, 2002, Kuroda et al., 2005); (Hubert et al., 2009). In addition, all Aire KO mouse strains lived to the expected age as their wild-type littermates, apart from the NOD mice, for whom the disease developed more rapidly, thus only enabling them to survive for up to 15 weeks compared to approximately 2 years for mice in captivity (Jiang et al., 2005). Furthermore, apart from the mice engineered by Kuroda's group, these groups reported spontaneous infiltration to multiple tissues, production of autoantibodies and infertility that accelerated with age. To be more specific, these mice were fertile, yet generated progenies only occasionally (Kuroda et al 2005).

Autoimmunity targeting multiple organs is an effect of the human APS-1 phenotype, yet these lymphocytic infiltrations never progress into any of the diagnostic classical triad: Chronic Mucocutaneous Candidiasis, Hypoparathyroidism and Addison's disease. It is well known that APS-1 has a stochastic element; hence, a difference in phenotypes among human twins (Li

et al., 2017). In addition, this might be due to the specificity of PTA, which is only regulated by human AIRE, such as the steroid 17α-monooxygenase (Cyp17a1) and the steroid 21-monooxygenase (Cyp21a1), in the adrenal cortex and not via Aire in mTEC of mice. This would explain the low pathological destruction caused by multiple antibodies and self-reactive T cells. Moreover, the capability of autoantigens to be taken up by APC for MHC binding limits the array of potential targets. Thus, MHC plays a major role in controlling the target-organ specificity (Kuroda et al., 2005). This suggests that constraints on MHC specificity, peripheral tolerance, and antigen handling can control the damage resulting from autoimmunity (Jiang et al., 2005). Epigenetics and differences in environment may also contribute to the phenotypic variations between humans and mice (Li et al., 2017).

To further investigate the effect of genetic variations on the phenotypic differences between human APS-1 patients, Jiang and colleagues backcrossed C57BL/6J, NOD/LtJ, BALB/cJ, and SJL/J backgrounds mice to Aire-deficient mice engineered by (Anderson et al., 2002, Jiang et al., 2005) [see *Table 3*]. Aire-deficient mice displayed lymphocytic infiltrates and autoimmune endocrine destruction with minor differences in the pathology generated between the different strains. Moreover, their study established that differences in genetic background have an impact on the overall disease intensity, which varies from mild as in the C57BL/6J background to more severe in the *Aire* NOD/LtJ background mouse, where it developed elevated autoantibody production and pneumonitis (Jiang et al., 2005).

However, it is important to highlight that contradictory evidence has caused debate over whether an Aire KO mouse is an ideal candidate for the study of

human APS-1 disease (Pöntynen et al., 2006, Kekäläinen et al., 2007). For instance, a study by (Pöntynen et al., 2006) revealed that an Aire KO mouse failed to generate autoantibodies identical to those in APS-1 patients. These results were observed in the C57BL/6 mouse (Hubert et al 2009). On the other hand, trials by (Kuroda et al., 2005) and (Jiang et al., 2005) on different mice strains showed that the genetic background is highly correlated with disease manifestations. This is possibly because Aire KO mice exhibit the same defects in self-tolerance, yet the clinical manifestations are against diverse targets. Take as an example candidiasis infection, in which there is a tendency for *Candida* to affect subclinical damage areas, potentially leading to the eventual causation of CMC in APS-1 patients. Therefore, the source of APS-1 could be initiated by the development of its clinical manifestations (Guo et al., 2018). As previously mentioned, Aire KO mice do not exhibit candidiasis. This might be due to the monitoring of the mice under aseptic conditions (Ramsey, 2002). These conclusions reflect the potentially vital role of environmental factors, such as acquirement of infections, on the impact on the immune responses in Aire KO mice (Hubert et al 2009). Finally, a recent study has reported that the gene expression profile is not similar in a two genetically identical wild-type(wt) mice whose mTECs were investigated via microarray profiling. This has shown the inter-individual variability in tolerance in mice (Meredith et al., 2015). Another study has described a similar observation from gene expression profiling of wt mTECs from two thymic lobes of the same mouse, in which they exhibited higher variability in Aire-dependent gene expression rather in Aire-independent expression (Venanzi et al., 2008). This diversity might assist in identifying the cause of the variations seen in clinical manifestations among

APS-1 patients. Thus, variations in genetic backgrounds have been observed in mice just as clinical manifestation can vary between APS-1 patients (Ahonen et al., 1990); (Ishii et al., 2000). Consequently, the Aire KO phenotype is categorised by the existence of various autoantibodies against non-endocrine and endocrine organs, as with APS-1 patients. Hence, it is possible to confidently conclude that Aire-deficient mice can be used to mimic the APS-1 phenotype and are a suitable model for APS-1 (Jiang et al 2005). Aside from the mouse models, Lovewell et al. (2018) established a recombinant AIRE expression variant of the human cell line TEC 1A3(TEC 1A3 AIREhi) to study genes that are regulated by AIRE. They were able to identify a large number of AIRE-dependent genes including STAT1. The AIRE-dependent expression in this model with murine datasheets elucidated 447 conserved genes, which demonstrates that only a relatively small number of genes are regulated by AIRE, as this number had previously been thought to be more than a thousand (Lovewell et al., 2018). This suggests a variation among the set of AIRE-regulated genes between mice and humans, meaning it is less inclusive than the mouse models. Using this model, they also observed that STAT1 is not phosphorylated, and that the expression profile of gene expression regulated by AIRE was not affected when TEC 1A3 AIREhi cells were treated with STAT1 phosphorylation inhibitors. Furthermore, TEC 1A3 AIREhi cells treated with STAT1 shRNA did not cause any variation in the expression of the downstream genes of unphosphorylated STAT1. This suggests that these genes were directly regulated via AIRE, not the unphosphorylated STAT1. The advantages of using such a model is that inhibitors and activators can be used to identify specific functions of a mechanism (Lovewell et al., 2018).

Mouse strain	Design	Autoantibodies(%of affected)	Tissue infiltration (%)	Reference
C57BL/6J	Neo cassette insertion on exon 6	Liver (20%) Testis (54%) Pancreas (47%) Adrenal (20%)	Liver (50%)	(Ramsey et al., 2002)
C57BL/6J	Cre-lox mediated deletion exon 2	Eyes (84%) Salivary gland (10%) Ovary (16%)	Eyes (67%) Stomach (9%) Salivary gland (100%) Lung prostate (100%)	(Anderson et al., 2002)
NOD/Shi Jic	Backcrossed with *AIRE*−/− C57BL/6	Pancreas (100%) Stomach (100%) Kidney (100%) Liver (100%)	Pancreas (81%) Liver (92%) Lung (96%)	(Niki et al., 2005)
C57BL/6J	Cre-lox mediated deletion of exon 8	Testis (50%) Lung (75%) Salivary gland (100%) Liver (75%) Eyes (75%) Pancreas (100%)	Eyes(80%) Parotid gland(80%) Subligual gland(40%)	(Hubert et al., 2009)
NOD/LTJ	Backcrossed with *AIRE*−/− C57BL/6	Pancreas (100%) Eyes (100%) Salivary gland (100%) Stomach (67%) Ovary (33%)	Eyes (93%) Stomach (87%) Salivary gland (100%) Lung (100%) Liver (86%) Prostate (100%) Ovary (88%) Pancreas (100%) Thyroid (54%)	(Jiang et al., 2005)
BALB/cJ	Backcrossed with *AIRE* -/- c57BL/6	Eyes (33%) Stomach (100%)	Eyes (80%) Stomach (100%) Salivary gland (20%) Lung (50%) Liver (40%) Prostate (100%) Ovary (100%)	(Jiang et al., 2005)

Table 3: Aire knockout mouse models

1.3 Autoimmune regulator gene (*AIRE*)

1.3.1 Structure and properties of AIRE

The *AIRE* gene in humans is located on chromosome 21q22.3 and has 14 exons covering 11.9 kb of the total genomic DNA, encoding a 57.7 kDa protein (Vogel et al., 2002). AIRE controls a variety of genes in different cells. In mTECs, it is primarily localised in the nucleus, where it is condensed in nuclear bodies proximal to nuclear speckles (Abramson et al., 2010). In addition, it is located in the cytoplasm as well as in the nucleoplasm. The nature of AIRE's nuclear distribution is central to its functional domains, emphasising the role of AIRE in regulating transcription (Halonen et al., 2001).The product of the *AIRE* gene is a nuclear protein that is 545 amino acids in length [see *Figure 1*]. It has five functional domains; namely, the caspase-recruitment domain (CARD), the putative DNA-binding domain (SAND), two plant homeodomain zinc fingers (PHD), four nuclear-receptor binding motifs, and a proline-rich region (PRR) (Perniola and Musco, 2014). AIRE also functions as a chromatin-binding protein (Mathis and Benoist, 2009). Autosomal-recessive missense mutations primarily occur within the CARD domain. The CARD domain exist in pro-apoptotic proteins and is included in the oligomerisation of AIRE (Hofmann et al.). The NLS is distinguished by the existence of importin-α as it binds to importin-β to facilitate the cycle of nuclear translocation (Pemberton and Paschal, 2005). The SAND domain plays a critical role in interacting with ATF7ip (Activating Transcription Factor 7–interacting protein) -associating Aire with further epigenetic processes (Ilmarinen et al., 2005, Waterfield et al., 2014). Mutations in the SAND domain have been observed in a dominant manner and the most frequent phenotype manifested in APS-1 patients is thyroiditis (Cetani et al., 2001). The PHD domain is located within the C-terminus, which is divided by

the PRR. The Zinc fingers, situated in the PHD domain are cysteine-rich domains that act as scanners for the epigenetic marker-specific modifications in histones (Musselman and Kutateladze, 2011). To be more specific, the PHD1 domain identifies the unmethylated lysine 4 H3 histones (H3K4), and therefore stimulates chromatin and orchestrates gene transcription via expressing PTA in mTECs (Org et al., 2008). The PHD2 domain interacts with protein partners that are included in regulating transcription and chromatin binding. Moreover, it maintains the mTEC transcriptome (Yang et al., 2013). Protein-protein interaction occurs in the four LXXLL motifs, which also contributes to *AIRE* gene transcription. The DNA binding activity is localised in the LXXLL motifs (Plevin et al., 2005).

Figure 1: Illustration of the functional domains of human *AIRE* protein. The CARD is responsible for homo-oligomerisation, inducing apoptosis, and the generation of AIRE nuclear bodies. The nuclear NLS is believed to be required for nuclear import. The SAND domain regulates the nuclear localisation and homomultimerisation of AIRE. In addition, it facilitates protein-protein interaction. PHD1 reacts with unmethylated histone H3K4 and shows E3-ubiquitin ligase activity. It has four LXXLL motifs that exist in transcriptional proteins and in proteins included in transcription in the PRR. Figure 1 redrawn from (Mathis and Benoist, 2009) using PowerPoint.

As proof that AIRE is a transcriptional regulator, a number of AIRE's molecular partners have been identified. They are classified into four functional classes: transcription, pre messenger ribonucleic acid (mRNA) processing, nuclear transport, and chromatin binding/structure. AIRE influences pre-mRNA processing, and this could possibly justify the efficiency of PTA expression in AIRE-expressing mTECs. The fully spliced RNA is thus stable and translation can be active for a prolonged period of time (Abramson et al., 2010). This facilitates the thymic display of peripheral tissue antigens, promoting self-tolerance and consequently avoiding organ-specific autoimmunity (Koh et al., 2008). The cyclic AMP response element-binding protein (CBP) is a protein partner that functions as a transcriptional coactivator, as it is involved in intrinsic histone acetyltransferase activity. To emphasise, it communicates with transcription factors, adjusting the intrinsic histone acetyltransferase to be proximal to specific locations of the gene promoter in order to facilitate the modifications in chromatin necessary to carry out transcription.

It has been reported that CBP interacts with numerous transcription factors, altering its function, including its role as a nuclear hormone receptor. The fact that CBP acts as a coactivator to transcription factors has been discovered due to its linkage with transcription machinery, including Transcription factor IIB (TFIIB) and RNA helicase A (Liiv et al., 2008). Furthermore, it has been hypothesised that the acetylation of AIRE is facilitated by the overexpression of CBP and its homologue p300. Finally, acetylation occurs within the SAND domain and the nuclear localisation signal (Saare et al., 2012). The nuclear phosphoprotein 63, p63, is another protein partner that functions in controlling cell proliferation and apoptosis. The interaction depends on the SAND domain

of AIRE. In mTECS, the interaction of AIRE-p63 interaction defines the level of transcription of the gene encoding the MHCII transactivator (CIITA), which is crucial for the transcriptional activity of the MHCII promoter (Perniola and Musco, 2014).

A Yeast two-hybrid (Y2H) screen was used to detect another protein partner known as Death-associated protein 6 (DAXX), a novel protein that interacts with the AIRE protein. DAXX is an adaptor protein that is included in both regulation of the transcription pathway and apoptosis. In terms of regulating apoptosis, DAXX acts as a pro-apoptotic protein, as it activates upstream of the Jun Nuclear Kinase (JNK) (Meloni et al., 2010). The Y2H system was also utilised to identify another protein partner: homeodomain-interacting protein kinase 2(HIPK2). Co-transfection has revealed that HIPK2 is colocalised with AIRE in nuclear domains. Moreover, it has been noted that it phosphorylates AIRE and represses its transcription coregulator activity. Furthermore, it controls the promiscuous gene expression and the development of mTECs (Rattay et al., 2015). The protein inhibitor of activated STAT 1 (PIAS1) localised in the nuclear bodies is also a partner that acts as a transcriptional co-regulator.

The interaction of PIAS1 with AIRE has been analysed in vitro and by a confocal microscope. The interaction primarily occurs via nuclear matrix components and is associated with the transcriptional events that are dependent on AIRE. As with PIAS1, they concurrently trigger the human insulin promoter, a target AIRE gene, so they interact to control the activities of AIRE target genes (Ilmarinen et al., 2008). DNA-dependent protein kinase (DNA-PK) is another interacting partner with AIRE. It phosphorylates AIRE at two different positions – Ser156 and Thr68 – therefore enhancing AIRE's transactivation ability (Liiv

et al., 2008). Moreover, the multiacetylation of AIRE-CARD facilitates the production of Bromodomain-containing protein 4 (BRD4) which attracts the positive transcription elongation factor-b (P-TEFb) and therefore stimulates transcriptional elongation. As a result, P-TEFb complex is recruited to RNA polymerase II, allowing the conversion of RNA polymerase II to the elongation form. This therefore activates gene expression (Peterson et al., 2008, Yoshida et al., 2015).

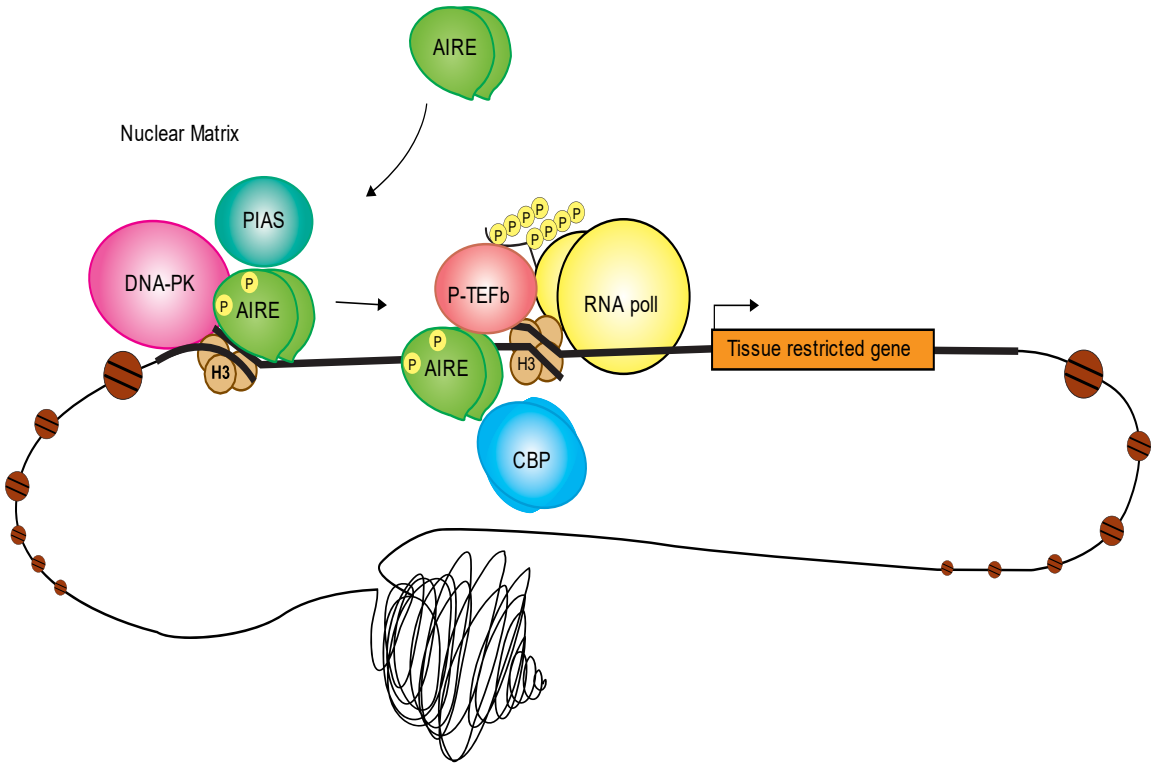


Figure 2. AIRE protein partners. AIRE has a role in the transcriptional elongation via binding to numerous protein partners. The binding of AIRE to DNA-dependent protein kinase (DNA-PK) and to the protein inhibitor of activated STAT-1(PIAS1) facilitates its transcriptional activity. DNA-PK phosphorylate AIRE is present at two sites, and as it binds to the nuclear matrix, it collaborates with AIRE to form chromatin loops (Liiv et al., 2008). AIRE then binds with CBP (cyclic AMP response element-binding protein). This acetylates histone tail residues, promoting AIRE's access to DNA and chromatin (Saare et al., 2012). Lastly, AIRE binds to the positive transcription elongation factor-b (P-TEFb), which in turn phosphorylates the serine residues of RNA polymerase II, converting it to the elongating form. Therefore, gene expression is activated (Oven et al., 2007). Diagram inspired by (Peterson et al., 2008) and created using Adobe Iillustrator.

1.3.2 AIRE expression and its impact on thymic (mTEC) development

As previously mentioned, *AIRE* is prominently expressed in the thymic medulla, particularly in mTECs (Nagamine et al., 1997); (Gillard et al., 2007). It is also expressed in a relatively low amount in DCs in humans (Gotter et al., 2004, Zuklys et al., 2000). Thymic medullary compartments were analysed in *AIRE*-deficient mice. Consequently, mTECs stained negative for the marker Ulex europaeus agglutinin 1 evidenced the absence of immature mTECs (Gillard et al., 2007). Note that it is the terminally-differentiated mTECs that express Aire (Matsumoto, 2011). It was also reported that the distribution of the thymus of clinically generated Aire-deficient mice is mainly in the centre of the medulla, not in the cortico-medullary junction, as compared to the *Aire*-expressed thymus in wild-type mice. Furthermore, the Aire-deficient thymus lacks the thymic corpuscles; hence, its expression is vital for the growth of the normal thymus (Yano et al., 2008). In contrast, it has been stated that components of the tight junctions claudin-4 and claudin-3 are normally localised in *Aire*-expressing mTECs, but their expression is independent of Aire. As the thymus develops, the positive cells of claudin-4 and claudin-3 act as progenitors for cells that express Aire, confirming that *Aire*-expressing cells follows a unique differentiation pathway (Hamazaki et al., 2007). In addition, in Aire-deficient mice, the growth of thymocytes seizes at the final stages of its differentiation (Li et al., 2007), whilst peripheral T cells appear to be normal in mice lacking the *Aire* gene. Hence, defects in the *Aire* gene do not have an effect on non-self-antigens that recognise the TCR lineage. Actually, *Aire* follows a specific selection process, ensuring that balance is maintained in both central and peripheral T cells (Kedzierska et al., 2010). Analysis of the thymus-derived invariant natural killer T cells (iNKT cells) in APS-1 patients and *Aire*-deficient

mice revealed that it was reduced in both patients and mice yet only dysfunctional in mice (Lindh et al., 2010). Defects in Aire result in the reduction of the Scurfy mutation of the FoxP3 gene ($Foxp3^{sf}$) in the thymus, affecting the function of the regulatory T cells (Aricha et al., 2011, Chen et al., 2005). Another study emphasised that the production of the X-C motif chemokine ligand 1 (XCL1) by *Aire* facilitates the migration of thymic DCs to the medulla, and thus facilitates the regulation of T cell development (Lei et al., 2011).

1.3.3 Peripheral tissue restricted antigens

Micro array analysis of mTECs from Aire-deficient and Aire-expressed mice has revealed that PTA expression depends on the sufficiency of Aire (Anderson et al., 2002, Kont et al., 2008). It has also been discovered that in Aire-deficient mice, PTA is the target for autoimmune reactions (Kuroda et al., 2005, Ramsey, 2002). The expression of PTA and the clonal deletion of thymocytes that reacts with antigens occurs in negative selection (Liston et al., 2003) *[see Figure 3].* However, defects in Aire expression do not have any effects on the expression of ovalbumin but result in the evasion of the ovalbumin T cell clones to the periphery, leading to autoimmune diabetes (Myers et al., 2000). Hence, Aire has an effect on the depletion of clonal T cells. As mentioned, defects in Aire have an impact on regulating specific genes, such as those linked with the processing and homing of antigens including cytokines (Anderson et al.). The use of systemic H-2K promoters in transgenic models has indicated that lack of *Aire* gene has no impact on the deletion of the antigen-specific T cells (Liston, 2004). Moreover, deletion of such cells regulated by the thyroglobulin promoter was eradicated in Aire heterozygous mice. The T cells eventually migrated to the periphery, causing no disease. In the thymus, lack

of the Aire gene and hence the loss of PTA expression could also have an effect on the function of peripheral tolerance (Kuroda et al., 2005). mTECS expressing Aire multiply rapidly in a way that enhances the expression of PTA. This therefore reduces the possibility of the DCs in the thymus engulfing the apoptotic cells, cross-presenting the PTA for apoptosis (Gallegos and Bevan, 2004). It has also been noted that Aire is responsible for the circulation of the DCs to the medulla in the thymus as well as the shift of the antigens from the mTECs to the DCs in the thymus (Hubert et al., 2011).

DCs can carry self and non-self-antigens during its migration to the thymus. As mentioned, the transfer of antigens to the DCs is dependent on Aire, emphasising the function of Aire in negative selection (Derbinski and Kyewski, 2010). In humans, it has been reported that AIRE affects the acetylcholine receptor gene *CHRNA1*'s expression in the thymus. This gene is the central target for autoantibodies in myasthenia gravis (Giraud et al., 2007, Berrih-Aknin, 2016). Expression of Aire in monocyte's up-regulated compounds is involved in cytokine signalling and cell adhesion, implying that this expression is essential for DC homing. Moreover, DCs in Aire-deficient mice and APS-1 patients activate a relatively significant amount of T cell hybridomas (Ramsey et al., 2006). In APS-1 patients, DCs originating from monocytes secrete a high rate of tumor necrosis factor α (TNF-α) as well as express more CD86, a co-stimulatory factor, with no prior stimulation (Ryan et al., 2008). Another group claimed that APS-1 patients produce reduced levels of cytokines in response to microbial stimuli, emphasising that DCs strive to fully mature without the expression of AIRE (Sillanpää et al., 2004). In addition, APS-1 patients exhibit immunodeficiency to *Candida* spp as a consequence of defects

in AIRE signaling pathway in DCs (Brännström et al., 2006). Defects in the functioning of DCs arise from monocytes in APS-1 patients as they exhibit a delay in the receptor-mediated endocytosis of *Candida* antigens, resulting in the decreased activation of the intracellular signaling pathway. In innate immunity signalling, defects in AIRE seem to be restricted to DCs (Hong et al., 2009).

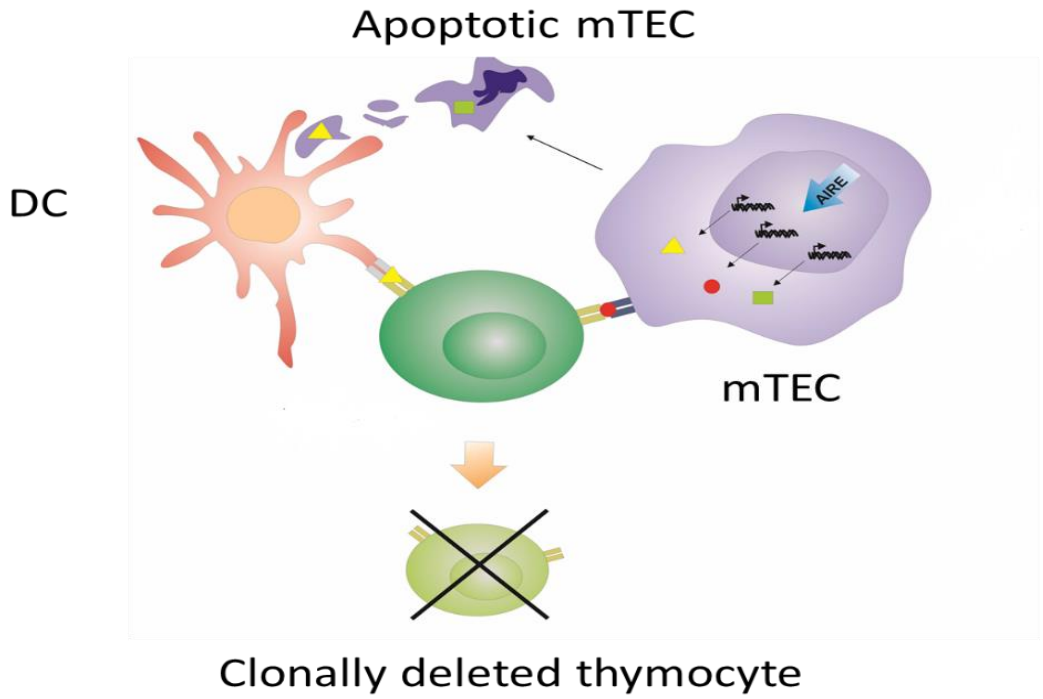


Figure 3. Autoimmune Regulator Protein (*AIRE*) regulates promiscuous antigen expression and induces Central Tolerance. AIRE controls the expression of peripheral tissue antigens (PTAs). PTA in the medullary thymic epithelial cells (mTECs) is bound onto MHC molecules for direct antigen presentation. mTECs die via apoptosis as soon as PTA and AIRE are activated. This results in the uptake of PTA for indirect antigen presentation by dendritic cells (DCs). Developing lymphocytes percolate via the medulla, and if the T cell receptors (TCR) sense the MHC: PTA complexes at the suitable affinity/avidity level, they are deleted by clonal deletion. Figure 3 adapted from (Mathis and Benoist, 2009) using Adobe Illustrator.

1.3.4 Regulation of *AIRE* Expression

In immature mTECs, up-regulation of AIRE is coordinated via the Cluster of differentiation 40 (CD40) protein and the receptor activator of nuclear factor k B (RANK).The three-dimentional meshwork of the mTECs, regulated by an epithelial cell-autonomous gene forkhead box N1 (FoxN1), is central for Aire expression in cells (Akiyama et al. 2008).

Mutations in the FoxN1 gene affect the differentiation of such cells, resulting in two-dimensional structure mTECS. This in turn reduces the expression of AIRE (Guo et al., 2011). Expression of Aire in mTECS follows a distinctive lineage that arises during the development of the thymus, emphasising the fact that mature lymphocytes are not required for AIRE expression (White et al., 2010). Tolerance loss in the Aire-dependent antigens interphotoreceptor retinoid-binding protein (IRBP) (eye antigen) and mucin 6 (stomach antigen) have also been described in Aire-deficient mice. In the thymus, lack of expression of such antigens leads to autoimmunity against the stomach and the eyes, delineating a correlation between the expression of PTA triggered by Aire and autoimmunity in Aire-deficient mice due to the initiation of harmful antibodies that targets these antigens (DeVoss et al., 2006); Gavanescu et al., 2007). In line with the evidence of PTA expression regulated by AIRE, another group demonstrated that expression of proinsulin in the epithelium of the thymus is associated with the amount of functional Aire alleles (Liston, 2004). Moreover, defects in the genes; Nuclear factor kappa-light-chain-enhancer of activated B cells (NFkB) , Transcription factor RelB (Relb), and Tumor necrosis factor receptor (TNFR)-associated factor 6 (TRAF6) that are included in the expression of *AIRE* gene have given some insight in terms of defining the

pathways that are crucial in regulating *AIRE,* and hence the initiation of autoimmunity (Li et al., 2007). Another study has established a direct association among the expression of insulin and AIRE in the thymus, reflecting the regulation of insulin expression by AIRE, which therefore has an impact on T1D (Sabater et al., 2005).

1.3.5 Effect of AIRE on Treg development

The termination of immune reactions facilitated by regulatory T cells is essential for maintaining tolerance in the periphery (Cobbold et al., 2003). Regulatory T (Treg) cells primarily represent $CD24^{+}CD25^{+}$ with the potential to supress autoimmunity. Other T cells involving T helper type 1 (Th1) and T helper type 2 (Th2) mediate regulatory factors by producing specific cytokines (O'Garra and Vieira, 2004). There is also another type of regulatory T cell that is involved in the secretion of Interleukin (IL) 10 as well as the transforming growth factor beta (TGFβ) (Oida et al., 2003). The expression Foxp3 Treg cells is thought to be generated by AIRE/Aire (Malchow et al., 2016). There is also a body of evidence suggesting that Aire-dependent promiscuous gene expression can induce both clonal deletion and Treg cell development in the thymus. Malchow et al. (2016) concluded that Aire promotes immune tolerance by associating these mechanisms, which involves eliminating self-reactive clonotypes from the Tconv cell compartment and promoting the differentiation of these cells into the Foxp3+Treg cells. In this study, upon dysregulation of the former process in Aire-deficient mice, the T cell clonotype that usually differentiates into Treg cells is switched into Tconv cells, directing the self-reactive T cell to enter the prostate. These type of T cells are referred to as rogue cells, which are Treg

cell-biased clonotypes that ‘‘go rogue’’ in the absence of Aire (Metzger and Anderson, 2011); (Malchow et al., 2016).

Moreover, organ-specific autoimmunity generated via continuous ablation of Tregs determines that the T cell pool consists of self-reactive T cells with the potential of being pathogenic, further reflecting the incomplete process of immune tolerance (Kim et al., 2007). (Malchow et al., 2016) also hypothesised that Aire-deficient mice gain autoimmunity as a result of functional deficiency in the compartment of Treg cells, which is possibly due to the absence of TRA-reactive specificities. In line with this notion, (Yang et al., 2015) further showed that the relocation of perinatally-tagged Treg cells from Aire-efficient donors serves as protection from autoimmunity, whilst shifting Tregs from Aire-deficient mice does not prevent autoimmunity. Thus, the potency of developing autoimmunity in an Aire-deficient setting is linked with the generation of rogue cells in the compartment of Tconv cells. Recently, (Takaba et al., 2015) established an essential role for Fezf2 in immune tolerance. Fezf2 is a transcription factor that is induced by mTECs and promotes the promiscuous expression of PTA that is independent of Aire. Moreover, Fezf2-deficient mice display lowered fractions of Tregs in the thymus and exhibit organ-specific autoimmunity revealed by infiltration of immune cells autoantibodies in serum. Therefore, rogue cells could also initiate organ-specific autoimmunity in Fezf2-deficient mice (Takaba et al., 2015).

The hypothesis and the aim of the study are outlined in the following section.

1.4 Hypotheses:

AIRE is involved in several cellular and molecular mechanisms and, as such, our hypothesis is that AIRE interacts with several proteins that have not yet been identified. Our second hypothesis is that APS-1 mutations are associated with APS-1 features, and therefore, it is possible to establish a relationship between AIRE interactions and APS-1 phenotypes.

1.5 Aims and objectives:

Given the hypotheses presented above, this thesis has the following aims and objectives.

1) Identification of new partners of AIRE.

   This will be done by screening the AIRE-PHD1 protein against the universal normalised-human cDNA library in yeast-two-hybrid assay to identify proteins interacting with AIRE.

2) To establish a correlation between APS1 AIRE mutations and AIRE interactions.

   This will be done by analysing the effect of APS1 mutations that are in AIRE, CARD and SAND domains that interact with known proteins.

Chapter 2
Materials and methods

2. Materials and Methods

2.1 Materials

2.1.1 Solutions

TAE- 0.04M Tris-acetate (Sigma).0.001M EDTA (Sigma)

1.0% agarose gel – 1.0 % (w/v) agarose (Bioline) in TAE buffer,0.5µg/ml ethidium bromide (Sigma)

1.1XTE/LiAc – 1X Tris-EDTA (TE buffer, 10mM Tris and 1mM EDTA), 0.11M lithium acetate (Clontech)

Aureobasidin A solution (Aba) – 0.05 %(w/v) Aureobasidin A (Clontech) in absolute ethanol (Fisher Scientific)

X-α-Gal solution – 2% 5-Bromo-4-chloro-3-indolyl-α-D-galactopyranoside (Clontech) in dimethylformamide (Fisher Scientific)

Running buffer - 25mM Tris Base (Sigma), 250mM Glycine (Sigma), 0.1%SDS (Sigma)

Separating Buffer – 1.5M Tris (Sigma), pH 8.8

Stacking Buffer – 0.5M Tris (Sigma), pH 6.8

Separating gel – water, 30% Acrylamide (Protogel), Separating Buffer, 10% SDS (Sigma), 10% APS(Sigma), TEMED(Sigma)

Stacking gel – water, 30% Acrylamide (Protogel), Stacking Buffer, 10% SDS (Sigma), 10% APS (Sigma), TEMED (Sigma)

Tris-buffered saline (TBS) – 50mM Tris-Base, pH 7.5, 150mM NaCl

Anti-HA Magnetic Beads- Pierce™ Anti-HA Magnetic Beads (Thermofisher)

Phosphate-buffered saline (PBS) - 100 mM sodium phosphate; 9.0% NaCl; pH 6.8 (Sigma)

Dual-Luciferase Reporter Assay System- Luciferase Assay Buffer 2 (Promega), Luciferase Assay Substrate (Promega), Stop & Glo Buffer (Promega), Stop & Glo Substrate (Promega), Passive Lysis Buffer (Promega)

BCA assay components – BCA Reagent A, BCA Reagent B, and Albumin Standard Ampules (2 mg/ml)

2.1.2 Bacteria and Yeast media:

LB agar and broth - 35g of LB agar powder (Sigma) or 20g of LB broth powder in 1L of H2O, supplemented with either 50µg/ml Kanamycin Monosulfate or 100µg/ml Ampicillin (Sigma) after autoclave.

2.1.3 Yeast media recipes:

2X YPDA, YPDA, 0.5X YPDA agar and broth.10g, 5g, or 2.5 g of YPD (Yeast Extract Peptone Dextrose) (Sigma) per litre of H2O supplemented with 0.2% of adenine hemisulfate (Sigma) giving YPDA, pH 6.5. For YPDA (Yeast Extract Peptone Dextrose + Adenine) agar, including 2 %(w/v) of bacteriological agar. Synthetically defined (SD) dropout medium – 6.8g of yeast nitrogen base without amino acids (Sigma), 2% (w/v) of D-glucose (AnalaR), 0.2% adenine hemisulfate (Sigma), pH 5.8. For SD agar, including 2% (w/v) of bacteriological agar *[see Table 4].*

Supplement	Selectivity
Without tryptophan(-Trp)	pGBKT7 in Y2H Gold Diploid mated with pGBKT7
Without leucine(-Leu)	pGADT7 in Y187 Diploid mated with pGADT7
Without tryptophan and leucine (-DDO)	Diploid containing pGBKT7 + pGADT7 pGBKT7 in Y2H Gold pGADT7 in Y187
Without adenine, histidine, tryptophan and leucine(-QDO)	Diploid containing pGBKT7 + pGADT7 pGBKT7 in Y2H Gold pGADT7 in Y187
-Trp + X-α-gal	pGBKT7 that activate *MEL*1 reporter gene
-Leu + X-α-gal	pGADT7 that activate MEL1 reporter gene
-Trp +X-α-gal +Aba	pGBKT7 that auto-activate *AUR1-C and MEL1* reporter genes
-DDO +X-α-gal +Aba	Diploids that activate AUR1-C and MEL1 reporter genes
-QDO + X-α-gal + Aba	Diploids that activate , HIS3, ADE2 , AUR1-C, and MEL1 reporter genes

Table 4: A list of all the synthetic drop-out medium supplements used in the study.

2.1.4 Plasmids

Plasmid	Company
pGBKT7	Clontech
pGADT7	Clontech
pcDNA5/FRT	thermofisher
pBIND	Promega
pACT	Promega
pG5Luc	Promega
pCMV3-TP63-HA	Stratech
pCMV3-DAXX-HA	Stratech
pCMV3-PIAS1-HA	Stratech
pCMV3-BRD4-HA	Stratech
pCMV3-AIRE-CMYC	Stratech

Table 5: A list of the all the plasmids that were used in this research.

2.2 Methods

2.2.1 Nucleic acid techniques

2.2.1.1 Amplification of insert

The cDNA was amplified from a previously cloned cDNA template using High Fidelity Taq DNA polymerase (Thermo fisher) following the manufacturers protocol. Reactions were incubated in a PCR thermal cycler under the following program: initial denaturation for 2 minutes at 94°C followed by 35 cycles of 30 seconds at 94°C (denaturation), then 30 seconds at an annealing temperature *[see Table 6]* and at 72°C (extension) for 1 min/kb. Reactions were later visualised on 1% (w/v) agarose gel, the product was then excised and purified using QIAquick Gel Extraction Kit (Qiagen). Primers used in amplification are listed in Table 7.

cDNA (Insert)	Temperature
PHD1	63
PHD1	66
SAND (600bp)	65
SAND (300 bp)	65
CARD	63
AIRE	66
fp63	55
tp63	62
PCNA	63
ATP1β1	62
NPC2	63
STX8	63
RNF2	63
PGM1	63

Table 6: The annealing temperatures for the amplified cDNA inserts used in the study.

Primer	Forward	Reverse
pGBKT7PHD1	5` AGTAAC**CCATGG**CTCTGGCCCTCCCCAGTGA 3`	5`ATTGTA**GTCGAC** CTGCACCTCCTGGACTGTTG 3`
pFRTPHD11bp	5`CTCAAG**GGATCC** CGATGGAGGAGCAGAAGCTGA 3`	5` CTCGAT**GCGGCCGC**TCATTACAGGCAGCTGGAGCACC 3`
FRTprey	5` TGTGAA**GGATCC**atggagtacccatacgacgtacc 3`	(FRTATP) 5` TTCAGT**GCGGCCGC**TTATCA GCTCTTAACTTCAATTT 3` (FRTNPC2) 5` TTGGAT**GCGGCCGC**TCATTAGAGATGAGAAACGATCTG 3` (FRTSTX) 5` TTCTCA**GCGGCCGC**TCAGTTGGTCGGCCAGACTGCAA 3` (FRTPGM1) 5`TTGTGA**GCGGCCGC**TTAGGTGATGACAGTGGGTGCAGTG 3` (FRTRNF2) 5` TTGTGA**GCGGCCGC**TCACTGACCAATTCCAAAGAAA 3` (FRT PCNA) 5` TTCTCA**GCGGCCGC**CTAAGATCCTTCTTCATCCT 3`
FRTCARD	5` TGTGAA**GGATCC**ATGGAGGAGCAGAAGCTGA 3`	5` TGTGAT**GCGGCCGC**TCAGAAGCTGTCCAGGATG GGCT 3`
FRTSAND	5` TGTGAG**GCTAGC**ATGGAGGAGCAGAAGCTGATC 3`	5` TGTGAT**GCGGCCGC**TCAGGGAACGCTGCCCTGCTG 3`
pBINDSANDG228W2bp	5` GTGAGA**ACGCGT**CAATGAGAGCTGTGGCCATG 3`	5` CTCTTC**GGTACC**AGGGAACGCTGCCCTGCTG 3`
pBINDAIRE2bp	5` CTCAGT**ACGCGT**CTATGGCGACGGACGC 3`	5` CTGAGT**GCGGCCGC**TTACAGATCCTCTTCTGAGATG 3`
pACTfp632bp	5` AACAACT**GGATCC**CTATGAATTTTGAAACTTCA 3`	5 `TTCCAC**GCGGCCGC**TCACTCCCCCTCCTCTTTGATGCGC 3`
pACTtp632bp	5` AACAACA**GGATCC**CAATGCAGCATATCTGGGA 3`	5 `TTCCAC**GCGGCCGC**TCACTCCCCCTCCTCTTTGATGCGC 3`

Table 7: a list representing all the primers used in the project.

2.2.1.2 Restriction digest and ligation of purified samples into a mammalian vector

The plasmid and the purified PCR products were digested using HF-restriction enzymes (NEB) following the manufacturer protocol. Digested products were run on agarose gel to visualise the excise bands that correspond to digested empty plasmid and the digested cDNA insert. All DNA samples were purified using QIAquick Gel Extraction Kit (Qiagen). Each digested-purified insert was ligated to the linearised empty vector at a 3:1 ratio and incubated with T4 DNA ligase and Ligase buffer (Promega).

2.2.1.3 Site-directed mutagenesis

To introduce mutations by replacing one amino acid in the wild type AIRE-CARD and AIRE-SAND, the QuikChange II XL Site-Directed Mutagenesis kit (Agilent Technologies) was used according to the manufacturer's instructions. Primers were designed using the QuickChange Primer Design website. The site-directed mutagenesis method was carried out using a high-fidelity DNA

polymerase for mutagenic primer-directed replication of both plasmid strands. Basically, the procedure involved using a supercoiled double-stranded DNA plasmid with an insert of interest and two synthetic oligonucleotide primers including the desired mutation. The primers used were between 25 and 28 bases in length and were generated to anneal to the same sequence on opposite strands of the vector. Both primers included the desired mutation in the middle flanked by sequences complementary to the template sequence.

The primers were extended during temperature cycling by the DNA polymerase. Extension of the primers produced a mutated vector involving nicks. Following temperature cycling, the PCR product was treated with DpnI endonuclease, which is specific for methylated and hemi-methylated DNA and was used to digest the parental DNA template and to select for mutation harbouring synthesised DNA. The nicked plasmid DNA incorporating the desired mutation was then transformed into E. coli XL10 Gold cells *[Table 8].*

Mutagenic primers	Forward	Reverse
FRTG228W	5'-gaagtgcatccaggttggttgggagttctacac-3'	5'-gtgtagaactcccaaccaacctggatgcacttc-3'
FRTP252L	5'-cagcagcagtggcctgaagcctctggttc-3'	5'-cttggtctccgaagtccggtgacgacgac-3'
FRTT16M	5'- aggctgcaccgcatggagatcgcgg-3'	5'- ccgcgatctccatgcggtgcagcct-3'
FRTL29P	5'- ccttcccactgccgcacgcgctggc-3'	5'- gccagcgcgtgcggcagtgggaagg-3'
FRTL93R	5'- gctatggccggcggcagcccatcct-3'	5'- aggatgggctgccgccggccatagc-3'

Table 8: A list representing the mutagenic primers used in this research.

2.2.1.4 Bacterial cell transformation

All products were then transformed using α-select Gold Efficiency chemically competent E. coli (Bioline) following the manufacture protocol and spread on LB-agar plates including antibiotics. Plasmids were then isolated by QIA quick mini prep (Qiagen) according to the manufacturer's instructions, and the cDNA insert was verified by sequencing.

2.2.2 Yeast-2-Hybrid techniques

2.2.2.1 Yeast cell transformation

Prior to transformation, the DNA concentration of empty pGBKT7 and pGBKT7-PHD1 plasmid solutions clones were quantified using a NanoDrop. Y2HGold and Y187 cells (Clontech) were used to generate competent yeast cells following the protocol in the Yeastmaker™ Yeast Transformation System2 (Clontech). 100 µl aliquots of 1/10 and 1/100 dilutions were spread on the appropriate SD agar plates.

2.2.2.2 Testing bait (pGBKT7) for auto-activation

100µ aliquots of 1/10 and 1/100 dilutions of the transformation mixtures pGBKT7 PHD1 were spread on SD/-Trp, SD/-Trp/X-α-gal, and SD/-Trp/ X-α-gal, AbA agar plates and incubated for 3 days at 30°C.

2.2.2.3 Testing bait (pGBKT7) for toxicity

100µl aliquots of 1/10 and 1/100 dilutions of Y2HGold cell bearing the empty pGBKT7 and the Y2HGold transformation mixtures including either pGBKT7 or PHD1 were spread onto SD/-Trp agar plates and incubated for 3 days at 30°C.

2.2.2.4 Two-Hybrid library screening

Mate & Plate libraries are ready-made libraries that are prepared from mRNA obtained from adult human tissue lysates and pre-transformed into MAT α haploid yeast (Y187). They are compatible for mating with the yeast MAT a two-hybrid reporter strain, Y2HGold, to express bait constructs. These libraries look for interactions between a known protein of interest (bait) and their interaction partners (preys) that exist in the cDNA library (Chien et al., 1991).

Prior to screening, kanamycin was added to the Two-Hybrid screening broths at a concentration of 50µg/ml. Library screening was carried out via yeast mating following the protocol in the matchmaker Gold Yeast Two-Hybrid System user manual. pGBKT7-PHD1 AIRE cDNA passed the bait-auto-activation test and hence was screened against the normalised universal human Mate & Plate library.

2.2.2.5 Prey-plasmid segregation

All blue colonies that grew on SD/-QDO/X-α-gal/Aba were individually re-streaked on SD/-DDO/X-α-gal agar plates. Single isolated colonies from the previous streaking were further re-streaked on SD/-DDO/X-α-gal to increase the chance of rescuing the positive prey plasmid.

2.2.2.6 Yeast plasmid isolation and transformation

Yeast plasmid was rescued using the Easy Yeast Plasmid Isolation Kit. All products were then transformed using α-select Gold Efficiency chemically competent E. coli (Bioline) and spread on LB-agar plates including ampicillin. The prey plasmid was then isolated by QIA quick mini prep (Qiagen) and the cDNA insert was determined by sequencing.

2.2.3 Protein analysis

2.2.3.1 Cell Line and Cell Culture

Human embryonic kidney 293 (HEK293) cells were used for transfection and biochemical assays. Dulbecco's Minimum Essential Medium (DMEM with 4.5 g/L Glucose, L-glutamine, Lonza) supplemented with 10% Foetal Bovine Serum (FBS, Lonza) was used to culture the cells. Cells were cultured in a 37°C incubator supplemented with 5% CO2. Cells were cultured when reaching 70-90% confluence. Initially, cells were washed with Dulbecco's phosphate buffered saline (PBS) and then incubated with Trypsin/EDTA for 5 minutes in a 37°C incubator. A medium was added to the trypsined cells to neutralise their activity. Then, the cells were washed by centrifugation for 5 minutes at 1500 rpm. After discarding the supernatants, the cells were diluted 1:3 up to 1:10 using a fresh complete culture media and transferred to new tissue culture flasks.

2.2.3.2 Transfection of HEK293

The transfection efficiency of the pCMV-GFP was assessed in HEK293 cells (Human Embryonic Kidney Cell Line). In 6-well plates, 4 $x10^5$ cells were seeded, and after 24 hours, the cells were transfected with 1 µg of DNA using FuGENE 6® (Promega) at a ratio of 3:1 (3 µl of FuGENE6: 1 µg of DNA), according to manufacturer's protocol. 24 hours after transfection, cells were washed with PBS and observed under fluorescence microscope (Leica AF6000LX inverted microscope) for GFP expression. For protein extraction, cells were harvested and lysed in IP buffer.

2.2.3.3 Preparation of Cell Lysates

HEK293 Cells were washed with PBS and pelleted in 1.5 Eppendorf tubes before adding the IP buffer, which consisted of 3M NaCl (Sigma Aldrich), 0.5M

PO4 (Sigma Aldrich), 1% Triton x-100 (Sigma-Aldrich). 1% Nonidet P40 (Sigma Aldrich) was supplemented with 1% Protease Inhibitor Cocktail (Sigma-Aldrich). Briefly, the cell pellet was suspended in 100 µl of IP buffer and rotated and incubated for 30 minutes in a cold room to lyse the cells. The lysates were centrifuged at 18,000g at 4°C for 20 minutes, and then the supernatant containing the protein was separated from cells debris and collected into a labelled 1.5 Eppendorf tube.

2.2.3.4 Protein quantification assay

Bicinchoninic Acid Protein (BCA) assay kit (Thermo Fisher) was used to determine the protein concentration in each sample. As per the manufacturer's guidelines, nine serially diluted protein standards in addition to 10 µl of 1:3 diluted samples were added individually and in triplicates to the aliquot BCA working reagent. The working reagent was prepared by adding reagent A to reagent B in a 50:1 ratio, then pipetted into a 96-well plate in a quantity of 100 µl per well. The plate was incubated for 30 minutes at 37°C. A microplate spectrophotometer read the absorbed light at 570 nm. The absorbance values were correlated to the protein concentration. According to the standard curve equation produced by plotting the absorbance readings, the unknown protein concentrations were calculated using Microsoft Excel.

2.2.3.5 De-glycolysation test

20µg of glycoprotein was combined with 1 µl of Glycoprotein Denaturing Buffer (10X) and water. The denatured glycoprotein was denatured by heating 100°C for 10 minutes, chilled on ice, and centrifuged for 10 seconds. The following products were added to the glycoprotein: 2 µl GlycoBuffer 2 (10X), 2 µl 10%

NP-40, 6 µl water and 1 µl PNGase F and mixed. The reaction was incubated for 1 hour at 37°C.

2.2.3.6 De-phosphorylation test

The protein sample was combined with water to a volume of 40 µl followed by 5 µL of 10X NEBuffer for Protein MetalloPhosphatases (PMP), 5 µL of 10 mM $MnCl_2$, and 1 µl of Lambda Protein Phosphatase. The total reaction was incubated for 30 min at 30°C.

2.2.3.7 Immunoprecipitation

Cells were washed in PBS, resuspended in 0.1 ml IP buffer (3M NaCl (Sigma Aldrich), 0.5M PO4 (Sigma Aldrich), and 1% Triton x-100 (Sigma-Aldrich). 1% Nonidet P40 (Sigma Aldrich) was supplemented with 1% Protease Inhibitor Cocktail (Sigma-Aldrich), transferred to a 1.5 ml Eppendorf tube, and rotated for 30 min at 4°C. The lysed cells were then spun at 84,000 g (-128) for 30 minutes at 4°C.

Next, 5 µg of immunoprecipitating antibody was added to the lysate followed by one-hour incubation at 4°C with rotation. Then, 30 µl pre–washed protein G beads were added to the protein lysate-antibody mix and left to incubate for 15 minutes at 4°C with rotation. The sample was then placed in a DynaMag magnets (Thermofisher) to remove non-specific immunoglobulin/protein G-binding proteins (unbound-proteins). Samples were washed 3 times with PBS. After each wash, they were placed in a DynaMag magnet to remove unwanted (supernatant) residues. 30 µl of Laemmli Sample Buffer was added to elute the protein complexes. To analyse the proteins, the beads were boiled in Laemmli sample buffer in order to remove the bound proteins from the protein G beads.

Protein samples were boiled and loaded onto a SDS-PAGE gel for separation and western blotting.

2.2.3.8 Western Blotting

The samples were denatured by heating at 95°C for 5 minutes with Laemmli buffer (Bio-Rad) and loaded onto a 15% SDS-PAGE gel. The gels were electrophoresed in Tris/Glycine/SDS running buffer at 60V for 20 minutes then 90 minutes at 120V. For immunodetection, proteins were transferred onto a PVDF membrane (Thermofisher) using an iBlot machine as per the manufacturer's protocol. The anode transfer stack (bottom) containing the PVDF membrane was located on the blotting surface of the iBlot machine. Pre-run gel was placed on the PVDF membrane of the anode stack. The PVDF was covered with water pre-soaked iBlot fiter paper. The iBlot Cathode Stack (Top) was placed on top of the pre-soaked filter paper. Air bubbles was removed via the Blotting Roller. The iBlot Disposable Sponge was placed on the inner side of the lid to absorb any excess liquid generated during blotting. The iBlot™ Disposable Sponge was then placed on the inner side of the lid (between the small protrusions on the lid that hold the sponge in its place) such that the white side was facing the experimenter and metal contact was to the top right, as described below for the Mini Transfer Stack. The sponge absorbed any excess liquid generated during blotting and exerted an even pressure on the stack surface. The iBLot Lid was closed and latch was secured. Program P2 was selected to start the transfer. Membranes were placed in a tube containing 5% milk TBS to be blocked for 60 minutes at room temperature on a roller (Milton) and washed for 30 minutes with TBS buffer. Primary and secondary antibodies were diluted as per the manufacturer's protocol in 5% milk TBS-blocking buffer.

The membrane was incubated with the primary antibody at 4°C overnight with agitation, then washed for 30 minutes with TBS. HRP (horseradish peroxidase) secondary antibodies were used to incubate the correspondent membrane and incubated for 1 hour at room temperature on a roller (Milton). Before the primary and secondary antibody incubations, the membranes were washed three times for 10 minutes in TBS at room temperature in the roller (Milton) to eliminate un-bound antibodies. The proteins were visualised using an ECL Plus chemiluminescence detection kit (GE Healthcare) for HRP. The membrane was visualised using the G-BOX Image Capture System (G-BOX, Syngene).

Antibody	Company
Rabbit anti-c-myc Antibody, Affinity Purified	Sigma
HA-probe Antibody (Y-11)	Santa Cruz
Anti-Rabbit IgG (whole molecule)–Peroxidase antibody produced in goat	Sigma

Table 9: A list of the antibodies used in this research.

2.2.3.9 Mammalian-2-hybrid system

The mammalian two-hybrid system and the dual-luciferase reporter assay system (Promega) was used to semi-quantify the interaction between AIRE-SAND and P63. The positive control vectors used in the experiment express two proteins known to interact in vivo, MyoD and Id. These were included in the assay kit for use as pACT- MyoD (VP16-MyoD) and pBIND-Id (GAL4-Id) fusion constructs. HEK293 cells were seeded in 12-well plates, and each well was transiently transfected with the luciferase plasmids containing the required inserts. 24 hours after transfection, the cells were lysed with 1X passive lysis buffer (Promega) and the activities of firefly (Photinus pyralis) and renilla

(Renilla reniformis) Luciferases were measured sequentially. Initially, the activity of the firefly luciferase reporter was determined by adding Luciferase assay reagent II (LAR II) to produce a luminescent signal. After quantifying the firefly luminescence, the reaction was quenched, and the Renilla luciferase reaction was measured simultaneously by adding Stop & Glo® reagent to the same sample.

2.2.4 Statistical analysis

The band intensity data was exported in an Excel file where the band intensity of each variable was measured and normalised to the total protein lysate. Comparison of data between samples in each experiment was carried using a two-tailed unpaired t-test. The luciferase data was exported in Excel where the mean firefly luciferase activity was normalised to the renilla luciferase control of four independent replicates, each performed in triplicates. Comparison of data between samples in each experiment was carried using a two-way ANOVA test. For both experiments, the analysis was performed using Graphprism software version 8 with a P value < 0.0001. Graphs were generated using Graphprism software version 8.

Chapter 3
Identification of proteins interacting with AIRE-PHD1 using the yeast two hybrid system

3. Identification of proteins interacting with AIRE-PHD1 using yeast two hybrid system

3.1 Outline

In this chapter, we will focus on the PHD1 domain of the AIRE protein. This is relevant to this thesis for the following reasons. Firstly the surface of PHD1 is negatively charged, suggesting a role in protein interactions more than in nucleic-acid binding (Perniola et al., 2014). Moreover, there is a link between PHD1–H3 interactions and autoimmune disease, which we will explore via a discussion of the reasons for APS-1 causing PHD1 mutations during *in vitro* binding and *in vivo* functional experiments (Koh et al., 2008). This emphasises its crucial role in AIRE function and is therefore a central domain for its ability to prevent mutli-organ autoimmunity (Bottomley et al., 2005; Koh et al., 2008; 2010; Žumer et al., 2012).

The aim of this present study is to identify novel proteins interacting with AIRE-PHD1 using yeast-two-hybrid library in order to widen our understanding of the function of the AIRE protein. To clone and express AIRE-PHD1 protein, the cDNA sequence encoding AIRE-PHD1 protein domains was amplified. The resultant PCR products were then cloned into the pGBKT7 plasmid, transforming them into E. coli. The purified recombinant plasmids were then sequenced and checked for genetic variations.

The plasmid that expresses the AIRE-PHD1 domain was then transformed into Y2H gold cells to be ready for library screening. This method allowed us to identify six protein candidates, including the proliferating cell nuclear antigen

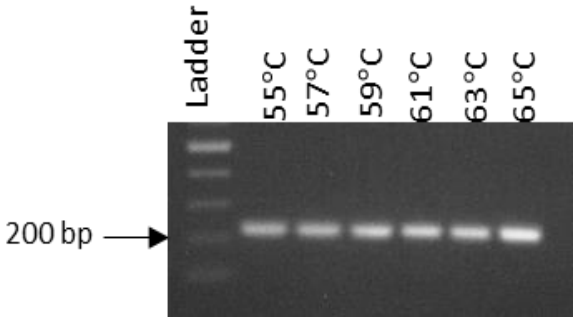
(PCNA), which had already been identified by Abramson's group and therefore validated our screening (Abramson et al., 2010). PCNA has a main role in DNA replication and in the assembly of chromatins, and thus has an impact on transcription. The second section of this chapter describes our testing of the protein expression of the newly identified samples via Western blot. We then describe the immunoprecipitation experiments that were optimised by carrying out a direct-IP for the different protein candidates before commencing to co-immunoprecipitation, a technique used to detect the protein-protein physical interaction, to further confirm interaction of protein candidates with the PHD1 domain.

3.2 Generation of pGBKT7 PHD1 AIRE cDNA for Y2H screening

3.2.1 Amplification of AIRE-PHD1 cDNA insert

In order to screen our library, first we had to clone the sequence encoding PHD1 domain of the AIRE protein into the pGBKT7 bait vector. The domain was amplified by PCR from pcDNA/FRT *AIRE* cDNA using domain-specific primers that allowed the introduction of Nco1 and Sal1 restriction sites at the domain's 5' and 3' ends, which are also situated at the multiple cloning site of the pGBKT7 domain *[see Figures 4A and 4B].*

(A)



(B)

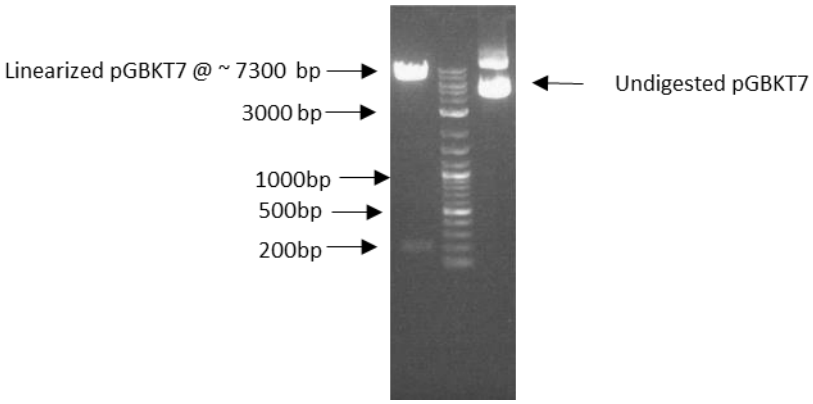


Figure 4. Construction of pGBKT7/AIRE-PHD1.

A. Optimisation of PHD1 AIRE cDNA.
   PCR was optimized for the PHD1 domain at the following annealing temperatures: 55°C, 57°C, 59°C, 61°C, 63°C, and 65°pGBKT7 PHD1 AIRE-cDNA digestion using Nco1 and Sal1.

B. The 200 bp bands correspond to AIRE-PHD1 samples. The 7300 bp corresponds to the linearised pGBKT7.

3.2.2 Testing AIRE-PHD1 for auto-activation and toxicity

Before screening the AIRE-PHD1 domain against the Universal Human Mate & Plate™ Library, it was important to confirm that the bait, which is bound to a DNA-binding domain, does not autonomously activate the reporter genes in the absence of the GAL4 activation domain. Thus, auto-activation and toxicity tests were carried out to rule out any false positive results. The SD/-Trp represents a synthetic drop-out media omitting tryptophan and was used as a control, reflecting the ability of pGBKT7 plasmid to synthesise this amino acid on its own. The SD/-Trp can be supplemented with X-α-gal, a chromogenic substrate for the enzyme alpha-galactosidase. Hydrolysis of X-α-gal by the alpha-galactosidase, encoded by the *MEL1* gene, causes the yeast colony to develop a blue colour, indicating a yeast-two-hybrid interaction. Moreover, we added Aureobasidin A (AbA), a potent and unique yeast antibiotic that kills *S. cerevisiae* at low concentrations (Takesako et al., 1993). It acts by inhibiting inositol phosphorylceramide synthase, an essential yeast enzyme. Upon yeast-two-hybrid interactions, the mutant enzyme encoded by the reporter gene AURE1-C was activated, allowing the yeast colony to confer resistance to Aba. Activation of this reporter gene also confirmed the yeast-two-hybrid interaction. Results are summarised in *Table 10.*

Plasmid	Selective agar media		
	SD/-Trp	SD/-Trp/X-α-gal	SD/-Trp/ X-α-gal/Aba
Empty pGBKT7	White colonies	White colonies	No growth
pGBKT7 PHD1	White colonies	White colonies	No growth

Table 10: Summary of auto-activation test results. Both Empty PGBKT7 and PGBKT7 PHD1 transformed in Y2HGold cells gave white colonies when spread on SD/-Trp/X-α-gal agar plates. No colonies were observed when plated on SD/-Trp/X-α-gal/Aba plates, thus AIRE-PHD1 is unable to autoactivate the MEL1 and AUR1-C reporter genes.

3.3 Library Screening

Once we confirmed that the AIRE-PHD1 did not autoactivate the reporter genes *MEL1* and *AUR1-C*, we moved on to screen the AIRE-PHD1 bait fusion protein against the Universal Normalised Human cDNA. Mated cultures were plated on 150mm SD/-DDO/X-α-gal/Aba agar plates, which could then activate the reporter genes *AUR1-C* and *MEL1* upon an interaction. Blue colonies, observed after 4 days of incubation at 30°C, were plated onto the more stringent SD/-QDO/X-α-gal/Aba agar plates. These agar plates lacked the amino acids, Histidine (His3) and Adenine (Ade2). As Y2HGold is not able to synthesise histidine, it is unable to grow on media that lack this amino acid. However, when bait and prey proteins interact, Gal4-responsive His3 expression allows the cell to synthesise histidine and grow on media that lack histidine. Furthermore, Y2HGold is also unable to grow on media that lacks the amino acid Ade2. Upon protein interaction, however, the amino acid Ade2 expression is activated, allowing these cells to grow on media lacking Ade2. Adding X-α-gal and Aba to these agar plates resulted in the activation of the four reporter genes: *AUR-1 C, MEL1, HIS3,* and *ADE2*. This had confirmed the positive interactors and confined them into six *[see Table 11].* These interactors appeared as blue colonies in the SD/-QDO/X-α-gal/Aba agar plates *[see Figure 5].*

Plasmid	Selective agar media	
AIRE-PHD1	SD/-DDO/ X-α-gal /Aba	SD/-QDO/X-α-gal /Aba
	12	6

Table 11: The number of screened clones in AIRE-PHD1 domains. 12 blue colonies were observed when screened against the AIRE-PHD1 domain, which reduced to 6 clones when colonies were streaked on the high stringency plate SD/-QDO/ X-α-gal /Aba plates.

Figure 5: Positive interactions on an SD/-QDO/X/A plate. Positive interactions appear as blue colonies in the high stringency plates (SD/-QDO/ X-α-gal /Aba), delineating confirmation of interaction.

3.3.1 Segregation, isolation, and transformation of interacting prey plasmid

To segregate the blue positive prey plasmid from any potential contaminants, colonies from the highest stringency agar plates were streaked twice on SD/-DDO/X-α-gal agar plates *[see Figure 6].*

Figure 6: Segregation of interacting prey plasmid. A single blue colony was streaked two times on DDO/X-α-gal plates to further segregate the prey plasmid.

Positive prey colonies were then extracted from the identified yeast diploids via streaking on SD/-Leu plates. The reason for streaking onto this type of plate is that because they are solely selective for prey colonies transformed into E. coli and hence to extract the plasmid for sequencing to identify the prey protein. Colonies appear white, as at this stage, we were not looking for an interaction. Thus, we did not include X-α-gal in the media *[see Figure 7].*

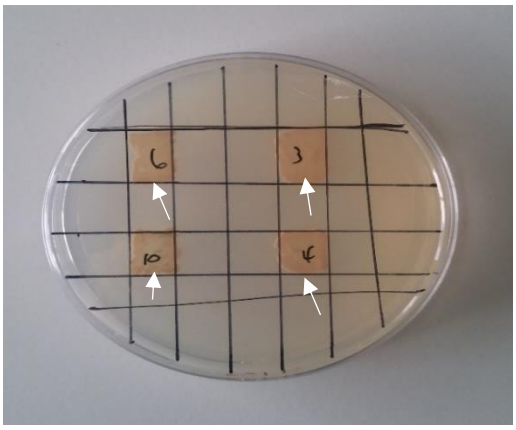


Figure 7: isolation of prey cDNA plasmid. A single blue colony was picked and streaked on SD/-Leu plates. The sequences were entered NCBI blastx to compare against non-redundant protein sequences.

Analysis of the sequences revealed that the proteins that interacted with AIRE-PHD1 were the following: Ring finger 2 (RNF2), Syntaxin8 (STX8), ATPase Sodium/potassium transporting subunit beta 1 (ATP1B1), Niemann-Pick type C2 (NPC2), Phosphoglucomutase 1(PGM1) and Proliferating cell nuclear antigen (PCNA).

3.3.2 Prey DNA-binding assay

The prey proteins transformed in Y187 cells were individually streaked on SD/-Leu/X-α-gal agar plates to determine whether these could auto-activate the *MEL1* reporter gene. White colonies were observed from all prey proteins, Proliferating cell nuclear antigen (PCNA), ATPase Na+/K+ Transporting Subunit Beta 1 (ATP1β1), Niemann-Pick Disease Type C2 Protein (NPC2), Phosphoglucomutase1 (PGM1), Ring Finger Protein 2 (RNF-2), Syntaxin 8 (STX8), suggesting that they did not activate the promoter of the *MEL1* reporter, which would have led to expression of α-galactosidase. Prey proteins are classified based on their known function in *Table 12.*

Protein	Function	Amino acid Position	References
PCNA	DNA replication, DNA excision repair, and in assembly of chromatins	Aspartic acid 150 to Serine 261	(Abramson et al., 2009)
ATP1β1	Regulating the electrochemical gradients of Na+ and K+ ions across the plasma membrane.	Cysteine 159 to Serine 300	(Bab-Dinitz et al., 2009)
NPC2	Cellular cholesterol regulation in the innate immune system	Proline 90 to Leucine 151	(Frolov et al., 2013)
PGM1	Neutrophil degranulation	Glutamic acid 272 to Threonine 365	(Lyons et al., 2015)
RNF2	E3-ubiquitin-protein ligase	Glutamine 11 to Phenylalanine 309	(Bergink et al., 2006)
STX8	Protein trafficking and ubiquitin ligase activity	Methionine 1 to Asparagine 236	(Pattu et al., 2012)

Table 12: **Potential interacting proteins identified in Y2H screening**. The table shows a list of proteins suspected to interact with the PHD1 domain of the AIRE protein, along with their functions. PCNA has role in DNA replication, DNA excision repair, and in chromatin assembly. ATP1β1 regulates the electrochemical gradients of sodium and potassium ions across the plasma membrane. NPC2 regulates cholesterol in the innate immune system. PGM1 functions in de-granulating neutrophils. RNF2 is a E3-ubiquitin-protein ligase. STX8 facilitates protein trafficking and ubiquitin ligase activity.

3.4 Transfection efficiency

The pCMV-GFP expression plasmid was used as control plasmid. The visualisation of HEK293-transfected cells under a fluorescent microscope served as a tool to evaluate the transfection efficiency. HEK293 cells were transiently transfected with the plasmid, pCMV-GFP. The CMV promoter displayed a strong GFP expression in HEK293 cells 24 hours after transfection *[see Figure 8]*.

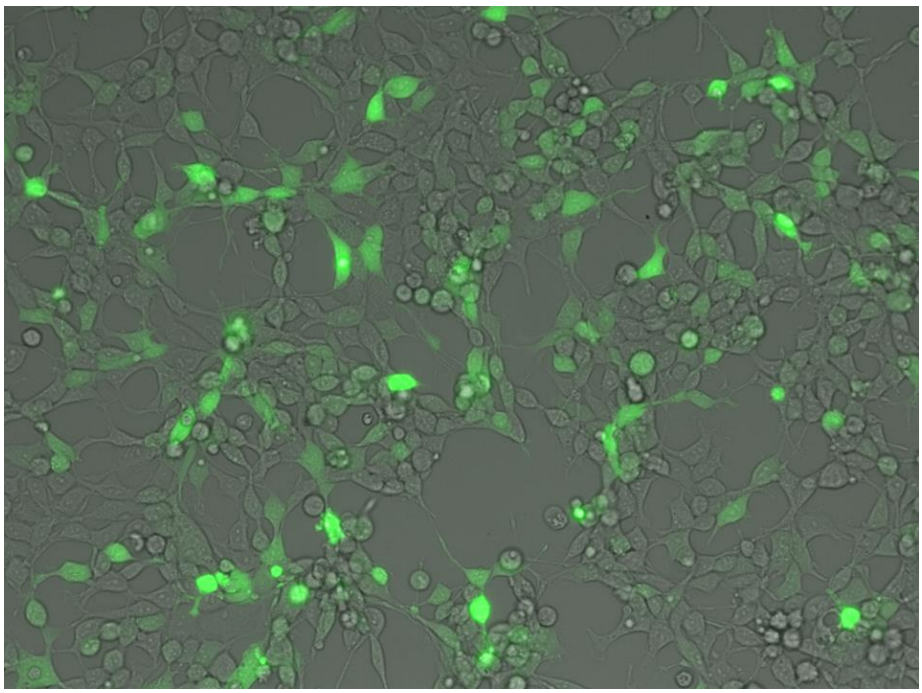


Figure 8: Visualisation of GFP transfected cells. HEK293 cells were seeded on 6-well plate and transfected with a pCMV GFP on the next day. 24 hours after transfection, the cells were assessed under a fluorescent microscope at 10x magnification. Scale bar represents 100µm.

3.5 Confirmation of protein expression

In order to validate the results that we obtained from the Y2H system via co-IP, we needed to confirm the protein expression of the isolated proteins and to optimise our immunoprecipitation experiment by carrying out a direct-IP. HEK293 cells were seeded in a trans-well plate, as described in Section 2.2.3.2. Following 24-hour post seeding, they were individually transfected with a plasmid bearing our prey cDNA. Cells were lysed and direct immunoprecipitation (IP) was carried out. IP with Cmyc-tagged beads (MYCIP) (Thermofisher) were used for AIRE-PHD1 and IP with HA-tagged beads (HAIP) (Thermofisher) were used for ATP1B1-NPC2-STX8-PCNA-PGM1-RNF2.

The results are for overexpression of total protein lysate (positive control), direct immunoprecipitation for each of the pulled down proteins, and un-transfected lysate (negative control). We managed to detect expression and to optimise a direct-IP experiment for the following proteins: AIRE-PHD1 ~ 10 kDa ; ATP1β1 ~ 20 kDa ; NPC2 ~ 10 kDa ; PCNA ~ 15 kDa ; PGM1 ~ 15 kDa ; RNF2 ~ 40 kDa; STX8 ~ 30 kDa *[see Figure 9].*

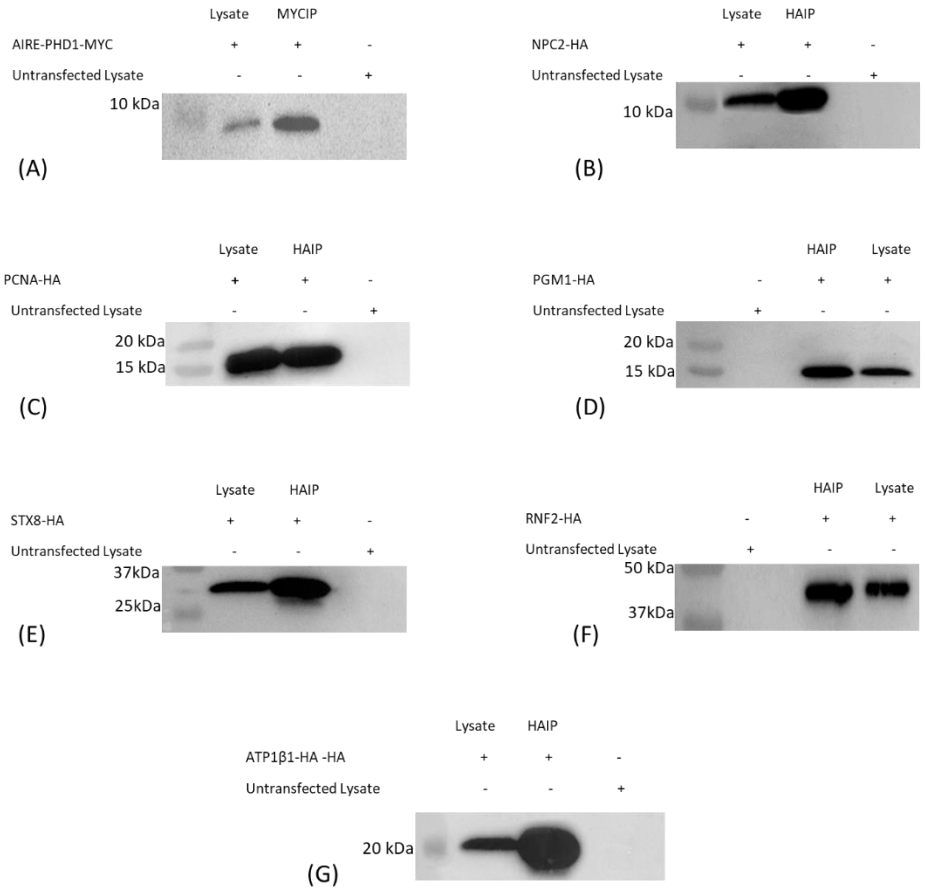


Figure 9: In vitro transfection of the pcDNA5/FRT plasmid containing one of the following inserts: AIRE-PHD1/ATP1B1/NPC2/STX8/PCNA/PGM1/RNF2 expression plasmids. The figure shows a representative western blot of the transfected cells. HEK293 cells were individually transfected with either one of the mentioned plasmids and were lysed using IP buffer and then direct immunoprecipitation was carried out using CMYC-tagged beads for AIRE-PHD1 and HA-tagged beads for ATP1B1-NPC2-STX8-PCNA-PGM1-RNF2. The results are shown for overexpression of total protein lysate (positive control), direct immunoprecipitation for each of the pulled down proteins, and un-transfected lysate (negative control). AIRE-PHD1 ~ 10 kD; ATP1β1 ~ 20 kD; NPC2 ~ 10 kDa; PCNA ~ 15 kDa; PGM1 ~ 15 kDa; RNF2 ~ 40 kDa; STX8 ~ 30 kDa . MYCIP: proteins pulled down with MYC tagged IP; HAIP: proteins pulled down with HA tagged IP.

3.6 Co-immunopreciptation

After confirming protein expression in the previous section, we carried out Co-IP in order to validate the interactions that were detected in the Y2H-system. HEK293 cells were seeded in a trans-well plate, as described previously. The mammalian expression plasmids containing the isolated prey cDNA's were individually co-transfected with a plasmid containing AIRE-PHD1 cDNA. This plasmid was also co-transfected with an empty bacterial plasmid, Pet28, in order to carry out a direct IP, which was used a control. 24-hour post transfection, the cells were lysed, and a co-IP was carried out using HA-tagged beads, HAIP, to pull out the prey protein along with AIRE-PHD1 if they interacted. In contrast, a direct-IP was done using a cMyc-tagged bead, MYCIP, to pull down AIRE-PHD1 protein, used as a positive control. A total protein lysate containing AIRE-PHD1 was also loaded into the SDS-PAGE to confirm its overexpression in the sample. All blots were probed in the c-Myc antibody. According to the western blot results in *Figure 10* below, we did not manage to confirm any of the interactions using this method.

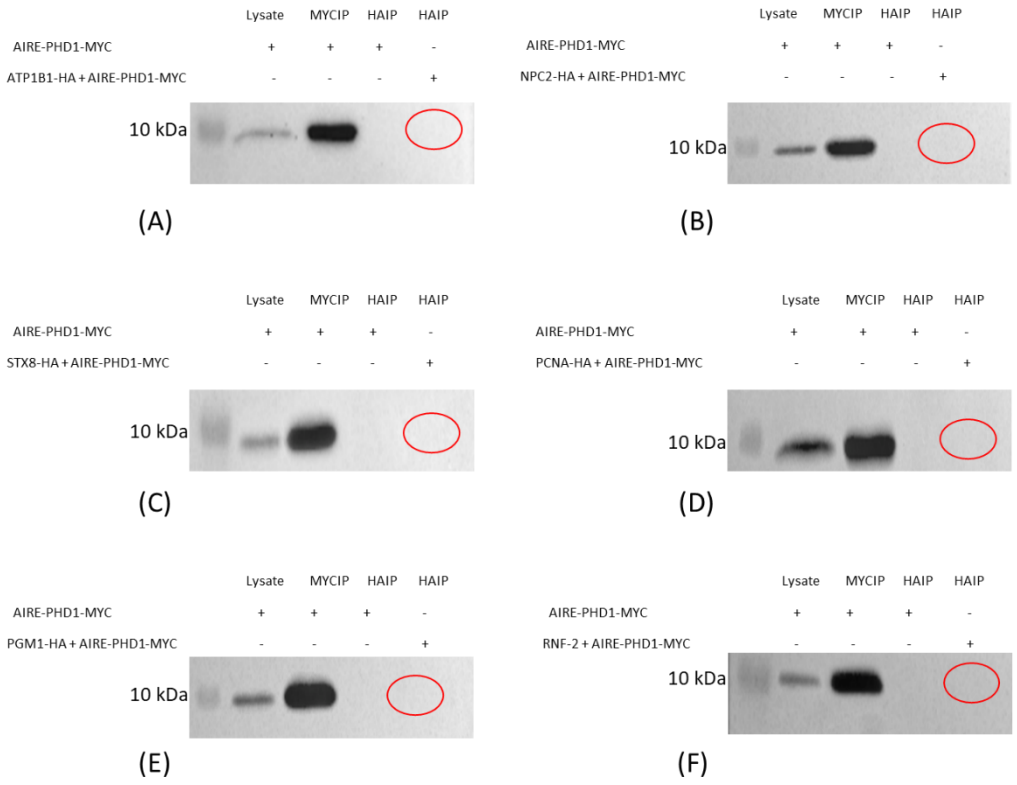


Figure 10: Western blot detection of co-immunoprecipated AIRE-PHD1 with each of the following proteins: PCNA-PGM1-RNF2-NP2-ATP1β1-STX8. co-IP was carried out using HA-tagged beads. Membranes were immunoblotted using the c-Myc antibody. Interaction of MYC-tagged AIRE-PHD1 with the HA-tagged prey proteins PCNA-PGM1-RNF2-NP2-ATP1β1-STX8 were validated via co-immunoprecipitation and analysed via western blotting. HA-tagged beads were used to immunoprecipitate the prey proteins and, if interaction existed, pull-down the AIRE-PHD1. The results are shown for direct immunoprecipitation of AIRE-PHD1 using HA-tagged beads (negative control), co-immunoprecipitation of AIRE-PHD1, and overexpression of AIRE-PHD1 total protein lysate (positive control) with a band size of around 10 kD. None of the interactions were observed via co-IP.

Since we did not manage to detect any interaction between AIRE-PHD1 and the identified novel prey proteins, we decided to chemically cross-link the proteins prior to lysis and then carry out co-IP in order to check for interaction. Lysis was carried out as explained previously, and the co-IP experiment was repeated using the same controls explained in the previous section. IP samples were loaded into an SDS-PAGE gel and transferred to iBlot PVDF blotting membrane. The washing and incubation steps were described in the previous section. For protein detection, the horseradish peroxidase (HRP) substrate was added to the membrane. The results, in the following page, indicate that we did not establish any interaction, even after cross-linking the proteins. It might be that this type of test is not suitable for such interactions, especially as it has been confirmed that PCNA interacts with AIRE *[see Figure 11].*

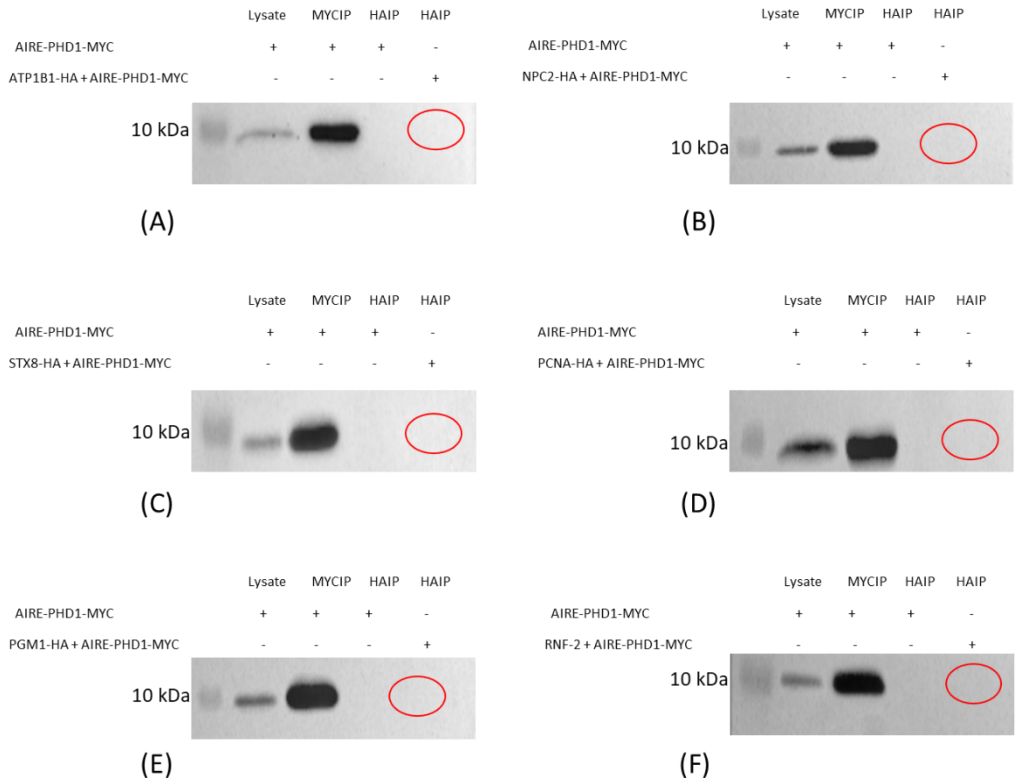


Figure 11: Western blot detection co-immunoprecipated AIRE-PHD1 with each of the following proteins: PCNA-PGM1-RNF2-NP2-ATP1β1-STX8 following cross-linking. Prior to cell lysis, the co-transfected proteins were cross-linked via an Amine-Reactive Crosslinker. co-IP was carried out using HA-tagged beads. Membranes were immunoblotted using the cmyc antibody. Interaction of cmyc-tagged AIRE-PHD1 with the HA-tagged prey proteins PCNA-PGM1-RNF2-NP2-ATP1β1-STX8 were validated via co-immunoprecipitation and analysed via western blotting. HA-tagged beads were used to immunoprecipitate the prey proteins and, if interaction existed, pull-down the AIRE-PHD1. The results are shown for direct immunoprecipitation of AIRE-PHD1 using HA-tagged beads (negative control), co-immunoprecipitation of AIRE-PHD1, and overexpression of AIRE-PHD1 total protein lysate (positive control) with a band size of around 10 kD. None of the interactions were observed via co-IP.

3.7 Discussion

3.7.1 Choosing Y2H screening

To enhance our understanding of the molecular function of AIRE-PHD1, we aimed to identify unknown binding partners using Y2H. This system was used as it allows screening of a cDNA library for interactions with the AIRE-PHD1 domain in an in vivo yeast environment (Luban & Goff, 1995). We used the human normalised library for our screening for several reasons. Firstly, there is evidence that AIRE also exists in the peripheral lymphoid tissues (Shcheglov et al., 2007). Thus, the use of the human universal library allowed us to pick up the proteins that interact with AIRE in other tissues aside from the thymus. Furthermore, the advantage of using normalised libraries is that each cDNA is equally represented regardless of the amount of mRNA's, therefore avoiding having multiple colonies bearing the same interacting prey protein and also to maximise the likelihood of detecting low-level expressed protein (Shcheglov et al., 2007).

3.7.2 Interpretation of Y2H screening results

This system allowed us to identify six potential partners of AIRE: namely, PCNA, ATP1B1, NPC2, PGM1, RNF2, and STX8. PCNA is a protein involved in transcription and was previously identified by Abramson et al. (2010) as interacting with AIRE. We detected this interaction when we used the PHD1 domain of AIRE, suggesting that the interaction is specific to PHD1 rather than full-length AIRE, as identified by Abramson's group. The interaction occurred with PCNA specifically from the amino acid aspartic acid at position 150 to serine at position 261. Thus far, we are confident that this represents a true interaction, as it has been previously been identified by Abramson's group. PCNA interacts with proteins that are included in the cell-cycle progression,

particularly in DNA replication, DNA excision repair, and in the assembly of chromatins (Maga & Hübscher, 2003).

Our results show that RNF2 interacts with AIRE-PHD1 via glutamine at position 11 to phenylalanine at position 309. It is a ubiquitin protein ligase that facilitates monoubiquitination of lysine-119 of histone H2A and may lead to the repression of PTA expression because H2A ubiquitination is known to repress gene expression. Thus, the RNF2 may act to stall RNAP2 and hence prevent methylation of H3K4. In this case, chromatin is inactivated, causing repression of PTA (Wang et al., 2004).

STX8 is another protein that is more likely to interact with the AIRE protein. The result from the Y2H screening indicate that it interacts with AIRE-PHD1 from methionine at position to asparagine at position 236. it is a family member of SNARE, which is expressed in lytic granules and is co-localised with the T cell receptor (TCR) upon formation of immunological synapse. It has also a role in protein trafficking from early to late endosomes via vesicle fusion and exocytosis. It also functions as a ubiquitin protein ligase binding (Bhat et al., 2016). In addition, STX8 is involved in the SNARE protein machinery, which also includes other proteins such as VAMP8, vti1b and syntaxin 7. These proteins are implicated in the maturation of T lymphocytes, as they organise the thymic epithelial cells and hence the proliferation and apoptosis of developing T lymphocytes (Kanwar et al., 2008). Hence, STX8 might interact indirectly with AIRE to aid in orchestrating the thymic structure (Passos et al., 2015)

NPC2 is another protein candidate that interacted with AIRE-PHD1 via the amino acid proline at position 90 to the amino acid leucine at position 151. it is

a protein localised in the endosome and functions to efflux the cholesterol from lysosomes. Mutations in this gene have been associated with Niemann-Pick disease, type C2 protein (Rosenbaum and Maxfield, 2011). It has been claimed that NPC2 interacts with LIM-domain binding protein 3 (LDB3) and joins protein kinase C to the cytoskeleton (Huttlin et al.). LDB3 interacts with ATP5f1, a protein that catalyses ATP synthesis bring about an electrochemical gradient of protons through the inner membrane during oxidative phosphorylation. (Wan et al., 2015) mentioned that ATP5F1 interacts with ATP1B1, another candidate that is suspected to interact with AIRE-PHD1. In the Y2H screening, it interacted with AIRE-PHD1 via the amino acid cysteine at position 159 to valine at position 297. It belongs to the family of Na+/K+ and H+/K+ ATPases and to the subfamily of Na+/K+-ATPases. It is a fundamental membrane protein responsible for maintaining the electrochemical gradients of Na+ and K+ ions across the plasma membrane, which are critical for osmoregulation, sodium-coupled transport of a variety of molecules, and for excitability of nerve and muscle. At the protein level, it interacts with ATP4A; a proton pump that mediates the hydrolysis of ATP coupled with the exchange of H+ and K+ across the plasma membrane and is thought to be regulated the AIRE protein (Bab-Dinitz et al., 2009). In addition, PGM1 was identified to interact with AIRE-PHD1 glutamic acid at position 272 to threonine at position 365. It is an enzyme that catalyses the bi-directional interconversion of glucose 1-phosphate (G-1-P) and glucose 6-phosphate (G-6-P).

In the first direction, G-1-P secreted from sucrose catabolism is converted to G-6-P, which is the first intermediate in glycolysis. The other direction involves the conversion of G-6-P to G-1-P, a substrate for synthesis of UDP-glucose, which

is central for synthesis of glycoproteins (Boros et al., 2002). It had been hypothesised that the former enzyme is post-translationally modified by cytoplasmic glycosylation and is involved in the localisation of the protein (Dey et al., 1994). Antigen receptors on T cell (TCR) and B cell (BCR) receptors, as well as major histocompatibility complex (MHC), are glycoproteins. The T cell co-receptors CD4 and CD8, which are critical in dictating T cells' fates including the receptors CTLA-4, are glycoproteins, and their expression and function are dependent upon normal glycosylation. Therefore, it might be that AIRE interacts with PGM1 directly or indirectly, thus affecting the fate of glycosylation, which might influence the T cell differentiation in vivo (Lyons et al., 2015). At this stage, it is difficult with the current data to identify which interactions are true and which interactions are false positive. However, we are confident in stating that the Y2H technique has been validated due to the fact that we were able to extract one of the proteins, PCNA, that has been identified before as interacting with AIRE. For more details see Chapter 5.

3.7.3 Validation of Y2H screening results

Prior to validation, the protein expression of the identified prey protein was confirmed in western blot. Coimmunoprecipitation experiments were carried out to validate the screening results. We did not manage to detect any interaction between AIRE-PHD1 and the identified novel prey proteins. Consequently, the proteins were chemically cross-linked prior to lysis. This technique involves the formation of covalent bonds between the two proteins via bi-functional reagents containing the reactive end group, homobifunctional N-hydroxysuccinimide ester (NHS ester), which react with functional groups such as primary amines (NH_2) of amino acid residues. Co-IP was utilised to detect the physical interaction of the covalently cross-linked proteins, which was visualised in a western blot. None of the cross-linked proteins seemed to interact. This is possibly due to the nature of interaction, which might be weak and transient, and therefore is difficult to detect in a co-IP experiment.

Chapter 4
Studying the effects of APS-1 mutations on the interaction with AIRE-CARD and SAND domain

4. Studying the effects of APS-1 mutations on the interaction with the AIRE-CARD and SAND domains

4.1 Outline

It is well known that AIRE functions via elongating transcription by unleashing RNA pol 2 transcription factor, and hence activating expression of several thousand genes, coding for TRA and regulating thymic promiscuous gene expression. The mechanism by which it accomplishes this function is poorly understood. Understanding this mechanism is vital, as it would help in generating an APS1-targeted therapy (Ulmanen et al., 2005). APS-1 appears in a straightforward genetic context, and thus provides an attractive model to understand the mechanism underlying autoimmunity (Martino et al., 2016).

Previous studies have determined that AIRE exists as a heteromeric complex in the nucleus, denoting that AIRE is capable of monitoring promiscuous gene expression via recruitment of protein partners. 66.6% of APS-1 mutations occur in the CARD domain and most of the known prey protein interactions occur via this domain (Ramsey et al., 2002). The initial section of this chapter focuses on confirming the interactions of known proteins that interact with AIRE via its CARD domain; namely, Bromodomain-containing protein 4 (BRD4), Death-associated protein 6 (DAXX), and Protein Inhibitor of Activated STAT 1 (PIAS1). Then, we describe how APS-1 mutants T16M, L29P and L93R were introduced to the CARD to test whether these pathogenic mutants affect the interactions (Abramson et al., 2016). The reason for choosing these particular mutations is that each one of them lie in a different compartment of the CARD domain. Thus, it is possible to compare if the location of the mutant could also have an impact on the interaction.

Previously, Halonen's research group used a mammalian-one-hybrid assay to test the ability of wild-type AIRE as well as a number of mutants, such as T16M, to activate the transcription of the reporter gene (Halonen et al., 2004). The T16M mutation is situated in the N-terminal of CARD domain and was used here as a control in the co-IP experiment, as it has no effect on the transactivation activity of AIRE, as revealed by Halonen et al. (2004). The L29 mutation lies at the core of the CARD domain 3D model, and it has been suggested that it can alter the conformation of the helices. It is therefore worthwhile checking whether it affects the interaction of AIRE with the proteins BRD4, DAXX and PIAS1. The L93R mutation, however, is situated at the surface of the CARD 3D model, and it has been suggested that it can disrupt the hydrophobic core of the four-helix bundle of the AIRE-CARD (Ferguson et al., 2008). The location of these mutations was confirmed by Pitkanen et al. (2000) by using a three-dimensional homology model for the CARD domain in order to predict the structural alterations of APS-1 mutants affecting the CARD domain of AIRE. Therefore, it was discovered that these mutations give a good coverage of the CARD domain.

This present thesis examines the SAND domain, due to its vital role in interacting with ATF7ip – linking Aire with further epigenetic processes (Waterfield et al., 2014). Thus, studying the function of this domain is crucial and will aid in further outlining the molecular function of AIRE (Yoshida et al., 2015).

The second section of this chapter thus focuses on confirming the interaction of the protein partner P63, which interacts with AIRE via its SAND domain. Two APS-1 mutants, G228W and P252L, were introduced to the SAND domain to

examine whether these mutants affect AIRE-p63 interaction. The reason for choosing the G228W mutant is because it is the only known APS-1 mutation with a unique mode of inheritance – dominant negative – which causes it to be ideal for the investigation of its effect on interaction with p63. Furthermore, as noted by Tonooka's research group, its interaction with p63 does not fit with how it exerts its affect in a dominant negative model (Tonooka et al., 2009). Therefore, the missense APS-1 recessive mutation P252L was used here as a control to test whether this mode of transmission behaves in the same way as with the dominant mutant G228W in terms of interaction with p63 (Ilmarinen et al., 2005).

Experiments described in this chapter involved the cloning of CARD and SAND cDNA sequences into the pcDNA5/FRT plasmid. Sanger sequencing was used to confirm the cloning prior to carrying out co-IP. Once the interactions were validated, site-directed mutagenesis was used to introduce a single-substitution APS-1 mutation into each of the cloned domains. Co-IP was then utilised again to detect any changes in interaction as a result of introducing mutants. Any detected effect on interaction as a result of introducing APS-1 mutations was further characterised by studying the functional impact of the interaction by using the mammalian-2-hybrid (M2H) system, which translates the potency of interaction as a dual-luciferase measure. Identifying known AIRE interactions affected by APS-1 mutations will shed light on the molecular mechanism underlying the pathogenesis of APS-1 and may lead to further understanding of autoimmunity (See Chapter 5 for details).

4.2 Amplification of AIRE cDNA domains

4.2.1 Amplification of CARD AIRE cDNA

The sequence encoding the CARD domain was amplified by PCR from pcDNA/FRT AIRE cDNA using domain-specific primers that allowed the introduction of BamH1 and NotI1 restriction sites at the domain's 5' and 3' ends respectively. A fragment of 300 bp was generated, as shown in *Figure 12*.

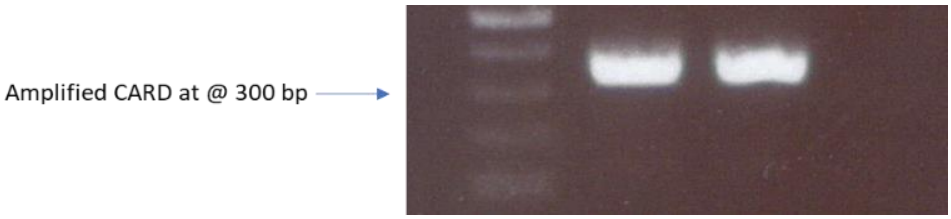


Figure 12: AIRE-CARD amplified. The CARD domain was amplified using specific primers at an annealing temperature of 61°C. The size of band, in duplicate, is approximately 300bp corresponding to the amplified AIRE-CARD sample.

4.2.2 Amplification of SAND AIRE cDNA

The sequence encoding the SAND domain of AIRE protein was cloned into pcDNA5/FRT. The SAND domain was then amplified by PCR from pcDNA/FRT AIRE cDNA using domain-specific primers that allowed the introduction of Nhe1 and NotI1 restriction sites at the domain's 5' and 3' ends respectively. A fragment of 600 bp was generated, as shown in *Figure 13.*

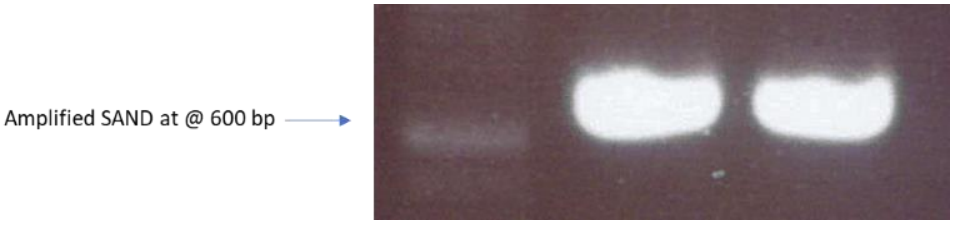


Figure 13: AIRE-SAND amplified. The SAND domain was amplified using specific primer at an annealing temperature of 63°C. The size of band, in duplicate, are around 600bp, corresponding to amplified AIRE-SAND sample.

4.3 Digestion of pcDNA5/FRT

The plasmids were digested using the restriction enzymes Nhe1 and Not1 for SAND, and BamH1 and Not1 for CARD. The sample was electrophoresed on a gel and the resultant bands were excised and purified *[see Figure 14].* Each digested-purified insert was ligated to the linearized empty pcDNA5/FRT and incubated with T4 DNA ligase. The resultant products were then transformed in E. coli. Plasmids were then isolated, and the cDNA inserts were verified by sequencing

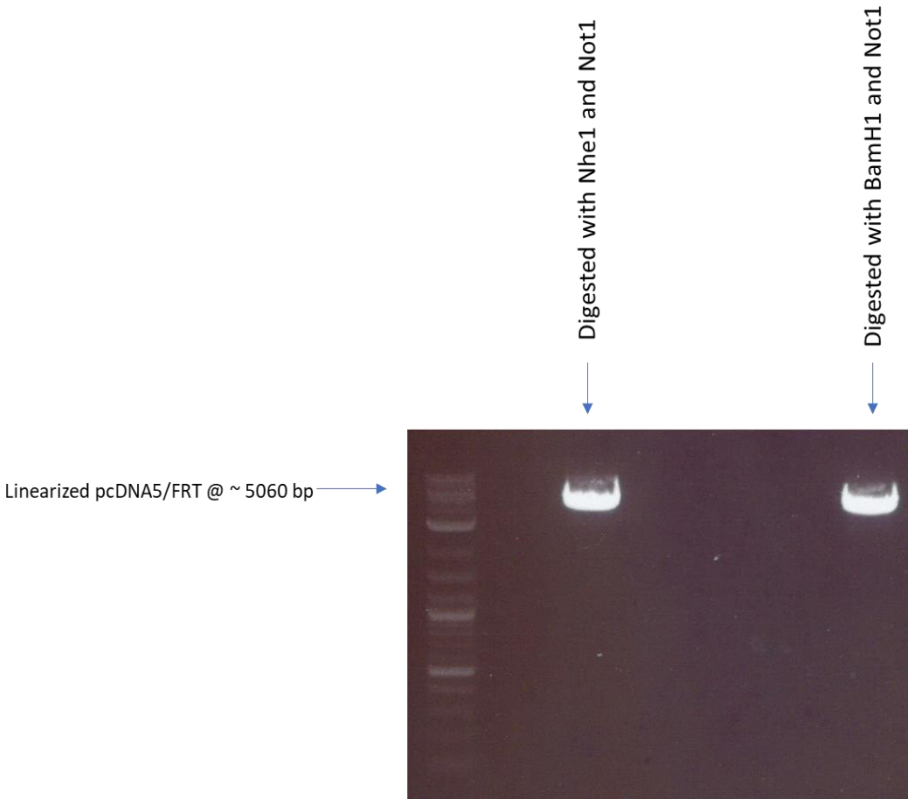
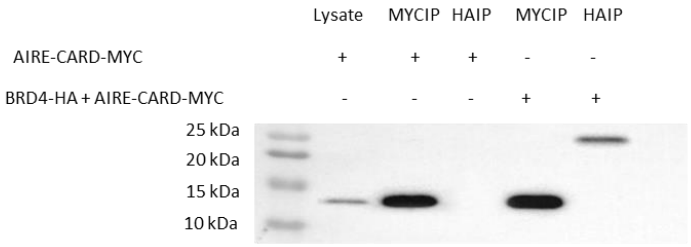


Figure 14: pcDNA5/FRT digested. Two samples of plasmid were digested: one of them was with the BamH1 and the Not1 restriction enzyme, and the other one was with the Nhe1 and the Not1 enzymes**.**

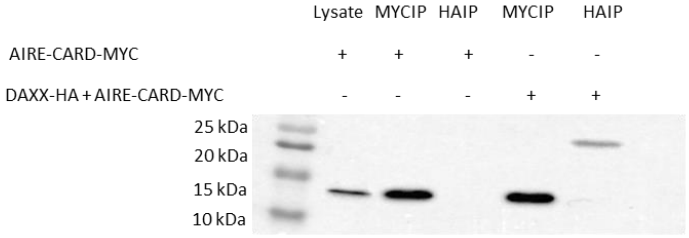
4.4 Co-immunoprecptation

In order to confirm the interaction of AIRE-CARD protein with BRD4, DAXX, and PIAS1 proteins, a co-IP experiment was conducted. Initially, the HA tagged-cDNA constructs were individually co-transfected with the cmyc-tagged AIRE-CARD construct in HEK-293 cells seeded in trans-well plates. As a control, the cmyc-tagged AIRE-CARD was transfected in a separate well. 24 hours post transfection, cells were lysed, and proteins were quantified. Co-IP was carried out by incubating the total protein cell lysate separately with either HA-tagged dyna beads or CMYC tagged dyna beads as a control.

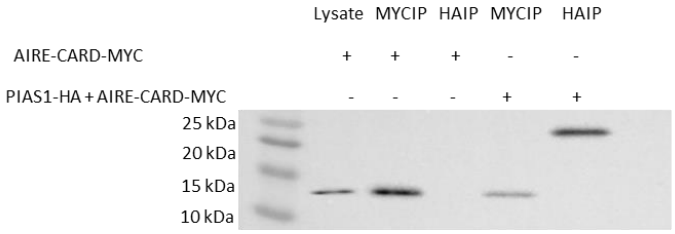
In Figures 15 A, B and C, from the left to right, the bands correspond to the AIRE-CARD total protein lysate (12 kD), direct cmyc-IP band of AIRE-CARD protein used as a positive control (12 kD), no band showing for Ha-IP AIRE-CARD protein (negative control), direct cmyc-IP band of protein-HA + AIRE-CARD-MYC band, another positive control (12 kD), and HA- coIP of protein-HA + AIRE-CARD-MYC band (25kD). Thus, we can confirm that AIRE interacts with BRD4, DAXX, and PIAS1 via its CARD domain. The band is approximately 25kD, suggesting that the binding to the proteins promotes dimerisation of CARD.



(A)



(B)



(C)

Figure 15: Blots presenting the co-IP of AIRE-CARD against BRD4(a)-DAXX(b)-PIAS1(c). Co-IP was carried out using HA-tagged beads. Membranes were immunoblotted using cmyc antibody. Interaction of cmyc-tagged AIRE-CARD with the HA-tagged prey proteins BRD4-DAXX-PIAS1 were validated via co-immunoprecipitation and analysed via western blotting. HA-tagged beads were used to immunoprecipitate the prey proteins. The results are as follows: overexpression of AIRE-CARD total protein lysate (12kD); direct immunoprecipitation of AIRE-CARD using cmyc-tagged beads (positive control) (12kD); direct immunoprecipitation of AIRE-CARD using ha-tagged beads (negative control); coimmunoprecipitation of AIRE-CARD via ha-tagged beads (25kD). MYCIP = protein pulled down with cmyc-tagged beads. HAIP = protein pulled down by HA-tagged beads.

Having confirmed the interaction between AIRE-CARD and BRD4, DAXX, PIAS1, we introduced the APS-1 mutations T16M, L29P, and L93R via site-directed mutagenesis, and the interaction was measured again via co-IP. In this experiment, each cmyc-tagged CARD cDNA mutant was individually co-transfected with each of the HA- tagged cDNA in HEK293 cells seeded in well plates. Cells were lysed and proteins were quantified. Therefore, an equal amount of IP lysate was either pulled out with HA-tagged beads (positive control) or CMYC-tagged beads. IP samples were heated to dissociate the proteins from its beads and were equally loaded in SDS-PAGE. In Figures16 A1, B1, and C1, the membrane was hybridised with an HA antibody in order to detect proteins captured by direct-IP. From left to right, the first four bands correspond to the direct-IP protein whilst the second four bands represent the total lysates of each protein (BRD4 ~100kD, DAXX ~ 120 kD, and PIAS1 ~ 75 kD). In *Figures 16 A2, B2, and C2*, the membrane was hybridised with a CMYC antibody in order to detect proteins captured by HA co-IP. From left to right, the first four bands correspond to the pulled-out proteins from co-IP samples whilst the second four bands represent the total lysates of the AIRE-CARD wild-type protein and its mutants; T16M, L29P, and L93R (CARD wild-type and its mutants' direct IP gave a band size ~ 12 kD; CARD wild-type and its mutants gave a band size co-IP ~ 25kD). We were unable to observe an effect on interaction upon introduction of these APS-1 mutations.

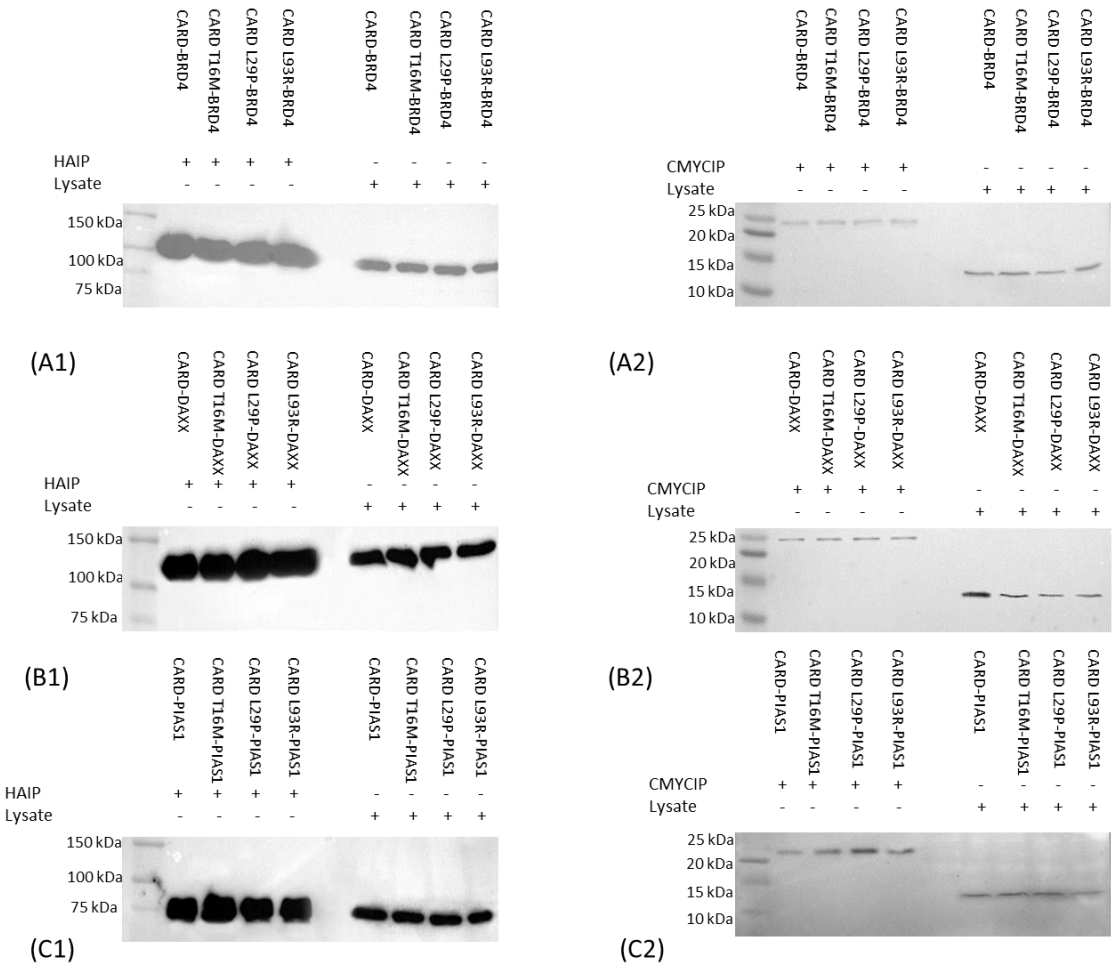


Figure 16: Blots presenting the co-IP of AIRE-CARD mutants against BRD4(a)-DAXX(b)-PIAS1(c). Co-IP was carried out using HA-tagged beads. Interactions of cmyc-tagged AIRE-CARD mutants with the HA-tagged prey proteins BRD4-DAXX-PIAS1 were validated via co-immunoprecipitation and analysed via western blotting. HA-tagged beads were used to immune-precipitate the prey proteins. The results are shown on blots A1-B1-C1: direct immunoprecipitation of BRD4-DAXX-PIAS using HA-tagged beads (positive control); overexpression of BRD4-DAXX- PIAS1 total protein lysate. Blots A2-B2-C2 represent coimmunoprecipitation of AIRE-CARD, AIRE-T16M, AIRE-L29P and AIRE-L93R as a dimer via ha-tagged beads and bands of overexpressed total protein lysate AIRE-CARD, AIRE-T16M, AIRE-L29P, AIRE-L93R.

The second most important domain that we selected to study is the SAND. Firstly, we confirmed that the protein p63 interacts with the AIRE protein through its SAND domain via co-IP. The cmyc-AIRE-SAND/HA-p63 protein lysate was pulled out using HA-tagged dynabeads, heated for 5 min at 93^{o}C to dissociate the IP and loaded in SDS-PAGE gel. The membrane was hybridised using cmyc rabbit antibody in order to detect the AIRE-SAND protein. In *Figure 17*, the order of bands from left to right indicate the following: total protein lysate of cmyc-AIRE-SAND (20kD + 25 kD), cmyc-AIRE SAND pulled down by cmyc antibody-tagged beads (20kD + 25 kD) used as a positive control; cmyc-AIRE SAND pulled down by HA antibody-tagged beads (negative control); cmyc-AIRE SAND + Ha-P63 protein pulled down by cmyc-tagged beads (20kD + 25 kD) used as a positive control; cmyc-AIRE SAND + Ha-P63 protein pulled down by HA-tagged beads (25kD). Only SAND's top band appears to interact with p63. It might be that SAND interacts with p63 via its post-translated modified (PTM) form. This will be further investigated in the next section.

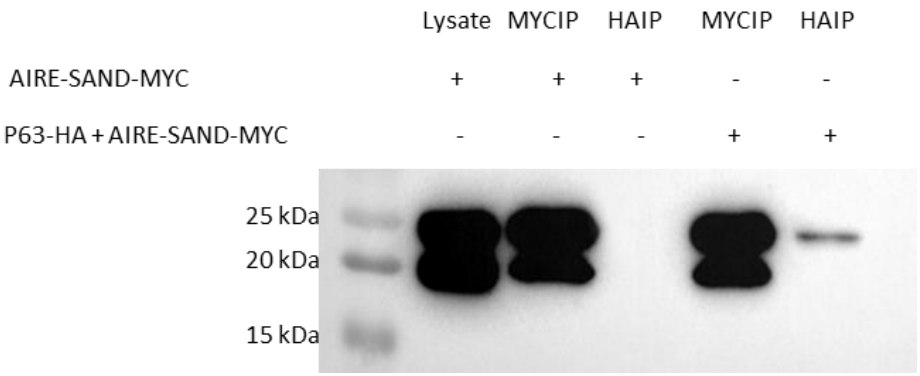


Figure 17: Blots presenting the co-IP of AIRE-SAND against P63. Co-IP was carried out using HA-tagged beads. Membranes were immunoblotted using cmyc antibody. Interaction of cmyc-tagged AIRE-SAND with the HA-tagged prey proteins P63 were validated via co-immunoprecipitation and analysed via western blotting. HA-tagged beads were used to immunoprecipitate the prey proteins. The results are as follows: overexpression of AIRE-SAND total protein lysate (20kD and 25kD); direct immunoprecipitation of AIRE-SAND using cmyc-tagged beads (20kD + 25kD) (positive control); direct immunoprecipitation of AIRE-SAND using HA-tagged beads (negative control); direct immunoprecipitation of AIRE-SAND using cmyc-tagged beads (20kD + 25 kD) (positive control); coimmunoprecipitation of AIRE-SAND via HA-tagged beads (25kD), confirming the interaction of p63 with AIRE-SAND.

4.5 Post-translational modification detection test

Since there is an obvious difference between the bands of SAND as a total protein lysate and upon protein interaction, we decided to conduct certain common PTM tests, including checking for phosphorylation and glycosylation. Firstly, we decided to dephosphorylate the AIRE-SAND to check whether interaction occurs via the phosphorylated site(s) of AIRE-SAND. Thus, the AIRE-SAND was combined with protein metallophosphatases, $MnCl_2$, and Lambda protein phosphatase, and then was incubated. A control sample was prepared for comparison, omitting the enzyme lambda protein phosphatase. Both samples were tested on SDS-PAGE gel giving 2 band sizes (20kD+25 kD) [see *Figure 18A1*]. The top band did not disappear after carrying out the dephosphorylation test, delineating that the AIRE-SAND protein did not undergo phosphorylation.

We then moved forward to test if the interaction occurs via the glycosylated site(s) of AIRE-SAND. Thus, the AIRE-SAND was combined with a glycoprotein denaturing buffer and allowed to be denatured by heating for 10 minutes. The protein was chilled, and the following components were added to the mixture; GlycoBuffer, NP-40, and the enzyme Peptide:N-glycosidase F (PNGase F) incubated at 37^{o}C. A control sample was also included for comparison, omitting the enzyme PNGase. Both the deglycosylated and the control sample were tested on SDS-PAGE gel (20kD + 25 kD) *[see Figure 18A2].* Similarly, the top band did not disappear after undertaking the deglycolysation test, revealing that the AIRE-SAND protein did not undergo glycolysation.

(A1) (A2)

Figure 18: Blots representing a treated AIRE-SAND protein with either lamba protein phosphatase or PNGase F. The AIRE-SAND was tested with specific enzymes to check whether it undergoes post-translational modifications (PTM). In A1 blot, the AIRE-SAND protein was treated with the enzyme lambda protein phosphatase for 30 min at 30°C. A control sample was also treated in the same conditions, excluding the enzyme (20kD + 25kD). In A2 blot, the AIRE-SAND protein was treated with the enzyme PNGase F for 1 hour at 37°C. A control sample for both tests were also treated in the same conditions, excluding the enzyme. The top band did not disappear post treatment, denoting that the AIRE-SAND protein did not undergo PTM in this case: glycosylation and phosphorylation (20 kd + 25kD).

4.6 Site-directed mutagenesis and coimmunoprecipitation

Having successfully confirmed AIRE-SAND/p63 interaction, the next step was to test whether specific mutations that exist in the SAND domain have an impact on the interaction. Therefore, we introduced the missense APS-1 mutation, G228W, which is a unique pathogenic mutation that acts in a dominant negative mode of inheritance. We also introduced another missense APS-1 mutation, P252L, with a recessive mode of inheritance. Co-IP and western blot were carried out in the same manner as previously mentioned. *Figures A1* and *B1* show the direct HAIP of the p63 protein along with the p63 total protein lysate. These membranes were probed with an HA antibody (p63 ~ 80 kD). In contrast, *Figures A2* and *B2* represent the co-IP of wt-SAND, SAND-G228W, and SAND-P252L via HA-tagged beads along with their total protein lysates. These membranes were probed with a cmyc antibody (wt- SAND and SAND-G228W size ~ 25 kD). Differences in interaction were only detected with the SAND-G228W mutation, reflecting its dominant negative manner. To confirm our finding, we conducted two other repeats for a total of n=3, which produced the same results [*See appendix section 7.16]*.

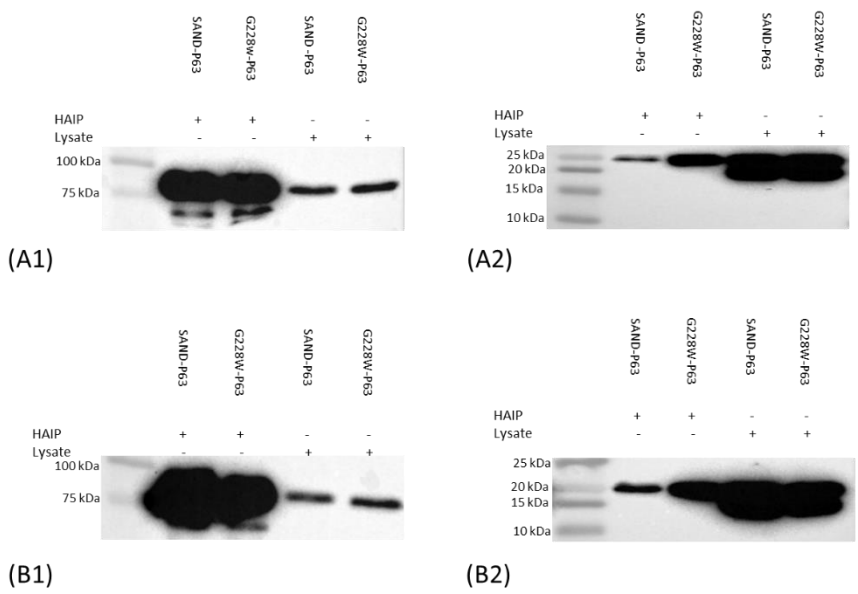


Figure 19: Blots presenting the co-IP of AIRE-SAND mutants against P63. Co-IP was carried out using HA-tagged beads. Interaction of cmyc-tagged AIRE-SAND mutants with the HA-tagged prey protein P63 was validated via co-immunoprecipitation and analysed via western blotting. HA-tagged beads were used to immunoprecipitate the interacting protein. The results are shown in blots A1 and B1: direct immunoprecipitation of P63 using HA-tagged beads (positive control) (size ~ 80 kD); overexpression of P63 total protein lysates (size ~ 80 kD). Blots A2 and B2 represent coimmunoprecipitation of wt-SAND, SAND-G228W, and SAND-P252L, which were immunoblotted using cmyc antibody (size ~ 25 kD): overexpression of total protein lysate wt-SAND, SAND-G228W and SAND-P252L (control) (size 20kD and 25 kD).

The results of the previous co-IP experiment of p63 interacting with either wt-SAND , SAND-G228W or SAND-P252L were analysed in Image J software and the results were displayed in GraphPad Prism 8 software. In Image J, the band intensity for the coIP samples wt-SAND , SAND-G228W and SAND-P252L were quantified individually and normalised to the total lysate samples of each. The relative band intensities for both samples were displayed in a bar chart using GraphPad Prism, and a two-way ANOVA test was used to calculate the p-value. We did not establish any difference between the relative intensity of IP with p63 among SAND-P252L and the control protien wt-SAND , which was expected (P value > 0.05). However, we found a higher relative intensity of IP when p63 interacted with SAND-G228W than with wt-SAND (P value < 0.0001). Thus, the SAND-G228W has a significantly higher affinity with p63, suggesting that this type of mutation fits with the model of the dominant negative inheritance as the mutant is stronger in terms of competing against wt-SAND to bind to p63 *[see Figure 20].* This result is encouraging, which led us to ask whether these differential interactions affect the transcriptional acitivity of the SAND domain of the AIRE protein. To examine that, we carried out a mammalian-2-hybrid test, which was translated through the measurement of the dual-luciferase activity. Details will be outlined in the following section *[see Figure 24].*

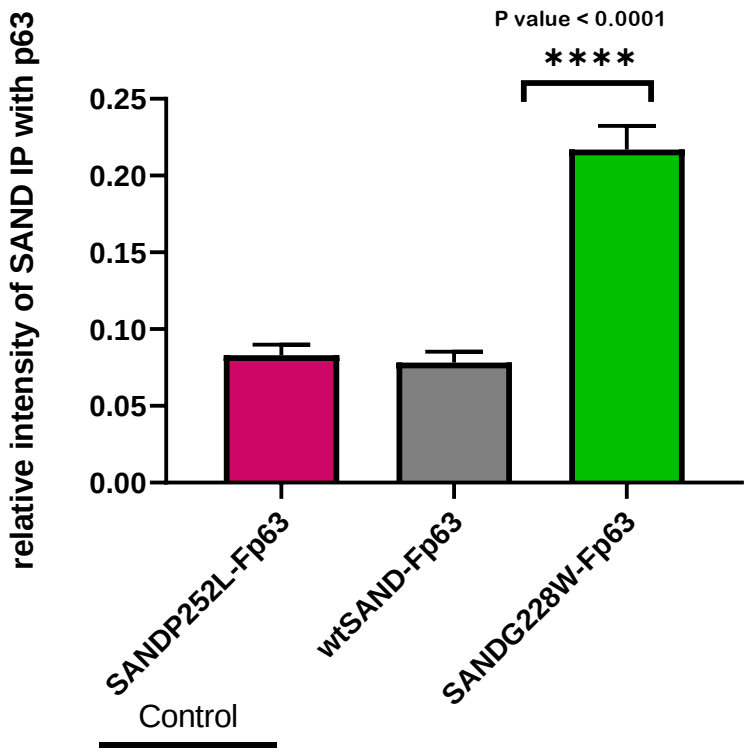
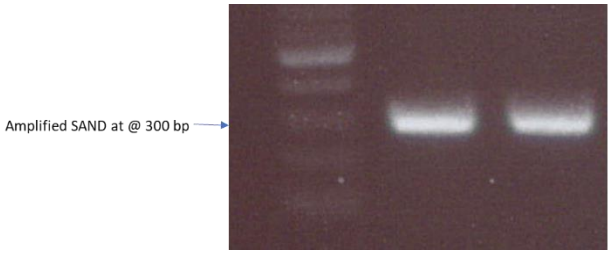
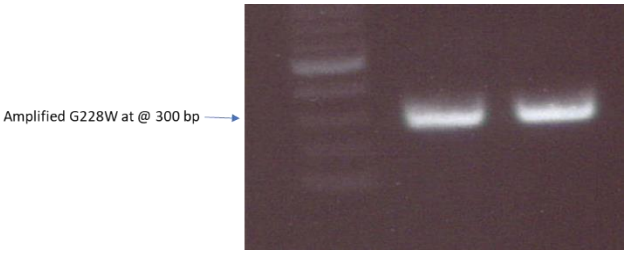


Figure 20: A bar-chart presenting the relative intensity of SAND IP with the protein p63. The bar chart results reflect the relative intensity of SAND-P252L (control), wtSAND, and SAND-G228W with the protein p63. Co-IP results for wt-SAND and SAND-G228W were repeated 3 times (n=3), and the band intensity was measured and normalised to the total protein lysate of each (P value <0.0001). The relative intensity of the control sample, SAND-P252L, was only measured once (n=1).

4.7 Amplification of SAND AIRE cDNA and G228W AIRE cDNA

To the measure the dual-luciferase activity, first we had to clone the sequence encoding SAND domain of the AIRE protein and its mutant W228 into the luciferase vector pBIND. They were amplified by PCR using domain-specific primers that allowed the introduction of Mlu1 and Kpn1 restriction sites at the domain's 5' and 3' ends that are also situated at the multiple cloning site of the pBIND vector *[see Figure 21].*





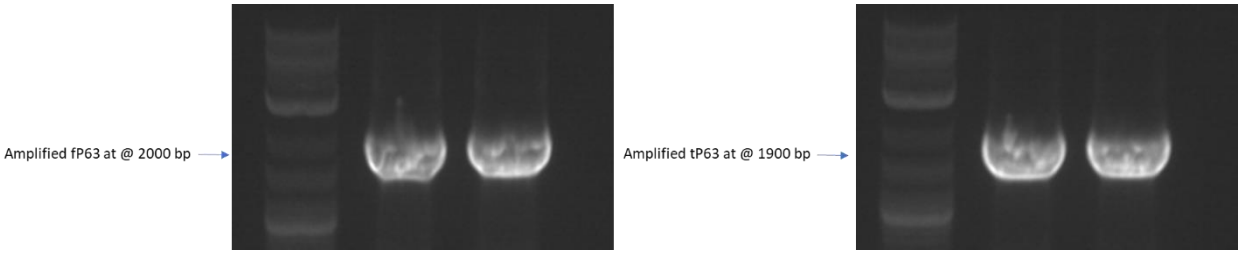
(21A) (21B)

Figure 21A: AIRE-SAND amplified. The SAND domain was amplified using specific primers at an annealing temperature of 63°C. The size of bands is approximately 600bp.

Figure 21B AIRE-G228W amplified. The SAND mutant domain was amplified using specific primer at an annealing temperature of 63°C. The size of bands is approximately 600bp.

4.8 Amplification of p63 cDNA

The full-length p63 cDNA and its truncated version, which excludes the TA domain, were cloned into the luciferase vector pACT. They were amplified by PCR using domain-specific primers that allowed the introduction of BamH1 and Not11 restriction sites at the domain's 5' and 3' ends that are also situated at the multiple cloning site of the pACT vector *[see Figure 22].*



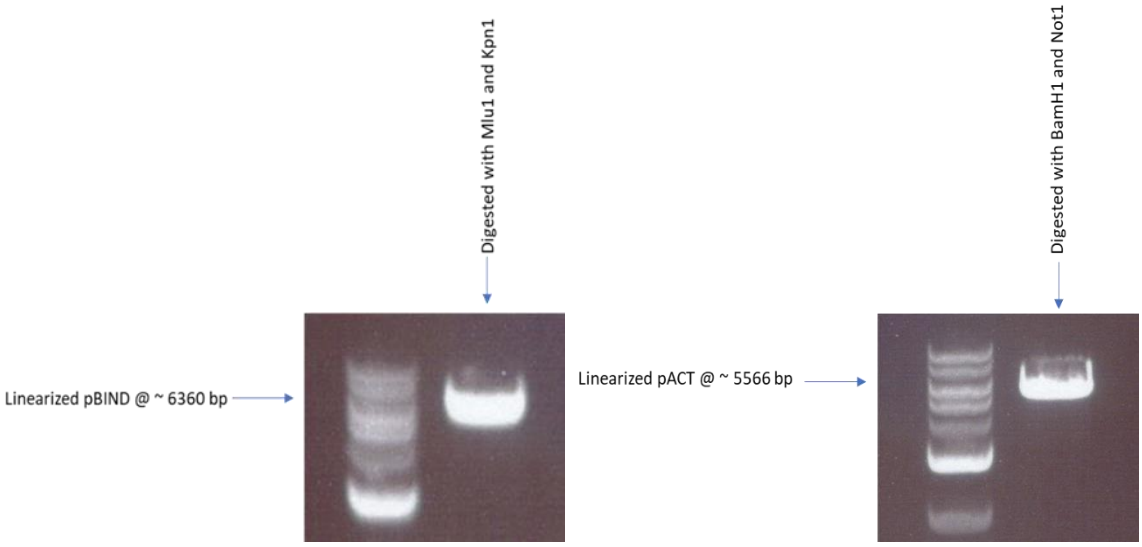
(22A) (22B)

Figure 22A: full-length p63 amplified. The full length p63 cDNA was amplified using specific primers at an annealing temperature of 63°C. The size of bands is approximately 2000bp.

Figure 22B: truncated p63 amplified. The truncated p63 cDNA was amplified using specific primer at an annealing temperature of 63°C. The size of bands is approximately 1900bp.

4.9 Digestion of pACT and pBIND

Prior to ligation and transformation, the plasmids and the inserts had to be digested using the required restriction enzymes. The sample was electrophoresed on a gel and the resultant bands were excised and purified *[see Figure 23].*



(23A) (23B)

Figure 23A: pBIND digested. The pBind was digested with the restriction enzymes Mlu1 and Kpn1.

Figure 23B: pACT digested. The pACt was digested with the restriction enzymes BamH1 and Not1.

4.10 Testing the functional interaction of AIRE-SAND/p63 using the mammalian-2-hybrid system

In order to test the transcriptional activity of the interaction, the mammalian two-hybrid system was used to semi-quantify the interaction between wt-SAND and its mutant SAND-G228W against the full-length (f) and truncated (t) p63 via the dual-luciferase reporter assay. HEK293 cells were seeded in trans-well plates, and each well was transiently transfected with the luciferase plasmids. 24 hours post-transfection, the cells were lysed with a passive lysis buffer and the activities of firefly luciferases were measured sequentially. As expected, we found a strong interaction between wt-SAND and F63, but this was significantly lower than the positive control MyoD and ID. In contrast, this interaction was significantly higher than the negative controls (P value < 0.0001). However, the interaction was abolished when wt-SAND interacted with the truncated version of p63 (Tp63), which is comparable to the negative controls (P value < 0.0001). This in turn, reflects the importance of the TA domain for p63 interaction. Interestingly, when the mutant SAND-G228W interacted with either Fp63 or Tp63, the interaction was also abolished, indicating that this mutation plays the same role as the Tp63 in impeding the interaction (P value < 0.0001) *[see Figure 24].*

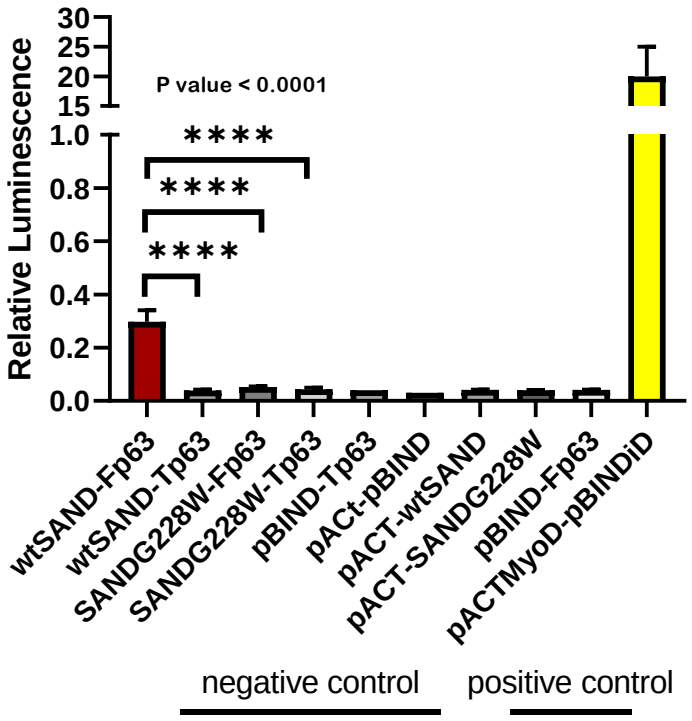


Figure 24: Dual-luciferase activity results of SAND and its mutant, G228W, against p63. The mammalian two-hybrid system Promega was used to semi-quantify the interaction between wt-SAND and its mutant G228W against and full-length (f) and truncated (t) p63 via the dual-luciferase reporter assay. We found a strong interaction between wt-SAND and F63, but this was significantly lower than the positive controls, MyoD and ID. Conversly, this interaction was significantly higher than the negative controls (P value < 0.0001). However, the interaction was abolished when wt-SAND interacted with Tp63, which is comparable to the negative controls (P value < 0.0001), reflecting the importance of the TA domain for p63 interaction. Moreover, when the mutant SAND-G228W interacted with either Fp63 or Tp63, the interaction was also abolished, indicating that this mutation plays the same role as the Tp63 in impeding the interaction (P value < 0.0001). Data shown in the bar chart represent the mean firefly luciferase activity normalised to the renilla luciferase control of four independent replicates, each performed in triplicates. The asterisks indicate statistical significance and P values from the two-way ANOVA test.

4.11 Amplification of AIRE cDNA and digestion of pBIND

We then moved on to test if the SAND-p63 interaction is affected by the other domains of AIRE. In order to test the AIRE-p63 interaction, we cloned the full length AIRE into the luciferase vector, pBIND, and then we introduced the APS-1 dominant negative mutation, G228W, via site-directed mutagenesis. The sample was amplified by PCR using domain-specific primers that allowed the introduction of Mlu1 and Not1 restriction sites at the domain's 5' and 3' ends that are also situated at the multiple cloning site of the pBIND vector *[see Figures 25 and 26]*.

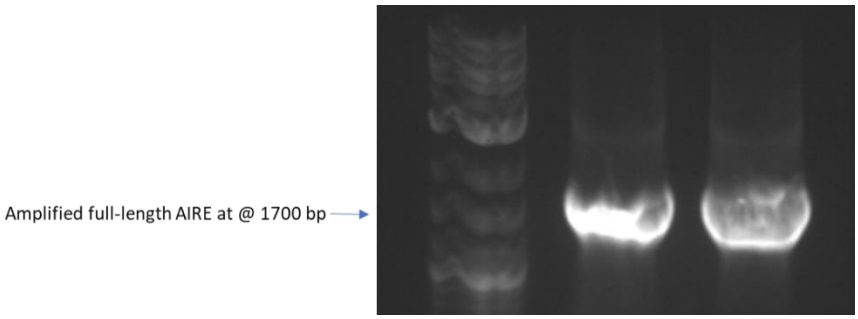


Figure 25: full-length AIRE amplified. The full-length AIRE cDNA was amplified using domain-specific primers at an annealing temperature of 57^{o}C. The size of bands is approximately 1700bpPrior to ligation and transformation, the pBIND plasmid and the amplified full-length AIRE insert had to be digested using the restriction enzymes Mlu1 and Not1.

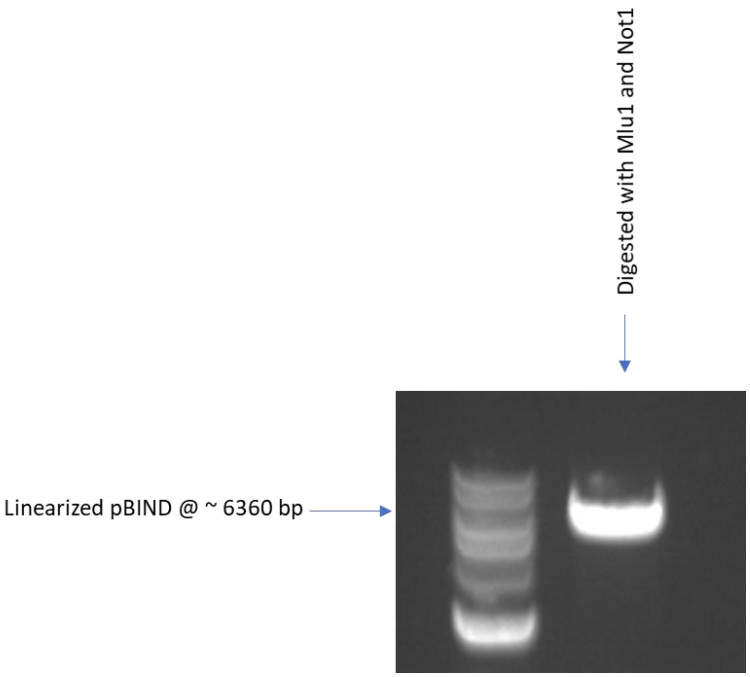


Figure 26: pBIND digested. The pBind plasmid was digested with the restriction enzymes Mlu1 and Not1.

4.12 Testing the functional interaction of full-length and mutated AIRE proteins using the mammalian-2-hybrid system

Once we confirmed the cloning of the full-length AIRE into the pBIND, we tested the interaction of wt-AIRE and AIRE-G228W against both the full-length and the truncated p63 to determine if the other domains of AIRE have a role in strengthening the interaction with p63. Therefore, we measured the transcriptional activity of the interaction using the mammalian-2-hybrid system, as discussed in the previous section. As mentioned before, the activity of the firefly luciferase reporter was quantified and then quenched so that the Renilla luciferase activity could be measured. The trancriptional activity of the full length AIRE interacting with Fp63 was significantly higher than when F63 interacts with the SANDdomain, suggesting the participation of other AIRE domains in enhancing the interaction (P value < 0.0001). However, this interaction was significantly lower than the positive control results (P value < 0.0001). In contrast, when either the full-length AIRE or the AIRE-SAND interacted with Tp63, the reaction was dimished, confirming the role of the TA domain of p63 in facilitating the interaction. This result is also comparable with the negative control results (P value < 0.0001). The interaction was also abolished when wt-AIRE or wt-SAND interacted with Fp63, which further confirms the role of this mutant, which has the same effect as when we truncated the TA from p63 (P value < 0.0001) *[see Figure 27].*

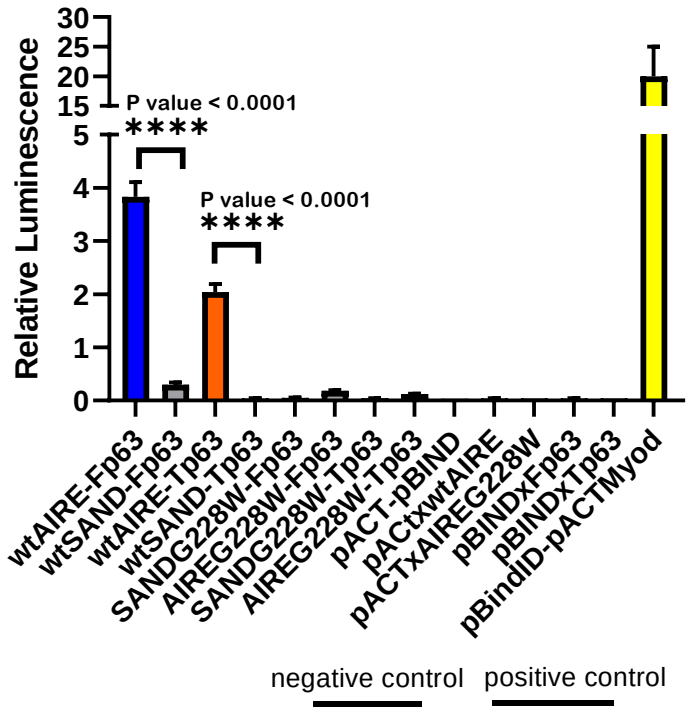


Figure 27: Dual-luciferase activity results of AIRE and the mutant G228W against p63. The mammalian two-hybrid system Promega was used to semi-quantify the interaction using the dual-luciferase assay between wt-AIRE and its mutant AIRE-G228W against and full-length (f) and truncated (t) p63. We then compared the results with the previous SAND-p63 M2H experiment. The trancriptional activity of wt-AIREG228 interacting with Fp63 was substantially higher than when F63 interacts with the wt-SAND domain, explaining the participation of other AIRE domains in facilitating the interaction (P value < 0.0001). However, this interaction was significantly lower than the positive control results (P value < 0.0001). However, when either wt-AIRE or wt-SAND interacted with Tp63, the reaction was abolished, confirming the role of the TA domain of p63 in facilitating the interaction. This result is also comparable with the negative control results (P value < 0.0001). The interaction was also dimished when AIRE-G228W or SAND-G228W interacted with Fp63, which further confirms the role of this mutant, which has the same activity as when we truncated the TA from p63 (P value < 0.0001). Data shown in the bar chart represents the mean firefly luciferase activity normalised to the renilla luciferase control of four independent replicates, each performed in triplicate. Asterisks indicate statistical significance and P values from the two-way ANOVA test.

4.13 Discussion

Our aim in these experiments was to determine whether APS-1 mutations affect AIRE protein interactions. Initially, we began by confirming that AIRE-CARD interacts with BRD4-DAXX-PIAS1 via co-IP. We confirmed the interactions and observed that the band size of pulled-out AIRE-CARD from a co-IP protein sample is greater than the band size of AIRE-CARD pulled out from a direct IP protein sample, indicating that the binding to the proteins actually promotes dimerisation of CARD. We have also shown that these interactions are not affected by the following recessive mutations: T16M, L29P and L93R *[see Figures 15 and 16]*. The reason for this might be that these mutations are involved in other AIRE-protein interactions, or it could be that the effects of these pathogenic mutants on the mentioned interactions are so weak that they cannot be revealed in a co-IP experiment. As mentioned earlier, the CARD domain is mainly linked to AIRE's dimerisation, so mutants at this domain exhibit very little transcriptional activity, especially if CARD was used on its own rather than the full-length AIRE (Sparks et al., 2016). Therefore, a difference in interaction might be detected if the mutants were introduced to the full-length AIRE rather than the CARD domain itself, as the existence of the remaining domains might aid in facilitating the interaction (Liiv et al., 2008).

We also confirmed that the interaction of the protein p63 with AIRE occurs via the SAND domain. It seems that the interaction only occurs via SAND's top band, as seen in the co-IP experiment *[see Figure 17]*. De-phosphorylation and de-glycolysation tests were applied to the SAND protein to check whether the interaction is via the post-translationally modified SAND protein. The result suggests that this is not the case, as there are not any considerable differences

in bands between the SAND and the post-translationally modified SAND upon interaction with p63 *[see Figure 18]*. However, we managed to demonstrate that the SAND mutant, G228W, interacts more strongly than the wt-SAND with the protein P63 due to its interference with the homodimerisation activity of the wild-type AIRE. It therefore influences the activity of the normal allele, hindering its function (Cetani et al., 2001). The recessive mutant P252L was also tested in terms of its interaction against p63 as a control *[see Figure 18]*. That is because, the mutant G228W follows a more dispersed pattern in the nucleus, whilst P252L exists as concentrated specks in the nucleus as with the wild-type AIRE, reflecting the level of transcriptional activation (Sparks et al., 2016). This is an exciting result, as it contradicts that of (Tonooka et al., 2009) previously. They conducted a co-IP experiment to show the interaction of wt-SAND with p63 and its mutant G228W with p63 and found that the interaction of the mutant with p63 is weaker than the wild-type SAND. Thus, correlates to the function (i.e. translated into functional), which does not make sense. What we found suggests that the interaction is not translated into function, which does give a rational explanation with the dominant negative inheritance of this APS-1 mutation.

Moreover, we wanted to determine the mechanism of interaction of SAND-p63 and to explain the unique model of dominant negative inheritance. Hence, we moved on to study the functional impact of this interesting interaction. To do this, the dual-luciferase activity was measured for the following interactions: wtSAND-p63 and G228W-p63. The results suggest that the interaction of the wt-SAND exerts a higher transcription activity when interacted with p63 than with its mutant, G228W. This in turn reflects the action of this mutant in

impeding AIRE function, hence reducing its transcriptional activation *[*see *Figure 24]* (Sparks et al., 2016). This is in line with the results of Halonen's group, where they visualised the subcellular localisation of wt-AIRE and its mutant, G228W. They cloned the cDNAs in mammalian expression vectors and expressed them in African green monkey kidney (COS-1) cells. These cells were examined using immunofluorescence. The wt-AIRE was localised in both the cytoplasm and the nucleus. In the cytoplasm, it is associated with filamentous structures whereas in the nucleus they exist as nuclear dots (Halonen et al., 2001). In contrast, G228W mutant greatly decreased the association of AIRE with the cytoplasmic filaments as well as in the nuclear dots, where they were detected as clump-liked aggregates (Halonen et al., 2001). Consequently, this mutation restricts the attachment of AIRE to the nuclear dots, leading to intracellular aggregation. This in turn results in the mislocalisation of the AIRE wild-type protein, which has been linked to the inhibition of its transactivation capacity (Sparks et al., 2016).

Our results are in keeping with the hypothesis of dominant negative inheritance effect, by which the G228W protein, through its association with wt-AIRE, prevents it from developing the complexes required for transactivation. Furthermore, we proved that AIRE-SAND binds to p63 via its TA domain (Tonooka et al., 2009). This was confirmed by measuring the functional impact of the interaction with the truncated version of p63 *[see Figure 24]*. According to the dual-luciferase activity results, the interaction was abolished. This implies that p63 which lacks its TA domain cannot bind to AIRE-SAND. We later decided to test the functional activity of the interaction of p63 with AIRE as a whole protein to test whether the interaction is similar. The pathogenic mutant

G228W was also introduced to the full-length AIRE rather than only the SAND domain. We can conclude that we managed to establish a higher interaction when the full-length AIRE interacted with p63 than with the SAND domain on its own, suggesting that other domains of AIRE have a role in prolonging the interaction with the protein p63, which was also observed when the mutant G228W was introduced to the full-length AIRE *[see Figure 27]*. In conclusion, the mutant G228W is a strong competitor to wt-SAND, yet the wild type can still bind to the protein complex, including p63, and be partially active. This explains the mild phenotype that we can see in APS-1 patients bearing the G228W mutant.

Chapter 5
General discussion and future work

5. General discussion and future work

The protein AIRE functions by elongating transcription as it unleashes RNA pol 2 transcription factor and thereby activates expression of thousands of genes that code for PTA and facilitate thymic promiscuous gene expression. However, the mechanism by which it accomplishes this function is still not clear. Characterising this mechanism is crucial, as it would help in producing an APS1-targeted therapy (Ulmanen et al., 2005).

APS-1 appears in a straightforward genetic context, and thus provides a promising model to understand the mechanism underlying autoimmunity (De Martino et al., 2016). The interaction of molecules in monogenic autoimmune diseases is perhaps involved in the same complex as in polygenic disorders, such as the links between Aire-expressing mTECS and the development of Foxp3-expressing regulatory T cells. This is because communications between developing T cells and mTECS thymic stromal cells can drive thymocytes into the Treg cell lineage by signalling through the TCR and accessory molecules, such as CD28, CD40, LFA-1 or B7 , which co-operate with their ligands expressed via mTECS. This activates Tregs and thus drive thymocytes into the Treg cell lineage (Nomura and Sakaguchi, 2007). Consequently, it activates expression of Foxp3, which is needed to control the expression of many other genes and to confer and stabilise the phenotype and suppressive activity of Treg cells. Genetic mutations in FOXP3 result in deficiency in Treg cells and the manifestation of the autoimmune disorder ‘IPEX syndrome’ (immune dysregulation, polyendocrinopathy, X-linked) in humans. Therefore, monogenic autoimmune diseases, such as APS-1 and IPEX, provide the fundamental basis

from which the credibility of theories concerning the cause of T cell–mediated organ specific autoimmune disease can be evaluated (Bacchetta et al., 2007).

In terms of this present research, dissecting the molecular interactions in APS-1 has aided in increasing the understanding of complex polygenic disorders by delineating the breakdown of central tolerance towards organ-specific antigens caused by mutations in the *AIRE* gene, leading to infiltration of lymphocytes in the affected organ and presence of tissue-specific antibodies (Anderson et al. 2005).

In the first stage of the research, we focused on screening the PHD1 domain of the AIRE protein with a yeast-two-hybrid (Y2H) library. The main purpose of the screening was to enhance our understanding of the molecular function of AIRE-PHD1 due to its crucial role in AIRE function. To be specific, AIRE-PHD1 interacts with unmodified histone H3K4 (H3K4me0) and DNA-dependent protein kinase (DNA-PK) via its PHD1 domain, which is a crucial step for activating the expression of tissue-specific antigens. This means it is a central domain for AIRE in terms of its ability to prevent multi-organ autoimmunity (Bottomley et al., 2005); (Koh et al., 2008); 2010; (Žumer et al., 2012). Thus, the first aim of the research was to identify novel proteins interacting with AIRE-PHD1 using a human-normalised cDNA library.

The Y2H library allowed us to identify six protein partners of AIRE: PCNA, ATP1β1, NPC2, PGM1, RNF2, and STX8. PCNA had been previously identified by (Abramson et al., 2010) as interacting with the full-length AIRE, therefore validated our screening. However, we confirmed that the interaction is confined to the PHD1 domain of AIRE. Four of the identified proteins – ATP1β1, NPC2,

PGM1, and RNF2 – may have been false positive results despite their high levels in yeast, as they were situated in a compartment that does not reflect their natural cellular environment (Bab-Dinitz et al., 2009); (Frolov et al., 2013); (Lyons et al., 2015). Furthermore, false positive results may occur when the prey interacts with the membrane anchors that are fused to the bait. Such results may also arise when tested proteins have domains that are involved in creating protein-protein interactions, but in reality, there is no physiological context for the interaction. However, the protein STX8 is likely to interact with AIRE-PHD1 due to its coordination with CBP and its role in regulating the thymic structure, and therefore requires further investigation (Lei et al., 2011, Passos et al., 2015, Kanwar et al., 2008). In the Y2H system, false negative results may also occur as the fused reporter proteins may result in a steric hindrance that might affect the nature of interaction, making it weak, and hence a false negative result (Brückner et al., 2009).

Protein-protein interactions (PPIs), which we define as proteins that physically interact, are vital in the biological processes inside the cell. Y2H screens have been used to isolate novel protein-protein interactions and thus can aid in generating protein interaction maps (Uetz et al., 2000). However, Y2H screens only display physical interaction under specific experimental conditions, and thus do not necessarily show an actual interaction in a biological process. Hence, expression of the prey proteins was confirmed in western blot and the identified interactions were validated using co-IP. We were not able to detect the identified interactions via this technique.

Therefore, we decided to covalently cross-link the proteins prior to testing their interaction via co-IP. The chemical cross-linking involves formation of covalent

bonds between the two proteins via bi-functional reagents containing the reactive end group homobifunctional N-hydroxysuccinimide ester (NHS ester), which reacts with the primary amines (NH2) of amino acid residues. Using this technique, we still did not manage to observe an interaction with the cross-linked proteins. This is possibly due to the nature of interaction, which might be weak and transient, and therefore is difficult to detect. Further optimisation might also help in detecting these interactions, such the use of the full-length of AIRE rather than the PHD1 domain on its own for the interaction, which might help to strengthen the interaction with the existence of the other domains of AIRE. Another possibility may be the technique itself, which might be unsuitable for this type interaction: using an alternative method may then be the solution, such as the protein-fragment complementation assay.

The second stage of the project involved studying the impact of APS-1 mutations on the interaction with AIRE-CARD and AIRE-SAND. The reason for choosing the CARD domain is that 66.6% of APS-1 mutations occur in this domain and most of the known prey protein interactions occur via AIRE-CARD (Ramsey, 2002). Our first aim in this stage was to focus on confirming the interactions of known proteins that interact with AIRE via its CARD domain; namely, Bromodomain-containing protein 4 (BRD4), Death-associated protein 6 (DAXX), and Protein Inhibitor of Activated STAT 1 (PIAS1). Then, APS-1 mutants, T16M, L29P, L93R were introduced to the CARD to test whether these pathogenic mutants affect the interactions (Abramson and Goldfarb, 2016). The rationale for specifically choosing these mutations is that each one of them lie in a different compartment of the CARD domain, and therefore we were able to compare if the location of the mutant also has an impact on the interaction. The

T16M mutation is in the N-terminal of CARD and was used as a control in the co-IP experiment as it has no effect on the transactivation activity of AIRE, as revealed by (Halonen et al., 2004). The L29P mutation lies at the core of the CARD domain, and it has been suggested that it can alter the conformation of the helices. We therefore decided that it would be worthwhile to check if it affects the interaction of AIRE with the proteins BRD4, DAXX, and PIAS1. The L93R mutation, however, is situated at the surface of the CARD domain 3D model, and it has been suggested that it can disrupt the hydrophobic core of the four-helix bundle of the AIRE-CARD (Ferguson et al., 2008). Pitkanen's research group confirmed the location of these mutations by using a three-dimensional homology model for the CARD domain in order to predict the structural alterations of APS-1 mutants affecting the CARD domain of AIRE (Pitkänen et al., 2000). Thus, these mutations give a good coverage of the CARD domain.

The aim here was to establish whether APS-1 mutations affect known protein interactions. Initially, we confirmed that AIRE-CARD interacts with BRD4, DAXX and PIAS1 via co-IP. We also showed that these interactions are not affected by these recessive mutations. It could be that these mutations are included in other AIRE-protein interactions, or it might be that the effect of these APS-1 mutants on these interactions is so weak that it cannot be revealed in a co-IP experiment. Moreover, the CARD domain is primarily linked to AIRE's dimerisation, so mutants at this domain exhibit very little transcriptional activity, especially if the CARD is tested on its own rather than testing the full-length AIRE (Sparks et al., 2016). Thus, differences in interaction between the wild-type CARD and its mutants might be detected if the mutants were introduced

to the full-length AIRE so that it can be compared to the full-length AIRE rather than the CARD domain, as the existence of the remaining domains might aid in enhancing the interaction (Liiv et al., 2008).

The second aim of this stage of the research was to confirm the interaction of AIRE-SAND with p63. The rationale for focusing on SAND was two-fold: It interacts with ATF7ip – linking Aire with further epigenetic processes and it contributes in regulating nuclear organisation, and its absence results in AIRE mis-localisation which impairs its transactivation activity (Waterfield et al., 2014; Yoshida et al., 2015). Hence, evaluating the function of this domain is vital and aids in further outlining the molecular function of AIRE. Once the interaction was confirmed via co-IP, we introduced two APS-1 mutants, G228W and P252L, to study the effect of these mutants on the interaction. The rationale for choosing the G228W mutant was that it is the only known APS-1 mutation with a unique mode of inheritance – dominant negative – making it more attractive to examine its effect on interaction with p63. Moreover, Tonooka et al. (2009) and Ilmarinen et al.'s (2005) discoveries regarding its interaction with p63 does not fit with how it applies its affect in a dominant negative model. The missense APS-1 recessive mutation P252L was used as a control to check if this mode of inheritance behaves in the same way as with the dominant mutant G228W in terms of interaction with p63 (Ilmarinen et al., 2005). The interaction of the protein p63 with AIRE-SAND was confirmed here. It appears that the interaction only occurs via SAND's top band. Thus, de-phosphorylation and de-glycolysation tests were applied to the SAND protein to test whether the interaction is via the post-translationally modified SAND protein. However, it appears that this is not the case, as there is not any significant difference

between the SAND and the post-translationally modified SAND upon interaction with p63. In future work, one of the solutions to identify the nature of AIRE-SAND interaction with p63 is to apply mass-spectrometry, which would enable us to study the dynamics of the interaction.

Furthermore, we demonstrated that the SAND mutant, G228W, interacts more strongly than the wt-SAND with the protein P63. This is because this type of mutant restricts the homodimerisation activity of the wt-AIRE, and hence affects the activity of the normal allele, hindering its function (Cetani et al., 2001). We already know that the mutant P252L does not interfere with the wt-AIRE's binding efficiency to p63. This mutant was therefore used as a control to test for interaction against p63 (Ilmarinen et al., 2005). Moreover, the mutant G228W follows a more dispersed pattern in the nucleus, whereas P252L exists as concentrated specks in the nucleus, as with the wild-type AIRE, allowing the level of transcriptional activation to be observed (Sparks et al., 2016). This result is exciting, as it contradicts previous findings by (Tonooka et al., 2009). They carried out a co-IP experiment to show the interaction of wt-SAND with p63 and the SAND mutant, G228W, with p63 and discovered that the interaction of the mutant with p63 is weaker than the wt-SAND. Therefore, correlates to the function (i.e. translated into function). However, our findings suggest that the interaction is not translated into function, which does give a reasonable explanation given the dominant negative inheritance of this APS-1 mutation.

In addition, we studied the mechanism of interaction of SAND-p63, which allowed us to justify this unique model of dominant negative inheritance. Thereby, the dual-luciferase activity was measured for the following

interactions: SAND-p63 and G228W-p63. The result suggests that the interaction of the wt-SAND utilises a higher transcription activity than with its mutant, G228W, reflecting the action of this mutant in hampering AIRE function, reducing its transcriptional activation (Sparks et al., 2016). This is in keeping with the findings of Halonen et al. (2001), who visualised and compared the subcellular localisation of the wild-type AIRE and its mutant, G228W. They cloned the cDNAs in mammalian expression vectors, expressing them in African green monkey kidney (COS-1) cells. The cells were then examined using immunofluorescence. The wild-type AIRE was localised in both the cytoplasm and the nucleus. In the cytoplasm, it was associated with filamentous structures, yet in the nucleus they occurred as nuclear dots. Conversely, the G228W mutant had greatly reduced the association of AIRE with the cytoplasmic filaments as well as in the nuclear dots, where they were detected as clump-liked aggregates (Halonen et al., 2004). Thus, this mutation limits the attachment of AIRE to the nuclear dots, leading to intracellular aggregation. This in turn leads to the mislocalisation of the wild-type AIRE, which has been linked to the inhibition of its transactivation capacity (Sparks et al., 2016).

Our results are in line with the theory of dominant negative inheritance effect by which the G228W protein via its association with the wt-AIRE, which hinders its ability to develop the complexes required for transactivation. Moreover, we managed to prove that AIRE-SAND binds to p63 via its TA domain (Tonooka et al., 2009). This was reinforced by measuring the functional impact of the interaction with the truncated version of p63, which is translated as a dual-luciferase activity. The interaction was abolished, indicating that p63 lacking its TA domain could not bind to AIRE-SAND. Furthermore, the functional activity

of the interaction of p63 with AIRE as a whole protein was tested to check whether it has an impact on the interaction, and this was compared with AIRE-G228W. We managed to establish a higher interaction when the full-length AIRE interacted with p63 than with the SAND domain on its own, signifying that other domains of AIRE have a role in strengthening the interaction with the protein p63. The interaction was also higher when the mutant G228W was introduced to the full-length AIRE. To conclude, the mutant G228W is a robust competitor to the wt-AIRE; however, the wild type can still bind to the protein complex, including p63, and be moderately active. This explains the mild phenotype that we can see in APS-1 patients bearing the G228W mutant (Capalbo et al., 2012). Therefore, we demonstrated the model that shows the molecular mechanism of the dominant negative mode of inheritance in APS-1

Future work will focus on further characterising the interaction of SAND-p63, comparing it with G228W-p63 in order to confirm the link between the phenotype and the molecular mechanism of the dominant negative mode of transmission in APS-1. Therefore, the equilibrium dissociation constant will be measured (Kd) between both interactions, and this will confirm the binding affinity of both proteins where a high Kd value corresponds to a greater binding affinity. Before the measuring the Kd, the proteins will be purified to ensure that no other proteins exist in solution. After confirming the Kd values, we will study the structural analysis for both SAND-p63 and G228W-p63 using X-ray crystallography. This technique provides atomic-resolution structures and hence the molecular details of how the interactions between the different components occur, and thus, will elucidate the molecular structure of the interaction and the atoms that are involved. This in turn will expand our

knowledge of the function of AIRE-SAND. In terms of the G228W-p63 interaction, it will give us an indication of the effect of this mutant on the conformation and the folding of the crystal structure. This will allow us to study how this change in structure affects the interaction with p63. In conclusion, dissecting the molecular mechanism underlying the pathogenesis of APS-1 will develop our understanding of autoimmunity. This will also contribute to our understanding of the mechanism of polygenic diseases such as type-1 diabetes and alopecia areata.

Chapter 6
References

6. References

ABRAMSON, J., GIRAUD, M., BENOIST, C. & MATHIS, D. 2010. Aire's Partners in the Molecular Control of Immunological Tolerance. *Cell,* 140**,** 123-135.

ABRAMSON, J. & GOLDFARB, Y. 2016. AIRE: From promiscuous molecular partnerships to promiscuous gene expression. *European Journal of Immunology,* 46**,** 22-33.

AHLGREN, K. M., MORETTI, S., LUNDGREN, B. A., KARLSSON, I., ÅHLIN, E., NORLING, A., HALLGREN, Å., PERHEENTUPA, J., GUSTAFSSON, J., RORSMAN, F., CREWTHER, P. E., RÖNNELID, J., BENSING, S., SCOTT, H. S., KÄMPE, O., ROMANI, L. & LOBELL, A. 2011. Increased IL-17A secretion in response to Candida albicans in autoimmune polyendocrine syndrome type 1 and its animal model. *European Journal of Immunology,* 41**,** 235-245.

AHONEN, P., MYLLÄRNIEMI, S., SIPILÄ, I. & PERHEENTUPA, J. 1990. Clinical Variation of Autoimmune Polyendocrinopathy–Candidiasis–Ectodermal Dystrophy (APECED) in a Series of 68 Patients. *New England Journal of Medicine,* 322**,** 1829-1836.

AKIRA, S., UEMATSU, S. & TAKEUCHI, O. 2006. Pathogen Recognition and Innate Immunity. *Cell,* 124**,** 783-801.

AKIYAMA, T., SHIMO, Y., YANAI, H., QIN, J., OHSHIMA, D., MARUYAMA, Y., ASAUMI, Y., KITAZAWA, J., TAKAYANAGI, H., PENNINGER, J. M., MATSUMOTO, M., NITTA, T., TAKAHAMA, Y. & INOUE, J.-I. The Tumor Necrosis Factor Family Receptors RANK and CD40 Cooperatively Establish the Thymic Medullary Microenvironment and Self-Tolerance. *Immunity,* 29**,** 423-437.

ALIMOHAMMADI, M., BJÖRKLUND, P., HALLGREN, Å., PÖNTYNEN, N., SZINNAI, G., SHIKAMA, N., KELLER, M. P., EKWALL, O., KINKEL, S. A., HUSEBYE, E. S., GUSTAFSSON, J., RORSMAN, F., PELTONEN, L., BETTERLE, C., PERHEENTUPA, J., ÅKERSTRÖM, G., WESTIN, G., SCOTT, H. S., HOLLÄNDER, G. A. & KÄMPE, O. 2008. Autoimmune Polyendocrine Syndrome Type 1 and NALP5, a Parathyroid Autoantigen. *New England Journal of Medicine,* 358**,** 1018-1028.

ANDERSON, M. S., VENANZI, E. S., CHEN, Z., BERZINS, S. P., BENOIST, C. & MATHIS, D. The Cellular Mechanism of Aire Control of T Cell Tolerance. *Immunity,* 23**,** 227-239.

ANDERSON, M. S., VENANZI, E. S., KLEIN, L., CHEN, Z., BERZINS, S. P., TURLEY, S. J., VON BOEHMER, H., BRONSON, R., DIERICH, A., XE, BENOIST, C. & MATHIS, D. 2002. Projection of an Immunological Self Shadow within the Thymus by the Aire Protein. *Science,* 298**,** 1395-1401.

ARICHA, R., FEFERMAN, T., SCOTT, H. S., SOUROUJON, M. C., BERRIH-AKNIN, S. & FUCHS, S. 2011. The susceptibility of Aire−/− mice to experimental myasthenia gravis involves alterations in regulatory T cells. *Journal of Autoimmunity,* 36**,** 16-24.

ATASSI, M. Z. & CASALI, P. 2008. Molecular mechanisms of autoimmunity. *Autoimmunity,* 41**,** 123-132.

AZZAM, H. S., GRINBERG, A., LUI, K., SHEN, H., SHORES, E. W. & LOVE, P. E. 1998. CD5 Expression Is Developmentally Regulated By T Cell Receptor (TCR) Signals and TCR Avidity. *Journal of Experimental Medicine,* 188**,** 2301-2311.

BAB-DINITZ, E., ALBECK, S., PELEG, Y., BRUMFELD, V., GOTTSCHALK, K. E. & KARLISH, S. J. D. 2009. A C-Terminal Lobe of the β Subunit of Na, K-ATPase and H, K-ATPase Resembles Cell Adhesion Molecules. *Biochemistry,* 48**,** 8684-8691.

BACCHETTA, R., GAMBINERI, E. & RONCAROLO, M.-G. 2007. Role of regulatory T cells and FOXP3 in human diseases. *Journal of Allergy and Clinical Immunology,* 120**,** 227-235.

BENNETT, C. L., CHRISTIE, J., RAMSDELL, F., BRUNKOW, M. E., FERGUSON, P. J., WHITESELL, L., KELLY, T. E., SAULSBURY, F. T., CHANCE, P. F. & OCHS, H. D. 2001. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nature Genetics,* 27**,** 20-21.

BENNETT, I., HARVEY, A. M., CARPENTER, C. C. J., SOLOMON, N., SILBERBERG, S. G., BLEDSOE, T., NORTHCUTT, R. C. & KLINENBERG, J. R. 1964. Schmidt's Syndrome (Thyroid and Adrenal

Insufficiency): A Review of the Literature and a Report of Fifteen New Cases Including Ten Instances of Coexistent Diabetes Mellitus. *Transactions of the American Clinical and Climatological Association,* 75**,** 27-36.

BERRIH-AKNIN, S. 2016. Role of the thymus in autoimmune myasthenia gravis. *Clinical and Experimental Neuroimmunology,* 7**,** 226-237.

BETTERLE, C., DAL PRA, C., MANTERO, F. & ZANCHETTA, R. 2002. Autoimmune Adrenal Insufficiency and Autoimmune Polyendocrine Syndromes: Autoantibodies, Autoantigens, and Their Applicability in Diagnosis and Disease Prediction. *Endocrine Reviews,* 23**,** 327-364.

BHAT, S. S., FRIEDMANN, K. S., KNÖRCK, A., HOXHA, C., LEIDINGER, P., BACKES, C., MEESE, E., KELLER, A., RETTIG, J., HOTH, M., QU, B. & SCHWARZ, E. C. 2016. Syntaxin 8 is required for efficient lytic granule trafficking in cytotoxic T lymphocytes. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1863**,** 1653-1664.

BJÖRSES, P., HALONEN, M., PALVIMO, J. J., KOLMER, M., AALTONEN, J., ELLONEN, P., PERHEENTUPA, J., ULMANEN, I. & PELTONEN, L. 2000. Mutations in the AIRE Gene: Effects on Subcellular Location and Transactivation Function of the Autoimmune Polyendocrinopathy-Candidiasis–Ectodermal Dystrophy Protein. *The American Journal of Human Genetics,* 66**,** 378-392.

BLECHSCHMIDT, K., SCHWEIGER, M., WERTZ, K., POULSON, R., CHRISTENSEN, H. M., ROSENTHAL, A., LEHRACH, H. & YASPO, M. L. 1999. The mouse Aire gene: comparative genomic sequencing, gene organization, and expression. *Genome research,* 9**,** 158-166.

BONILLA, F. A. & OETTGEN, H. C. 2010. Adaptive immunity. *Journal of Allergy and Clinical Immunology,* 125**,** S33-S40.

BOTTOMLEY, M. J., STIER, G., PENNACCHINI, D., LEGUBE, G., SIMON, B., AKHTAR, A., SATTLER, M. & MUSCO, G. 2005. NMR Structure of the First PHD Finger of Autoimmune Regulator Protein (AIRE1): INSIGHTS INTO AUTOIMMUNE POLYENDOCRINOPATHY-CANDIDIASIS-ECTODERMAL DYSTROPHY (APECED) DISEASE. *Journal of Biological Chemistry,* 280**,** 11505-11512.

BRATANIC, N., KISAND, K., AVBELJ STEFANIJA, M., BATTELINO, T. & TREBUSAK PODKRAJSEK, K. 2015. Clinical, Genetic and Immunological Characteristics of Paediatric Autoimmune Polyglandular Syndrome Type 1 Patients in Slovenia. *Zdravstveno varstvo,* 54**,** 112-118.

BRATLAND, E., MAGITTA, N. W. F., BØE WOLFF, A. S., EKERN, T., KNAPPSKOG, P. M., KÄMPE, O., HAAVIK, J. & HUSEBYE, E. S. 2013. Autoantibodies against aromatic amino acid hydroxylases in patients with autoimmune polyendocrine syndrome type 1 target multiple antigenic determinants and reveal regulatory regions crucial for enzymatic activity. *Immunobiology,* 218**,** 899-909.

BRUN, J. M. 1982. Juvenile Autoimmune Polyendocrinopathy. *Hormone Research in Paediatrics,* 16**,** 308-316.

BRÄNNSTRÖM, J., HÄSSLER, S., PELTONEN, L., HERRMANN, B. & WINQVIST, O. 2006. Defect internalization and tyrosine kinase activation in Aire deficient antigen presenting cells exposed to Candida albicans antigens. *Clinical Immunology,* 121**,** 265-273.

BRÜCKNER, A., POLGE, C., LENTZE, N., AUERBACH, D. & SCHLATTNER, U. 2009. Yeast Two-Hybrid, a Powerful Tool for Systems Biology. *International Journal of Molecular Sciences,* 10**,** 2763-2788.

BUNNELL, S. C., HONG, D. I., KARDON, J. R., YAMAZAKI, T., MCGLADE, C. J., BARR, V. A. & SAMELSON, L. E. 2002. T cell receptor ligation induces the formation of dynamically regulated signaling assemblies. *The Journal of cell biology,* 158**,** 1263-1275.

BÖCKLE, B. C., WILHELM, M., MÜLLER, H., GÖTSCH, C. & SEPP, N. T. 2010. Oral mucous squamous cell carcinoma—an anticipated consequence of autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED). *Journal of the American Academy of Dermatology,* 62**,** 864-868.

CAPALBO, D., DE MARTINO, L., GIARDINO, G., DI MASE, R., DI DONATO, I., PARENTI, G., VAJRO, P., PIGNATA, C. & SALERNO, M. 2012. Autoimmune polyendocrinopathy candidiasis ectodermal

dystrophy: insights into genotype-phenotype correlation. *International journal of endocrinology,* 2012**,** 353250-353250.
CASPI, R. R. 2008. Immunotherapy of autoimmunity and cancer: the penalty for success. *Nature Reviews Immunology,* 8**,** 970-976.
CETANI, F., BARBESINO, G., BORSARI, S., PARDI, E., CIANFEROTTI, L., PINCHERA, A. & MARCOCCI, C. 2001. A Novel Mutation of the Autoimmune Regulator Gene in an Italian Kindred with Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy, Acting in a Dominant Fashion and Strongly Cosegregating with Hypothyroid Autoimmune Thyroiditis. *The Journal of Clinical Endocrinology & Metabolism,* 86**,** 4747-4752.
CHEN, Z., BENOIST, C. & MATHIS, D. 2005. How defects in central tolerance impinge on a deficiency in regulatory T cells. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 14735.
CHIESA, M. D., SIVORI, S., CASTRICONI, R., MARCENARO, E. & MORETTA, A. 2005. Pathogen-induced private conversations between natural killer and dendritic cells. *Trends in Microbiology,* 13**,** 128-136.
COBBOLD, S. P., GRACA, L., LIN, C.-Y., ADAMS, E. & WALDMANN, H. 2003. Regulatory T cells in the induction and maintenance of peripheral transplantation tolerance. *Transplant International,* 16**,** 66-75.
COLLINS, S. M., DOMINGUEZ, M., ILMARINEN, T., COSTIGAN, C. & IRVINE, A. D. 2006. Dermatological manifestations of autoimmune polyendocrinopathy–candidiasis–ectodermal dystrophy syndrome. *British Journal of Dermatology,* 154**,** 1088-1093.
COSTENBADER, K. H., GAY, S., ALARCÓN-RIQUELME, M. E., IACCARINO, L. & DORIA, A. 2012. Genes, epigenetic regulation and environmental factors: Which is the most relevant in developing autoimmune diseases? *Autoimmunity Reviews,* 11**,** 604-609.
CUNNINGHAM, M. W. 2019. Molecular Mimicry, Autoimmunity, and Infection: The Cross-Reactive Antigens of Group A Streptococci and their Sequelae. *Microbiology spectrum,* 7**,** 10.1128/microbiolspec.GPP3-0045-2018.
DAVIDSON, A. & DIAMOND, B. 2001. Autoimmune Diseases. *New England Journal of Medicine,* 345**,** 340-350.
DE LUCA, F., VALENZISE, M., ALAGGIO, R., ARRIGO, T., CRISAFULLI, G., SALZANO, G., CERVATO, S., MARINIELLO, B., LAZZAROTTO, F. & BETTERLE, C. 2008. Sicilian family with autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED) and lethal lung disease in one of the affected brothers. *European Journal of Pediatrics,* 167**,** 1283-1288.
DE MARTINO, L., CAPALBO, D., IMPRODA, N., LORELLO, P., UNGARO, C., DI MASE, R., CIRILLO, E., PIGNATA, C. & SALERNO, M. 2016. Novel Findings into AIRE Genetics and Functioning: Clinical Implications. *Frontiers in pediatrics,* 4**,** 86-86.
DELVES, P. J. & ROITT, I. M. 2000. The immune system - First of two parts. *New England Journal of Medicine,* 343**,** 37-49.
DERBINSKI, J. & KYEWSKI, B. 2010. How thymic antigen presenting cells sample the body's self-antigens. *Current Opinion in Immunology,* 22**,** 592-600.
DEVOSS, J., HOU, Y., JOHANNES, K., LU, W., LIOU, G. I., RINN, J., CHANG, H., CASPI, R. R., FONG, L. & ANDERSON, M. S. 2006. Spontaneous autoimmunity prevented by thymic expression of a single self-antigen. *Journal of Experimental Medicine,* 203**,** 2727-2735.
DZHAGALOV, I. L., CHEN, K. G., HERZMARK, P. & ROBEY, E. A. 2013. Elimination of self-reactive T cells in the thymus: a timeline for negative selection. *PLoS biology,* 11**,** e1001566-e1001566.
EISENBARTH, G. S. & GOTTLIEB, P. A. 2004. Autoimmune Polyendocrine Syndromes. *New England Journal of Medicine,* 350**,** 2068-2079.
ERNST, B., LEE, D.-S., CHANG, J. M., SPRENT, J. & SURH, C. D. 1999. The Peptide Ligands Mediating Positive Selection in the Thymus Control T Cell Survival and Homeostatic Proliferation in the Periphery. *Immunity,* 11**,** 173-181.

FERGUSON, B. J., ALEXANDER, C., ROSSI, S. W., LIIV, I., REBANE, A., WORTH, C. L., WONG, J., LAAN, M., PETERSON, P., JENKINSON, E. J., ANDERSON, G., SCOTT, H. S., COOKE, A. & RICH, T. 2008. AIRE's CARD Revealed, a New Structure for Central Tolerance Provokes Transcriptional Plasticity. *Journal of Biological Chemistry,* 283**,** 1723-1731.

FRIDKIS-HARELI, M. 2008. Immunogenetic mechanisms for the coexistence of organ-specific and systemic autoimmune diseases. *Journal of Autoimmune Diseases,* 5**,** 1-1.

FROLOV, A., DONG, H., JIANG, M., YANG, L., COOK, E. C., MATNANI, R., HAMMOCK, B. D. & CROFFORD, L. J. 2013. Niemann-Pick Type C2 Deficiency in Human Fibroblasts Confers Robust and Selective Activation of Prostaglandin E2 Biosynthesis. *Journal of Biological Chemistry,* 288**,** 23696-23703.

FU, G., RYBAKIN, V., BRZOSTEK, J., PASTER, W., ACUTO, O. & GASCOIGNE, N. R. J. 2014. Fine-tuning T cell receptor signaling to control T cell development. *Trends in immunology,* 35**,** 311-318.

FUJIO, K., OKAMURA, T., SUMITOMO, S. & YAMAMOTO, K. 2012. Regulatory T cell-mediated control of autoantibody-induced inflammation. *Frontiers in immunology,* 3**,** 28-28.

GALLEGOS, A. M. & BEVAN, M. J. 2004. Central Tolerance to Tissue-specific Antigens Mediated by Direct and Indirect Antigen Presentation. *The Journal of Experimental Medicine,* 200**,** 1039-1049.

GERMAIN, R. N. 2001. The Art of the Probable: System Control in the Adaptive Immune System. *Science,* 293**,** 240.

GILLARD, G. O., DOOLEY, J., ERICKSON, M., PELTONEN, L. & FARR, A. G. 2007. Aire-Dependent Alterations in Medullary Thymic Epithelium Indicate a Role for Aire in Thymic Epithelial Differentiation. *The Journal of Immunology,* 178**,** 3007-3015.

GIRAUD, M., TAUBERT, R., VANDIEDONCK, C., KE, X., LEVI-STRAUSS, M., PAGANI, F., BARALLE, F. E., EYMARD, B., TRANCHANT, C., GAJDOS, P., VINCENT, A., WILLCOX, N., BEESON, D., KYEWSKI, B. & GARCHON, H.-J. 2007. An IRF8-binding promoter variant and AIRE control CHRNA1 promiscuous expression in thymus. *Nature,* 448**,** 934-937.

GOODNOW, C. C., SPRENT, J., DE ST GROTH, B. F. & VINUESA, C. G. 2005. Cellular and genetic mechanisms of self tolerance and autoimmunity. *Nature,* 435**,** 590-597.

GOTTER, J., BRORS, B., HERGENHAHN, M. & KYEWSKI, B. 2004. Medullary Epithelial Cells of the Human Thymus Express a Highly Diverse Selection of Tissue-specific Genes Colocalized in Chromosomal Clusters. *The Journal of Experimental Medicine,* 199**,** 155-166.

GOUGH, S. C. L. & SIMMONDS, M. J. 2007. The HLA Region and Autoimmune Disease: Associations and Mechanisms of Action. *Current genomics,* 8**,** 453-465.

GUO, C.-J., LEUNG, P. S. C., ZHANG, W., MA, X. & GERSHWIN, M. E. 2018. The immunobiology and clinical features of type 1 autoimmune polyglandular syndrome (APS-1). *Autoimmunity Reviews,* 17**,** 78-85.

GUO, J., RAHMAN, M., CHENG, L., ZHANG, S., TVINNEREIM, A. & SU, D.-M. 2011. Morphogenesis and maintenance of the 3D thymic medulla and prevention of nude skin phenotype require FoxN1 in pre- and post-natal K14 epithelium. *Journal of Molecular Medicine,* 89**,** 263-277.

GYLLING, M., TUOMI, T., BJÖRSES, P., KONTIAINEN, S., PARTANEN, J., CHRISTIE, M. R., KNIP, M., PERHEENTUPA, J. & MIETTINEN, A. 2000. β-Cell Autoantibodies, Human Leukocyte Antigen II Alleles, and Type 1 Diabetes in Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy*. *The Journal of Clinical Endocrinology & Metabolism,* 85**,** 4434-4440.

HALONEN, M., ESKELIN, P., MYHRE, A.-G., PERHEENTUPA, J., HUSEBYE, E. S., KÄMPE, O., RORSMAN, F., PELTONEN, L., ULMANEN, I. & PARTANEN, J. 2002. AIRE Mutations and Human Leukocyte Antigen Genotypes as Determinants of the Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy Phenotype. *The Journal of Clinical Endocrinology & Metabolism,* 87**,** 2568-2574.

HALONEN, M., KANGAS, H., RÜPPELL, T., ILMARINEN, T., OLLILA, J., KOLMER, M., VIHINEN, M., PALVIMO, J., SAARELA, J., ULMANEN, I. & ESKELIN, P. 2004. APECED-causing mutations in AIRE reveal the functional domains of the protein. *Human Mutation,* 23**,** 245-257.

HALONEN, M., PELTO–HUIKKO, M., ESKELIN, P., PELTONEN, L., ULMANEN, I. & KOLMER, M. 2001. Subcellular Location and Expression Pattern of Autoimmune Regulator (Aire), the Mouse Orthologue for Human Gene Defective in Autoimmune Polyendocrinopathy Candidiasis Ectodermal Dystrophy (APECED). *Journal of Histochemistry & Cytochemistry,* 49**,** 197-208.

HAMAZAKI, Y., FUJITA, H., KOBAYASHI, T., CHOI, Y., SCOTT, H. S., MATSUMOTO, M. & MINATO, N. 2007. Medullary thymic epithelial cells expressing Aire represent a unique lineage derived from cells expressing claudin. *Nat Immunol,* 8**,** 304-311.

HAYTER, S. M. & COOK, M. C. 2012. Updated assessment of the prevalence, spectrum and case definition of autoimmune disease. *Autoimmunity Reviews,* 11**,** 754-765.

HOFMANN, K., BUCHER, P. & TSCHOPP, J. The CARD domain: a new apoptotic signalling motif. *Trends in Biochemical Sciences,* 22**,** 155-156.

HOLOPAINEN, P., ARVAS, M., SISTONEN, P., MUSTALAHTI, K., COLLIN, P., MÄKI, M. & PARTANEN, J. 1999. CD28/CTLA4 gene region on chromosome 2q33 confers genetic susceptibility to celiac disease. A linkage and family-based association study. *Tissue Antigens,* 53**,** 470-475.

HONG, M., RYAN, K. R., ARKWRIGHT, P. D., GENNERY, A. R., COSTIGAN, C., DOMINGUEZ, M., DENNING, D. W., MCCONNELL, V., CANT, A. J., ABINUN, M., SPICKETT, G. P., SWAN, D. C., GILLESPIE, C. S., YOUNG, D. A. & LILIC, D. 2009. Pattern recognition receptor expression is not impaired in patients with chronic mucocutanous candidiasis with or without autoimmune polyendocrinopathy candidiasis ectodermal dystrophy. *Clinical & Experimental Immunology,* 156**,** 40-51.

HUBERT, F.-X., KINKEL, S. A., CREWTHER, P. E., CANNON, P. Z. F., WEBSTER, K. E., LINK, M., UIBO, R., BRYAN, M. K., MEAGER, A., FOREHAN, S. P., SMYTH, G. K., MITTAZ, L., ANTONARAKIS, S. E., PETERSON, P., HEATH, W. R. & SCOTT, H. S. 2009. Aire-Deficient C57BL/6 Mice Mimicking the Common Human 13-Base Pair Deletion Mutation Present with Only a Mild Autoimmune Phenotype. *The Journal of Immunology,* 182**,** 3902.

HUBERT, F.-X., KINKEL, S. A., DAVEY, G. M., PHIPSON, B., MUELLER, S. N., LISTON, A., PROIETTO, A. I., CANNON, P. Z. F., FOREHAN, S., SMYTH, G. K., WU, L., GOODNOW, C. C., CARBONE, F. R., SCOTT, H. S. & HEATH, W. R. 2011. Aire regulates the transfer of antigen from mTECs to dendritic cells for induction of thymic tolerance. *Blood,* 118**,** 2462-2472.

HUTTLIN, EDWARD L., TING, L., BRUCKNER, RAPHAEL J., GEBREAB, F., GYGI, MELANIE P., SZPYT, J., TAM, S., ZARRAGA, G., COLBY, G., BALTIER, K., DONG, R., GUARANI, V., VAITES, LAURA P., ORDUREAU, A., RAD, R., ERICKSON, BRIAN K., WÜHR, M., CHICK, J., ZHAI, B., KOLIPPAKKAM, D., MINTSERIS, J., OBAR, ROBERT A., HARRIS, T., ARTAVANIS-TSAKONAS, S., SOWA, MATHEW E., DE CAMILLI, P., PAULO, JOAO A., HARPER, J. W. & GYGI, STEVEN P. The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell,* 162**,** 425-440.

ILMARINEN, T., ESKELIN, P., HALONEN, M., RÜPPELL, T., KILPIKARI, R., TORRES, G. D., KANGAS, H. & ULMANEN, I. 2005. Functional analysis of SAND mutations in AIRE supports dominant inheritance of the G228W mutation. *Human Mutation,* 26**,** 322-331.

ILMARINEN, T., KANGAS, H., KYTÖMAA, T., ESKELIN, P., SAHARINEN, J., SEELER, J.-S., TANHUANPÄÄ, K., CHAN, F. Y.-L., SLATTERY, R. M., ALAKURTTI, K., PALVIMO, J. J. & ULMANEN, I. 2008. Functional interaction of AIRE with PIAS1 in transcriptional regulation. *Molecular Immunology,* 45**,** 1847-1862.

ISHII, T., SUZUKI, Y., ANDO, N., MATSUO, N. & OGATA, T. 2000. Novel Mutations of the Autoimmune Regulator Gene in Two Siblings with Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy1. *The Journal of Clinical Endocrinology & Metabolism,* 85**,** 2922-2926.

IWASAKI, A. & MEDZHITOV, R. 2010. Regulation of Adaptive Immunity by the Innate Immune System. *Science,* 327**,** 291.

JIANG, W., ANDERSON, M. S., BRONSON, R., MATHIS, D. & BENOIST, C. 2005. Modifier loci condition autoimmunity provoked by Aire deficiency. *The Journal of Experimental Medicine,* 202**,** 805-815.

KANWAR, N., FAYYAZI, A., BACKOFEN, B., NITSCHE, M., DRESSEL, R. & FISCHER VON MOLLARD, G. 2008. Thymic alterations in mice deficient for the SNARE protein VAMP8/endobrevin. *Cell and Tissue Research,* 334**,** 227-242.

KEDZIERSKA, K., VALKENBURG, S. A., GUILLONNEAU, C., HUBERT, F.-X., CUKALAC, T., CURTIS, J. M., STAMBAS, J., SCOTT, H. S., KEDZIERSKI, L., VENTURI, V. & DAVENPORT, M. P. 2010. Diversity and clonotypic composition of influenza-specific CD8+ TCR repertoires remain unaltered in the absence of Aire. *European Journal of Immunology,* 40**,** 849-858.

KEKÄLÄINEN, E., TUOVINEN, H., JOENSUU, J., GYLLING, M., FRANSSILA, R., PÖNTYNEN, N., TALVENSAARI, K., PERHEENTUPA, J., MIETTINEN, A. & ARSTILA, T. P. 2007. A Defect of Regulatory T Cells in Patients with Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy. *The Journal of Immunology,* 178**,** 1208.

KIM, K. D., ZHAO, J., AUH, S., YANG, X., DU, P., TANG, H. & FU, Y.-X. 2007. Adaptive immune cells temper initial innate responses. *Nature medicine,* 13**,** 1248-1252.

KINDT, T. J., GOLDSBY, R. A., OSBORNE, B. A. & KUBY, J. 2007. *Kuby immunology*, Macmillan.

KISAND, K., LILIC, D., CASANOVA, J.-L., PETERSON, P., MEAGER, A. & WILLCOX, N. 2011. Mucocutaneous candidiasis and autoimmunity against cytokines in APECED and thymoma patients: Clinical and pathogenetic implications. *European Journal of Immunology,* 41**,** 1517-1527.

KISAND, K., LINK, M., WOLFF, A. S. B., MEAGER, A., TSEREL, L., ORG, T., MURUMÄGI, A., UIBO, R., WILLCOX, N., TREBUŠAK PODKRAJŠEK, K., BATTELINO, T., LOBELL, A., KÄMPE, O., LIMA, K., MELONI, A., ERGUN-LONGMIRE, B., MACLAREN, N. K., PERHEENTUPA, J., KROHN, K. J. E., SCOTT, H. S., HUSEBYE, E. S. & PETERSON, P. 2008. Interferon autoantibodies associated with AIRE deficiency decrease the expression of IFN-stimulated genes. *Blood,* 112**,** 2657-2666.

KLEIN, J. & SATO, A. 2000. The HLA System. *New England Journal of Medicine,* 343**,** 702-709.

KLEIN, L., KYEWSKI, B., ALLEN, P. M. & HOGQUIST, K. A. 2014. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nature reviews. Immunology,* 14**,** 377-391.

KLEMETTI, P., BJÖRSES, P., TUOMI, T., PERHEENTUPA, J., PARTANEN, J., RAUTONEN, N., HINKKANEN, A., ILONEN, J. & VAARALA, O. 2000. Autoimmunity to glutamic acid decarboxylase in patients with autoimmune polyendocrinopathy–candidiasis–ectodermal dystrophy (APECED). *Clinical and Experimental Immunology,* 119**,** 419-425.

KOH, A. S., KUO, A. J., PARK, S. Y., CHEUNG, P., ABRAMSON, J., BUA, D., CARNEY, D., SHOELSON, S. E., GOZANI, O., KINGSTON, R. E., BENOIST, C. & MATHIS, D. 2008. Aire employs a histone-binding module to mediate immunological tolerance, linking chromatin regulation with organ-specific autoimmunity. *Proceedings of the National Academy of Sciences,* 105**,** 15878-15883.

KONDO, M., WEISSMAN, I. L. & AKASHI, K. 1997. Identification of Clonogenic Common Lymphoid Progenitors in Mouse Bone Marrow. *Cell,* 91**,** 661-672.

KONT, V., LAAN, M., KISAND, K., MERITS, A., SCOTT, H. S. & PETERSON, P. 2008. Modulation of Aire regulates the expression of tissue-restricted antigens. *Molecular Immunology,* 45**,** 25-33.

KUKREJA, A. & MACLAREN*, N. K. 2000. Current cases in which epitope mimicry is considered as a component cause of autoimmune disease: immune-mediated (type 1) diabetes. *Cellular and Molecular Life Sciences CMLS,* 57**,** 534-541.

KURODA, N., MITANI, T., TAKEDA, N., ISHIMARU, N., ARAKAKI, R., HAYASHI, Y., BANDO, Y., IZUMI, K., TAKAHASHI, T., NOMURA, T., SAKAGUCHI, S., UENO, T., TAKAHAMA, Y., UCHIDA, D., SUN, S., KAJIURA, F., MOURI, Y., HAN, H., MATSUSHIMA, A., YAMADA, G. & MATSUMOTO, M. 2005. Development of Autoimmunity against Transcriptionally Unrepressed Target Antigen in the Thymus of Aire-Deficient Mice. *The Journal of Immunology,* 174**,** 1862-1870.

LEBOVITZ, H. E. 2013. Autoimmune Polyglandular Syndromes: Interplay Between the Immune and the Endocrine Systems Leading to a Diverse Set of Clinical Diseases and New Insights into Immune Regulation. *Diabetes Technology & Therapeutics,* 15**,** S2-21-S2-28.

LEI, Y., RIPEN, A. M., ISHIMARU, N., OHIGASHI, I., NAGASAWA, T., JEKER, L. T., BÖSL, M. R., HOLLÄNDER, G. A., HAYASHI, Y., DE WAAL MALEFYT, R., NITTA, T. & TAKAHAMA, Y. 2011. Aire-dependent production of XCL1 mediates medullary accumulation of thymic dendritic cells and contributes to regulatory T cell development. *The Journal of Experimental Medicine,* 208**,** 383-394.

LERNMARK, A. 2001. Autoimmune diseases: are markers ready for prediction? *The Journal of clinical investigation,* 108**,** 1091-1096.

LERNMARK, Å. & LARSSON, H. E. 2013. Immune therapy in type 1 diabetes mellitus. *Nature Reviews Endocrinology,* 9**,** 92-103.

LI, D., STREETEN, E. A., CHAN, A., LWIN, W., TIAN, L., PELLEGRINO DA SILVA, R., KIM, C. E., ANDERSON, M. S., HAKONARSON, H. & LEVINE, M. A. 2017. Exome Sequencing Reveals Mutations in AIRE as a Cause of Isolated Hypoparathyroidism. *The Journal of Clinical Endocrinology & Metabolism,* 102**,** 1726-1733.

LI, J., LI, Y., YAO, J.-Y., JIN, R., ZHU, M.-Z., QIAN, X.-P., ZHANG, J., FU, Y.-X., WU, L., ZHANG, Y. & CHEN, W.-F. 2007. Developmental pathway of CD4(+)CD8(−) medullary thymocytes during mouse ontogeny and its defect in Aire(−/−) mice. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 18175-18180.

LIANG, S. C., TAN, X.-Y., LUXENBERG, D. P., KARIM, R., DUNUSSI-JOANNOPOULOS, K., COLLINS, M. & FOUSER, L. A. 2006. Interleukin (IL)-22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides. *Journal of Experimental Medicine,* 203**,** 2271-2279.

LIIV, I., REBANE, A., ORG, T., SAARE, M., MASLOVSKAJA, J., KISAND, K., JURONEN, E., VALMU, L., BOTTOMLEY, M. J., KALKKINEN, N. & PETERSON, P. 2008. DNA-PK contributes to the phosphorylation of AIRE: Importance in transcriptional activity. *Biochimica et Biophysica Acta,* 1783**,** 74-83.

LINDH, E., ROSMARAKI, E., BERG, L., BRAUNER, H., KARLSSON, M. C. I., PELTONEN, L., HÖGLUND, P. & WINQVIST, O. 2010. AIRE deficiency leads to impaired iNKT cell development. *Journal of Autoimmunity,* 34**,** 66-72.

LISTON, A. 2004. Gene dosage-limiting role of Aire in thymic expression, clonal deletion, and organ-specific autoimmunity. *J. Exp. Med.,* 200**,** 1015-1026.

LISTON, A., LESAGE, S., WILSON, J., PELTONEN, L. & GOODNOW, C. C. 2003. Aire regulates negative selection of organ-specific T cells. *Nature Immunology,* 4**,** 350-354.

LOVEWELL, T. R. J., MCDONAGH, A. J. G., MESSENGER, A. G., AZZOUZ, M. & TAZI-AHNINI, R. 2018. Meta-Analysis of Autoimmune Regulator-Regulated Genes in Human and Murine Models: A Novel Human Model Provides Insights on the Role of Autoimmune Regulator in Regulating STAT1 and STAT1-Regulated Genes. *Frontiers in Immunology,* 9**,** 1380.

LYONS, J. J., MILNER, J. D. & ROSENZWEIG, S. D. 2015. Glycans Instructing Immunity: The Emerging Role of Altered Glycosylation in Clinical Immunology. *Frontiers in pediatrics,* 3**,** 54-54.

LÄMMERMANN, T. 2016. In the eye of the neutrophil swarm—navigation signals that bring neutrophils together in inflamed and infected tissues. *Journal of Leukocyte Biology,* 100**,** 55-63.

LÄMMERMANN, T., AFONSO, P. V., ANGERMANN, B. R., WANG, J. M., KASTENMÜLLER, W., PARENT, C. A. & GERMAIN, R. N. 2013. Neutrophil swarms require LTB4 and integrins at sites of cell death in vivo. *Nature,* 498**,** 371-375.

MACKAY, I. R. 2005. The Etiopathogenesis of Autoimmunity. *Semin Liver Dis,* 25**,** 239-250.

MALCHOW, S., LEVENTHAL, D. S., LEE, V., NISHI, S., SOCCI, N. D. & SAVAGE, P. A. 2016. Aire Enforces Immune Tolerance by Directing Autoreactive T Cells into the Regulatory T Cell Lineage. *Immunity,* 44**,** 1102-1113.

MANDL, J. N., MONTEIRO, J. P., VRISEKOOP, N. & GERMAIN, R. N. 2013. T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens. *Immunity,* 38**,** 263-274.

MARRACK, P., KAPPLER, J. & KOTZIN, B. L. 2001. Autoimmune disease: why and where it occurs. *Nature Medicine,* 7**,** 899-905.
MATHIS, D. & BENOIST, C. 2009. Aire. *Annual Review of Immunology,* 27**,** 287-312.
MATSUMOTO, M. 2011. Contrasting models for the roles of Aire in the differentiation program of epithelial cells in the thymic medulla. *European Journal of Immunology,* 41**,** 12-17.
MATZINGER, P. 1994. Tolerance, Danger, and the Extended Family. *Annual Review of Immunology,* 12**,** 991-1045.
MATZINGER, P. 2002. The Danger Model: A Renewed Sense of Self. *Science,* 296**,** 301.
MAZZA, C., BUZI, F., ORTOLANI, F., VITALI, A., NOTARANGELO, L. D., WEBER, G., BACCHETTA, R., SORESINA, A., LOUGARIS, V., GREGGIO, N. A., TADDIO, A., PASIC, S., DE VROEDE, M., PAC, M., KILIC, S. S., OZDEN, S., RUSCONI, R., MARTINO, S., CAPALBO, D., SALERNO, M., PIGNATA, C., RADETTI, G., MAGGIORE, G., PLEBANI, A., NOTARANGELO, L. D. & BADOLATO, R. 2011. Clinical heterogeneity and diagnostic delay of autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy syndrome. *Clinical Immunology,* 139**,** 6-11.
MEAGER, A., VISVALINGAM, K., PETERSON, P., MÖLL, K., MURUMÄGI, A., KROHN, K., ESKELIN, P., PERHEENTUPA, J., HUSEBYE, E., KADOTA, Y. & WILLCOX, N. 2006. Anti-Interferon Autoantibodies in Autoimmune Polyendocrinopathy Syndrome Type 1. *PLoS Medicine,* 3**,** e289.
MELONI, A., FIORILLO, E., CORDA, D., INCANI, F., SERRA, M. L., CONTINI, A., CAO, A. & ROSATELLI, M. C. 2010. DAXX Is a New AIRE-interacting Protein. *Journal of Biological Chemistry,* 285**,** 13012-13021.
MELONI, A., FURCAS, M., CETANI, F., MARCOCCI, C., FALORNI, A., PERNIOLA, R., PURA, M., BØE WOLFF, A. S., HUSEBYE, E. S., LILIC, D., RYAN, K. R., GENNERY, A. R., CANT, A. J., ABINUN, M., SPICKETT, G. P., ARKWRIGHT, P. D., DENNING, D., COSTIGAN, C., DOMINGUEZ, M., MCCONNELL, V., WILLCOX, N. & MEAGER, A. 2008. Autoantibodies against Type I Interferons as an Additional Diagnostic Criterion for Autoimmune Polyendocrine Syndrome Type I. *The Journal of Clinical Endocrinology & Metabolism,* 93**,** 4389-4397.
MEREDITH, M., ZEMMOUR, D., MATHIS, D. & BENOIST, C. 2015. Aire controls gene expression in the thymic epithelium with ordered stochasticity. *Nature Immunology,* 16**,** 942-949.
METZGER, T. C. & ANDERSON, M. S. 2011. Control of central and peripheral tolerance by Aire. *Immunological Reviews,* 241**,** 89-103.
MICHELS, A. W. & EISENBARTH, G. S. 2009. Autoimmune polyendocrine syndrome type 1 (APS-1) as a model for understanding autoimmune polyendocrine syndrome type 2 (APS-2). *Journal of Internal Medicine,* 265**,** 530-540.
MITTAZ, L., ROSSIER, C., HEINO, M., PETERSON, P., KROHN, K. J. E., GOS, A., MORRIS, M. A., KUDOH, J., SHIMIZU, N., ANTONARAKIS, S. E. & SCOTT, H. S. 1999. Isolation and Characterization of the MouseAireGene. *Biochemical and Biophysical Research Communications,* 255**,** 483-490.
MORETTA, A. 2002. Natural killer cells and dendritic cells: rendezvous in abused tissues. *Nature Reviews Immunology,* 2**,** 957-965.
MUJAL, A. M. & KRUMMEL, M. 2015. The subtle hands of self-reactivity in peripheral T cells. *Nature immunology,* 16**,** 10-11.
MUSSELMAN, C. A. & KUTATELADZE, T. G. 2011. Handpicking epigenetic marks with PHD fingers. *Nucleic Acids Research,* 39**,** 9061-9071.
MYERS, M. A., DAVIES, J. M., TONG, J. C., WHISSTOCK, J., SCEALY, M., MACKAY, I. R. & ROWLEY, M. J. 2000. Conformational Epitopes on the Diabetes Autoantigen GAD65 Identified by Peptide Phage Display and Molecular Modeling. *The Journal of Immunology,* 165**,** 3830.
NAGAMINE, K., PETERSON, P., SCOTT, H. S., KUDOH, J., MINOSHIMA, S., HEINO, M., KROHN, K. J. E., LALIOTI, M. D., MULLIS, P. E., ANTONARAKIS, S. E., KAWASAKI, K., ASAKAWA, S., ITO, F. & SHIMIZU, N. 1997. Positional cloning of the APECED gene. *Nat Genet,* 17**,** 393-398.

NEUFELD, M., MACLAREN, N. K. & BLIZZARD, R. M. 1981. Two Types of Autoimmune Addison's Disease Associated with Different Polyglandular Autoimmune (PGA) Syndromes. *Medicine,* 60.

NIKI, S. 2006. Alteration of intra-pancreatic target-organ specificity by abrogation of Aire in NOD mice. *J. Clin. Invest.,* 116**,** 1292-1301.

NOMURA, T. & SAKAGUCHI, S. 2007. Foxp3 and Aire in thymus-generated Treg cells: a link in self-tolerance. *Nature Immunology,* 8**,** 333-334.

O'GARRA, A. & VIEIRA, P. 2004. Regulatory T cells and mechanisms of immune system control. *Nature Medicine,* 10**,** 801-805.

OBERMAYER-STRAUB, P. & MANNS, M. P. 1998. 6 Autoimmune polyglandular syndromes. *Baillière's Clinical Gastroenterology,* 12**,** 293-315.

OIDA, T., ZHANG, X., GOTO, M., HACHIMURA, S., TOTSUKA, M., KAMINOGAWA, S. & WEINER, H. L. 2003. CD4+CD25- T cells that express latency-associated peptide on the surface suppress CD4+CD45RBhigh-induced colitis by a TGF-beta-dependent mechanism. *J Immunol,* 170**,** 2516-22.

OMORI, S. A., CATO, M. H., ANZELON-MILLS, A., PURI, K. D., SHAPIRO-SHELEF, M., CALAME, K. & RICKERT, R. C. 2006. Regulation of Class-Switch Recombination and Plasma Cell Differentiation by Phosphatidylinositol 3-Kinase Signaling. *Immunity,* 25**,** 545-557.

ORG, T., CHIGNOLA, F., HETÉNYI, C., GAETANI, M., REBANE, A., LIIV, I., MARAN, U., MOLLICA, L., BOTTOMLEY, M. J., MUSCO, G. & PETERSON, P. 2008. The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. *EMBO Reports,* 9**,** 370-376.

OVEN, I., BRDIČKOVÁ, N., KOHOUTEK, J., VAUPOTIČ, T., NARAT, M. & PETERLIN, B. M. 2007. AIRE Recruits P-TEFb for Transcriptional Elongation of Target Genes in Medullary Thymic Epithelial Cells. *Molecular and Cellular Biology,* 27**,** 8815-8823.

PASSOS, G. A., MENDES-DA-CRUZ, D. A. & OLIVEIRA, E. H. 2015. Editorial: The Role of Aire, microRNAs and Cell-Cell Interactions on Thymic Architecture and Induction of Tolerance. *Frontiers in immunology,* 6**,** 615-615.

PEMBERTON, L. F. & PASCHAL, B. M. 2005. Mechanisms of Receptor-Mediated Nuclear Import and Nuclear Export. *Traffic,* 6**,** 187-198.

PENNELL, L. M., GALLIGAN, C. L. & FISH, E. N. 2012. Sex affects immunity. *Journal of Autoimmunity,* 38**,** J282-J291.

PEREIRA, L. E., BOSTIK, P. & ANSARI, A. A. 2005. The development of mouse APECED models provides new insight into the role of AIRE in immune regulation. *Clinical & developmental immunology,* 12**,** 211-216.

PERHEENTUPA, J. 2006. Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy. *The Journal of Clinical Endocrinology & Metabolism,* 91**,** 2843-2850.

PERNIOLA, R. & MUSCO, G. 2014. The biophysical and biochemical properties of the autoimmune regulator (AIRE) protein. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease,* 1842**,** 326-337.

PERSAUD, S. P., PARKER, C. R., LO, W.-L., WEBER, K. S. & ALLEN, P. M. 2014. Intrinsic CD4+ T cell sensitivity and response to a pathogen are set and sustained by avidity for thymic and peripheral complexes of self peptide and MHC. *Nature immunology,* 15**,** 266-274.

PETERSON, P., ORG, T. & REBANE, A. 2008. Transcriptional regulation by AIRE: molecular mechanisms of central tolerance. *Nature reviews. Immunology,* 8**,** 948-957.

PETERSON, P. & PELTONEN, L. 2005. Autoimmune polyendocrinopathy syndrome type 1 (APS1) and AIRE gene: New views on molecular basis of autoimmunity. *Journal of Autoimmunity,* 25, Supplement**,** 49-55.

PITKÄNEN, J., DOUCAS, V., STERNSDORF, T., NAKAJIMA, T., ARATANI, S., JENSEN, K., WILL, H., VÄHÄMURTO, P., OLLILA, J., VIHINEN, M., SCOTT, H. S., ANTONARAKIS, S. E., KUDOH, J., SHIMIZU, N., KROHN, K. & PETERSON, P. 2000. The Autoimmune Regulator Protein Has

Transcriptional Transactivating Properties and Interacts with the Common Coactivator CREB-binding Protein. *Journal of Biological Chemistry,* 275**,** 16802-16809.

PLEVIN, M. J., MILLS, M. M. & IKURA, M. 2005. The LxxLL motif: a multifunctional binding sequence in transcriptional regulation. *Trends in Biochemical Sciences,* 30**,** 66-69.

PLOTZ, P. H. 2003. The autoantibody repertoire: searching for order. *Nature Reviews Immunology,* 3**,** 73-78.

PODKRAJŠEK, K. T., MILENKOVIĆ, T., ODINK, R. J., CLAASEN-VAN DER GRINTEN, H. L., BRATANIČ, N., HOVNIK, T. & BATTELINO, T. 2008. Detection of a complete autoimmune regulator gene deletion and two additional novel mutations in a cohort of patients with atypical phenotypic variants of autoimmune polyglandular syndrome type 1. *European Journal of Endocrinology,* 159**,** 633-639.

PUEL, A., CYPOWYJ, S., BUSTAMANTE, J., WRIGHT, J. F., LIU, L., LIM, H. K., MIGAUD, M., ISRAEL, L., CHRABIEH, M., AUDRY, M., GUMBLETON, M., TOULON, A., BODEMER, C., EL-BAGHDADI, J., WHITTERS, M., PARADIS, T., BROOKS, J., COLLINS, M., WOLFMAN, N. M., AL-MUHSEN, S., GALICCHIO, M., ABEL, L., PICARD, C. & CASANOVA, J.-L. 2011. Chronic mucocutaneous candidiasis in humans with inborn errors of interleukin-17 immunity*. *Science (New York, N.Y.),* 332**,** 65-68.

PUEL, A., DÖFFINGER, R., NATIVIDAD, A., CHRABIEH, M., BARCENAS-MORALES, G., PICARD, C., COBAT, A., OUACHÉE-CHARDIN, M., TOULON, A., BUSTAMANTE, J., AL-MUHSEN, S., AL-OWAIN, M., ARKWRIGHT, P. D., COSTIGAN, C., MCCONNELL, V., CANT, A. J., ABINUN, M., POLAK, M., BOUGNÈRES, P.-F., KUMARARATNE, D., MARODI, L., NAHUM, A., ROIFMAN, C., BLANCHE, S., FISCHER, A., BODEMER, C., ABEL, L., LILIC, D. & CASANOVA, J.-L. 2010. Autoantibodies against IL-17A, IL-17F, and IL-22 in patients with chronic mucocutaneous candidiasis and autoimmune polyendocrine syndrome type I. *The Journal of Experimental Medicine,* 207**,** 291-297.

PÖNTYNEN, N., MIETTINEN, A., PETTERI ARSTILA, T., KÄMPE, O., ALIMOHAMMADI, M., VAARALA, O., PELTONEN, L. & ULMANEN, I. 2006. Aire deficient mice do not develop the same profile of tissue-specific autoantibodies as APECED patients. *Journal of Autoimmunity,* 27**,** 96-104.

RAMSEY, C. 2002. Aire deficient mice develop multiple features of APECED phenotype and show altered immune response. *Hum. Mol. Genet.,* 11**,** 397-409.

RAMSEY, C., HÄSSLER, S., MARITS, P., KÄMPE, O., SURH, C. D., PELTONEN, L. & WINQVIST, O. 2006. Increased antigen presenting cell-mediated T cell activation in mice and patients without the autoimmune regulator. *European Journal of Immunology,* 36**,** 305-317.

RATTAY, K., CLAUDE, J., REZAVANDY, E., MATT, S., HOFMANN, T. G., KYEWSKI, B. & DERBINSKI, J. 2015. Homeodomain-Interacting Protein Kinase 2, a Novel Autoimmune Regulator Interaction Partner, Modulates Promiscuous Gene Expression in Medullary Thymic Epithelial Cells. *The Journal of Immunology,* 194**,** 921-928.

RAUTEMAA, R., HIETANEN, J., NIISSALO, S., PIRINEN, S. & PERHEENTUPA, J. 2007. Oral and oesophageal squamous cell carcinoma – A complication or component of autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED, APS-I). *Oral Oncology,* 43**,** 607-613.

RILEY, W. J. 1992. Autoimmune Polyglandular Syndromes. *Hormone Research in Paediatrics,* 38(suppl 2)**,** 9-15.

ROSENBAUM, A. I. & MAXFIELD, F. R. 2011. Niemann-Pick type C disease: molecular mechanisms and potential therapeutic approaches. *Journal of neurochemistry,* 116**,** 789-795.

RYAN, K. R., HONG, M., ARKWRIGHT, P. D., GENNERY, A. R., COSTIGAN, C., DOMINGUEZ, M., DENNING, D., MCCONNELL, V., CANT, A. J., ABINUN, M., SPICKETT, G. P. & LILIC, D. 2008. Impaired dendritic cell maturation and cytokine production in patients with chronic mucocutanous candidiasis with or without APECED. *Clinical & Experimental Immunology,* 154**,** 406-414.

SAARE, M., REBANE, A., RAJASHEKAR, B., VILO, J. & PETERSON, P. 2012. Autoimmune regulator is acetylated by transcription coactivator CBP/p300. *Experimental Cell Research,* 318**,** 1767-1778.

SABATER, L., FERRER-FRANCESCH, X., SOSPEDRA, M., CARO, P., JUAN, M. & PUJOL-BORRELL, R. 2005. Insulin alleles and autoimmune regulator (AIRE) gene expression both influence insulin expression in the thymus. *Journal of Autoimmunity,* 25**,** 312-318.

SEGAL, A. W. 2004. HOW NEUTROPHILS KILL MICROBES. *Annual Review of Immunology,* 23**,** 197-223.

SFRISO, P., GHIRARDELLO, A., BOTSIOS, C., TONON, M., ZEN, M., BASSI, N., BASSETTO, F. & DORIA, A. 2010. Infections and autoimmunity: the multifaceted relationship. *Journal of Leukocyte Biology,* 87**,** 385-395.

SILLANPÄÄ, N., MAGUREANU, C. G., MURUMÄGI, A., REINIKAINEN, A., WEST, A., MANNINEN, A., LAHTI, M., RANKI, A., SAKSELA, K., KROHN, K., LAHESMAA, R. & PETERSON, P. 2004. Autoimmune regulator induced changes in the gene expression profile of human monocyte-dendritic cell-lineage. *Molecular Immunology,* 41**,** 1185-1198.

SOMMER, S. 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in zoology,* 2**,** 16-16.

SONAL, C., MICHAEL, M., DANIELE, T. & PAOLO, R. 2012. Autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy. *The Journal of clinical and aesthetic dermatology,* 5**,** 18-22.

SPARKS, A. E., CHEN, C., BRESLIN, M. B. & LAN, M. S. 2016. Functional Domains of Autoimmune Regulator (AIRE) Modulate INS-VNTR Transcription in Human Thymic Epithelial Cells. *The Journal of biological chemistry,* 291**,** 11313-11322.

SPRENT, J. & SURH, C. D. 2002. T Cell Memory. *Annual Review of Immunology,* 20**,** 551-579.

TAKABA, H., MORISHITA, Y., TOMOFUJI, Y., DANKS, L., NITTA, T., KOMATSU, N., KODAMA, T. & TAKAYANAGI, H. 2015. Fezf2 Orchestrates a Thymic Program of Self-Antigen Expression for Immune Tolerance. *Cell,* 163**,** 975-987.

TARAKHOVSKY, A., KANNER, S. B., HOMBACH, J., LEDBETTER, J. A., MULLER, W., KILLEEN, N. & RAJEWSKY, K. 1995. A role for CD5 in TCR-mediated signal transduction and thymocyte selection. *Science,* 269**,** 535.

TINIAKOU, E., COSTENBADER, K. H. & KRIEGEL, M. A. 2013. Sex-specific environmental influences on the development of autoimmune diseases. *Clinical immunology (Orlando, Fla.),* 149**,** 182-191.

TONOOKA, A., KUBO, T., ICHIMIYA, S., TAMURA, Y., ILMARINEN, T., ULMANEN, I., KIMURA, S., YOKOYAMA, S., TAKANO, Y., KIKUCHI, T. & SATO, N. 2009. Wild-type AIRE cooperates with p63 in HLA class II expression of medullary thymic stromal cells. *Biochemical and Biophysical Research Communications,* 379**,** 765-770.

UETZ, P., GIOT, L., CAGNEY, G., MANSFIELD, T. A., JUDSON, R. S., KNIGHT, J. R., LOCKSHON, D., NARAYAN, V., SRINIVASAN, M., POCHART, P., QURESHI-EMILI, A., LI, Y., GODWIN, B., CONOVER, D., KALBFLEISCH, T., VIJAYADAMODAR, G., YANG, M., JOHNSTON, M., FIELDS, S. & ROTHBERG, J. M. 2000. A comprehensive analysis of protein–protein interactions in Saccharomyces cerevisiae. *Nature,* 403**,** 623-627.

ULMANEN, I., HALONEN, M., ILMARINEN, T. & PELTONEN, L. 2005. Monogenic autoimmune diseases — lessons of self-tolerance. *Current Opinion in Immunology,* 17**,** 609-615.

VAN LITH, M., MCEWEN-SMITH, R. M. & BENHAM, A. M. 2010. HLA-DP, HLA-DQ, and HLA-DR Have Different Requirements for Invariant Chain and HLA-DM. *Journal of Biological Chemistry,* 285**,** 40800-40808.

VENANZI, E. S., MELAMED, R., MATHIS, D. & BENOIST, C. 2008. The variable immunological self: genetic variation and nongenetic noise in Aire-regulated transcription. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 15860-15865.

VIVIER, E. & MEDZHITOV, R. 2016. Editorial overview: Innate immunity. *Current Opinion in Immunology,* 38**,** v-vii.

VOGEL, A., STRASSBURG, C., OBERMAYER-STRAUB, P., BRABANT, G. & MANNS, M. 2002. The genetic background of autoimmune polyendocrinopathy–candidiasis–ectodermal dystrophy and its autoimmune disease components. *Journal of Molecular Medicine,* 80**,** 201-211.

VRISEKOOP, N., MONTEIRO, JOÃO P., MANDL, JUDITH N. & GERMAIN, RONALD N. 2014. Revisiting Thymic Positive Selection and the Mature T Cell Repertoire for Antigen. *Immunity,* 41**,** 181-190.

WAN, C. H., BORGESON, B., PHANSE, S., TU, F., DREW, K., CLARK, G., XIONG, X. J., KAGAN, O., KWAN, J., BEZGINOV, A., CHESSMAN, K., PAL, S., CROMAR, G., PAPOULAS, O., NI, Z. Y., BOUTZ, D. R., STOILOVA, S., HAVUGIMANA, P. C., GUO, X. H., MALTY, R. H., SAROV, M., GREENBLATT, J., BABU, M., DERRY, W. B., TILLIER, E. R., WALLINGFORD, J. B., PARKINSON, J., MARCOTTE, E. M. & EMILI, A. 2015. Panorama of ancient metazoan macromolecular complexes. *Nature,* 525**,** 339-+.

WANG, H., WANG, L., ERDJUMENT-BROMAGE, H., VIDAL, M., TEMPST, P., JONES, R. S. & ZHANG, Y. 2004. Role of histone H2A ubiquitination in Polycomb silencing. *Nature,* 431**,** 873-878.

WATERFIELD, M., KHAN, I. S., CORTEZ, J. T., FAN, U., METZGER, T., GREER, A., FASANO, K., MARTINEZ-LLORDELLA, M., POLLACK, J. L., ERLE, D. J., SU, M. & ANDERSON, M. S. 2014. The transcriptional regulator Aire coopts the repressive ATF7ip-MBD1 complex for the induction of immunotolerance. *Nature immunology,* 15**,** 258-265.

WHITE, A. J., NAKAMURA, K., JENKINSON, W. E., SAINI, M., SINCLAIR, C., SEDDON, B., NARENDRAN, P., PFEFFER, K., NITTA, T., TAKAHAMA, Y., CAAMANO, J. H., LANE, P. J. L., JENKINSON, E. J. & ANDERSON, G. 2010. Lymphotoxin Signals from Positively Selected Thymocytes Regulate the Terminal Differentiation of Medullary Thymic Epithelial Cells. *The Journal of Immunology,* 185**,** 4769-4776.

WITEBSKY, E., ROSE, N. R., TERPLAN, K., PAINE, J. R. & EGAN, R. W. 1957. CHRONIC THYROIDITIS AND AUTOIMMUNIZATION. *Journal of the American Medical Association,* 164**,** 1439-1447.

YANG, S., BANSAL, K., LOPES, J., BENOIST, C. & MATHIS, D. 2013. Aire's plant homeodomain(PHD)-2 is critical for induction of immunological tolerance. *Proceedings of the National Academy of Sciences,* 110**,** 1833-1838.

YANG, S., FUJIKADO, N., KOLODIN, D., BENOIST, C. & MATHIS, D. 2015. Regulatory T cells generated early in life play a distinct role in maintaining self-tolerance. *Science,* 348**,** 589.

YANO, M., KURODA, N., HAN, H., MEGURO-HORIKE, M., NISHIKAWA, Y., KIYONARI, H., MAEMURA, K., YANAGAWA, Y., OBATA, K., TAKAHASHI, S., IKAWA, T., SATOH, R., KAWAMOTO, H., MOURI, Y. & MATSUMOTO, M. 2008. Aire controls the differentiation program of thymic epithelial cells in the medulla for the establishment of self-tolerance. *The Journal of Experimental Medicine,* 205**,** 2827-2838.

YOSHIDA, H., BANSAL, K., SCHAEFER, U., CHAPMAN, T., RIOJA, I., PROEKT, I., ANDERSON, M. S., PRINJHA, R. K., TARAKHOVSKY, A., BENOIST, C. & MATHIS, D. 2015. Brd4 bridges the transcriptional regulators, Aire and P-TEFb, to promote elongation of peripheral-tissue antigen transcripts in thymic stromal cells. *Proceedings of the National Academy of Sciences,* 112**,** E4448.

ZLOTOGORA, J. & SHAPIRO, M. S. 1992. Polyglandular autoimmune syndrome type I among Iranian Jews. *Journal of Medical Genetics,* 29**,** 824.

ZUKLYS, S., BALCIUNAITE, G., AGARWAL, A., FASLER-KAN, E., PALMER, E. & HOLLÄNDER, G. A. 2000. Normal Thymic Architecture and Negative Selection Are Associated with Aire Expression, the Gene Defective in the Autoimmune-Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy (APECED). *The Journal of Immunology,* 165**,** 1976-1983.

ŽUMER, K., LOW, A. K., JIANG, H., SAKSELA, K. & PETERLIN, B. M. 2012. Unmodified Histone H3K4 and DNA-Dependent Protein Kinase Recruit Autoimmune Regulator to Target Genes. *Molecular and Cellular Biology,* 32**,** 1354.

ŽUMER, K., SAKSELA, K. & PETERLIN, B. M. 2013. The Mechanism of Tissue-Restricted Antigen Gene Expression by AIRE. *The Journal of Immunology,* 190**,** 2479-2482.

Chapter 7
Appendices

7. Appendices

7.1 In silico sequence of the precursor full-length AIRE cDNA clone

>pCMV-MYC AIRE

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGAGTACA
TTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTG
GCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGA
CTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATA
TGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCT
TACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAA
TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCCTCACTC
TCTTCCGCATCGCTGTCTGCGAGGGCCAGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCC
AGTACTCTTGGATCGGAAACCCGTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACCTGAGCGAGTCCGCA
TCGACCGGATCGGAAAACCTCTCGAGAAAGGCGTCTAACCAGTCACAGTCGCAAGGTAGGCTGAGCACCGTG
GCGGGCGGCAGCGGGTGGCGGTCGGGGTTGTTTCTGGCGGAGGTGCTGCTGATGATGTAATTAAAGTAGGCG
GTCTTGAGACGGCGGATGGTCGAGGTGAGGTGTGGCAGGCTTGAGATCCAGCTGTTGGGGTGAGTACTCCCT
CTCAAAAGCGGGCATTACTTCTGCGCTAAGATTGTCAGTTTCCAAAAACGAGGAGGATTTGATATTCACCTG
GCCCGATCTGGCCATACACTTGAGTGACAATGACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCA
GGTCCAAGTTTAAACTTTAATACGACTCACTATAGGGGCCGCCACCAAGCTTATGGCGACGGACGCGGCGCT
ACGCCGGCTTCTGAGGCTGCACCGCACGGAGATCGCGGTGGCCGTGGACAGCGCCTTCCCACTGCTGCACGC
GCTGGCTGACCACGACGTGGTCCCCGAGGACAAGTTTCAGGAGACGCTTCATCTGAAGGAAAAGGAGGGCTG
CCCCCAGGCCTTCCACGCCCTCCTGTCCTGGCTGCTGACCCAGGACTCCACAGCCATCCTGGACTTCTGGAG
GGTGCTGTTCAAGGACTACAACCTGGAGCGCTATGGCCGGCTGCAGCCCATCCTGGACAGCTTCCCCAAAGA
TGTGGACCTCAGCCAGCCCCGGAAGGGGAGGAAGCCCCCGGCCGTCCCCAAGGCTTTGGTACCGCCACCCAG
ACTCCCCACCAAGAGGAAGGCCTCAGAAGAGGCTCGAGCTGCCGCGCCAGCAGCCCTGACTCCAAGGGGCAC
CGCCAGCCCAGGCTCTCAACTGAAGGCCAAGCCCCCCAAGAAGCCGGAGAGCAGCGCAGAGCAGCAGCGCCT
TCCACTCGGGAACGGGATTCAGACCATGTCAGCTTCAGTCCAGAGAGCTGTGGCCATGTCCTCCGGGGACGT
CCCGGGAGCCCGAGGGGCCGTGGAGGGGATCCTCATCCAGCAGGTGTTTGAGTCAGGCGGCTCCAAGAAGTG
CATCCAGGTTGGCGGGGAGTTCTACACTCCCAGCAAGTTCGAAGACTCCGGCAGTGGGAAGAACAAGGCCCG
CAGCAGCAGTGGCCCGAAGCCTCTGGTTCGAGCCAAGGGAGCCCAGGGCGCTGCCCCCGGTGGAGGTGAGGC
TAGGCTGGGCCAGCAGGGCAGCGTTCCCGCCC

primers to clone AIRE-PHD1 in pGBKT7

CTCTGGCCCTCCCCAGTGACCCCCAGCTCCACCAGAAGAATGAGGACGAGTGTGCCGTGTGTCGGGACGGCG
GGGAGCTCATCTGCTGTGACGGCTGCCCTCGGGCCTTCCACCTGGCCTGCCTGTCCCCTCCGCTCCGGGAGA
TCCCCAGTGGGACCTGGAGGTGCTCCAGCTGCCTGCAGGCAACAGTCCAGGAGGTGCAGCCCCGGGCAGAGG
AGCCCCGGCCCCAGGAGCCACCCGTGGAGACCCCGCTCCCCCCGGGGCTTAGGTCGGCGGGAGAGGAGGTAA
GAGGTCCACCTGGGGAACCCCTAGCCGGCATGGACACGACTCTTGTCTACAAGCACCTGCCGGCTCCGCCTT
CTGCAGCCCCGCTGCCAGGGCTGGACTCCTCGGCCCTGCACCCCCTACTGTGTGTGGGTCCTGAGGGTCAGC
AGAACCTGGCTCCTGGTGCGCGTTGCGGGGTGTGCGGAGATGGTACGGACGTGCTGCGGTGTACTCACTGCG
CCGCTGCCTTCCACTGGCGCTGCCACTTCCCAGCCGGCACCTCCCGGCCCGGGACGGGCCTGCGCTGCAGAT
CCTGCTCAGGAGACGTGACCCCAGCCCCTGTGGAGGGGGTGCTGGCCCCCAGCCCCGCCCGCCTGGCCCCTG
GGCCTGCCAAGGATGACACTGCCAGTCACGAGCCCGCTCTGCACAGGGATGACCTGGAGTCCCTTCTGAGCG
AGCACACCTTCGATGGCATCCTGCAGTGGGCCATCCAGAGCATGGCCCGTCCGGCGGCCCCCTTCCCCTCCG
GGGGTGGAGGCTCTGAGCAGAAACTCATCTCAGAAGAGGATCTGTAAGCGGCCGCCGAATTCGGGCCCGTTT
AAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTT
CCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGA
GTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCA
GGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATC
CCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTG
CCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCAGGCTTTCCCCGTC
AAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTG
ATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCA

CGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATT
TATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATT
AATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAG
CATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAG
CATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAG
TTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTC
TGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCTCGGGAGCTTG
TATATCCATTTTCGGATCTGATCAGCACGTGATGAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTT
TCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAG
CTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTA
TGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAG
CCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGC
TGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGG
CCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCA
TGTGTATCACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGAT
GCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGAC
GGACAATGGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGC
CAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCC
GGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTT
GGTTGACGGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGG
GACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGC
CGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTGCTACGAGATTTCGATTC
CACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCTGGCTGGATGATCCTCCAGCG
CGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAG
CAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCAT
CAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTT
TCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTG
GGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCT
GTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC
TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGT
AATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG
GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCG
ACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCT
CGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC
GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCA
CGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACA
CGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGA
GTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCC
AGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT
TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTC
TGAGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAATTCTT
AGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAA
AAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGT
CTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTG
AGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTT
CAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCG
CCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCA
GGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTT
TCCCAGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAG
GCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCAT
GTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACAT
TATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACG
TTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATG
ACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAATCCGCGCACATTTCCCC
GAAAAGTGCCACCTGACGTC

7.2 In silico sequence of the pGBKT7 PHD1 used in Y2H screening

>pGBKT7 PHD1

CGGTGCGGGCCTCTTCGCTATTACGCCAGATCCTTTTGTTGTTTCCGGGTGTACAATATGGACTTCCTCTTT
TCTGGCAACCAAACCCATACATCGGGATTCCTATAATACCTTCGTTGGTCTCCCTAACATGTAGGTGGCGGA
GGGGAGATATACAATAGAACAGATACCAGACAAGACATAATGGGCTAAACAAGACTACACCAATTACACTGC
CTCATTGATGGTGGTACATAACGAACTAATACTGTAGCCCTAGACTTGATAGCCATCATCATATCGAAGTTT
CACTACCCTTTTTCCATTTGCCATCTATTGAAGTAATAATAGGCGCATGCAACTTCTTTTCTTTTTTTTTCT
TTTCTCTCTCCCCCGTTGTTGTCTCACCATATCCGCAATGACAAAAAAATGATGGAAGACACTAAAGGAAAA
AATTAACGACAAAGACAGCACCAACAGATGTCGTTGTTCCAGAGCTGATGAGGGGTATCTCGAAGCACACGA
AACTTTTTCCTTCCTTCATTCACGCACACTACTCTCTAATGAGCAACGGTATACGGCCTTCCTTCCAGTTAC
TTGAATTTGAAATAAAAAAAAGTTTGCTGTCTTGCTATCAAGTATAAATAGACCTGCAATTATTAATCTTTT
GTTTCCTCGTCATTGTTCTCGTTCCCTTTCTTCCTTGTTTCTTTTTCTGCACAATATTTCAAGCTATACCAA
GCATACAATCAACTCCAAGCTTGAAGCAAGCCTCCTGAAAGATGAAGCTACTGTCTTCTATCGAACAAGCAT
GCGATATTTGCCGACTTAAAAAGCTCAAGTGCTCCAAAGAAAAACCGAAGTGCGCCAAGTGTCTGAAGAACA
ACTGGGAGTGTCGCTACTCTCCCAAAACCAAAAGGTCTCCGCTGACTAGGGCACATCTGACAGAAGTGGAAT
CAAGGCTAGAAAGACTGGAACAGCTATTTCTACTGATTTTTCCTCGAGAAGACCTTGACATGATTTTGAAAA
TGGATTCTTTACAGGATATAAAAGCATTGTTAACAGGATTATTTGTACAAGATAATGTGAATAAAGATGCCG
TCACAGATAGATTGGCTTCAGTGGAGACTGATATGCCTCTAACATTGAGACAGCATAGAATAAGTGCGACAT
CATCATCGGAAGAGAGTAGTAACAAAGGTCAAAGACAGTTGACTGTATCGCCGGAATTTGTAATACGACTCA
CTATAGGGCGAGCCGCCATC

Primers to clone AIRE-PHD1 in pcDNA5/FRT

ATGGAGGAGCAGAAGCTGATCTCAGAGGAGGACCTGCATATGGCCATGGCTCTGGCCCTCCCCAGTGACCCC
CAGCTCCACCAGAAGAATGAGGACGAGTGTGCCGTGTGTCGGGACGGCGGGGAGCTCATCTGCTGTGACGGC
TGCCCTCGGGCCTTCCACCTGGCCTGCCTGTCCCCTCCGCTCCGGGAGATCCCCAGTGGGACCTGGAGGTGC
TCCAGCTGCCTGCAGGCAACAGTCCAGGAGGTGCAGGTCGACCTGCAGCGGCCGCATAACTAGCATAACCCC
TTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCGCGCTTGCAGCCAAGCTAATTCCGGGCGAATTTCTT
ATGATTTATGATTTTTATTATTAAATAAGTTATAAAAAAAATAAGTGTATACAAATTTTAAAGTGACTCTTA
GGTTTTAAAACGAAAATTCTTATTCTTGAGTAACTCTTTCCTGTAGGTCAGGTTGCTTTCTCAGGTATAGCA
TGAGGTCGCTCTTATTGACCACACCTCTACCGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCT
GTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGT
GCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCG
TGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCC
TCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATA
CGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAC
CGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGC
TCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG
CGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTT
TCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAA
CCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGAC
TTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTC
TTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT
ACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTT
TGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGAC
GCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATC
CTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACCTGAGGCTATGGCAGGGCCT
GCCGCCCCGACGTTGGCTGCGAGCCCTGGGCCTTCACCCGAACTTGGGGGGTGGGGTGGGGAAAAGGAAGAA
ACGCGGGCGTATTGGCCCCAATGGGGTCTCGGTGGGGTATCGACAGAGTGCCAGCCCTGGGACCGAACCCCG
CGTTTATGAACAAACGACCCAACACCGTGCGTTTTATTCTGTCTTTTTATTGCCGTCATAGCGCGGGTTCCT
TCCGGTATTGTCTCCTTCCGTGTTTCAGTTAGCCTCCCCCTAGGGTGGGCGAAGAACTCCAGCATGAGATCC
CCGCGCTGGAGGATCATCCAGCCGGCGTCCCGGAAAACGATTCCGAAGCCCAACCTTTCATAGAAGGCGGCG
GTGGAATCGAAATCTCGTGATGGCAGGTTGGGCGTCGCTTGGTCGGTCATTTCGAACCCCAGAGTCCCGCTC
AGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGA
GGAAGCGGTCAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGC
GGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGGCA
AGCAGGCATCGCCATGAGTCACGACGAGATCCTCGCCGTCGGGCATGCTCGCCTTGAGCCTGGCGAACAGTT

CGGCTGGCGCGAGCCCCTGATGCTCTTCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTAC
GTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCC
GCATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGGCA
CTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAGCACAGCTGCGCAAGGAACGCCCG
TCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTTGCAGTTCATTCAGGGCACCGGACAGGTCGGTCTTGA
CAAAAAGAACCGGGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTG
CCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGCAATCCATCTTGTTCAATCA
TACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAAT
GTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGAACGAAGCATCT
GTGCTTCATTTTGTAGAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACAAAGAATCTGAGCTGCATT
TTTACAGAACAGAAATGCAACGCGAAAGCGCTATTTTACCAACGAAGAATCTGTGCTTCATTTTTGTAAAAC
AAAAATGCAACGCGAGAGCGCTAATTTTTCAAACAAAGAATCTGAGCTGCATTTTTACAGAACAGAAATGCA
ACGCGAGAGCGCTATTTTACCAACAAAGAATCTATACTTCTTTTTTGTTCTACAAAAATGCATCCCGAGAGC
GCTATTTTTCTAACAAAGCATCTTAGATTACTTTTTTTCTCCTTTGTGCGCTCTATAATGCAGTCTCTTGAT
AACTTTTTGCACTGTAGGTCCGTTAAGGTTAGAAGAAGGCTACTTTGGTGTCTATTTTCTCTTCCATAAAAA
AAGCCTGACTCCACTTCCCGCGTTTACTGATTACTAGCGAAGCTGCGGGTGCATTTTTTCAAGATAAAGGCA
TCCCCGATTATATTCTATACCGATGTGGATTGCGCATACTTTGTGAACAGAAAGTGATAGCGTTGATGATTC
TTCATTGGTCAGAAAATTATGAACGGTTTCTTCTATTTTGTCTCTATATACTACGTATAGGAAATGTTTACA
TTTTCGTATTGTTTTCGATTCACTCTATGAATAGTTCTTACTACAATTTTTTTGTCTAAAGAGTAATACTAG
AGATAAACATAAAAAATGTAGAGGTCGAGTTTAGATGCAAGTTCAAGGAGCGAAAGGTGGATGGGTAGGTTA
TATAGGGATATAGCACAGAGATATATAGCAAAGAGATACTTTTGAGCAATGTTTGTGGAAGCGGTATTCGCA
ATATTTTAGTAGCTCGTTACAGTCCGGTGCGTTTTTGGTTTTTTGAAAGTGCGTCTTCAGAGCGCTTTTGGT
TTTCAAAAGCGCTCTGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTTCAAAGCGTTT
CCGAAAACGAGCGCTTCCGAAAATGCAACGCGAGCTGCGCACATACAGCTCACTGTTCACGTCGCACCTATA
TCTGCGTGTTGCCTGTATATATATATACATGAGAAGAACGGCATAGTGCGTGTTTATGCTTAAATGCGTACT
TATATGCGTCTATTTATGTAGGATGAAAGGTAGTCTAGTACCTCCTGTGATATTATCCCATTCCATGCGGGG
TATCGTATGCTTCCTTCAGCACTACCCTTTAGCTGTTCTATATGCTGCCACTCCTCAATTGGATTAGTCTCA
TCCTTCAATGCTATCATTTCCTTTGATATTGGATCATACTAAGAAACCATTATTATCATGACATTAACCTAT
AAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATG
CAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCA
GCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCAT
AGATCAACGACATTACTATATATATAATATAGGAAGCATTTAATAGAACAGCATCGTAATATATGTGTACTT
TGCAGTTATGACGCCAGATGGCAGTAGTGGAAGATATTCTTTATTGAAAAATAGCTTGTCACCTTACGTACA
ATCTTGATCCGGAGCTTTTCTTTTTTTGCCGATTAAGAATTAATTCGGTCGAAAAAAGAAAAGGAGAGGGCC
AAGAGGGAGGGCATTGGTGACTATTGAGCACGTGAGTATACGTGATTAAGCACACAAAGGCAGCTTGGAGTA
TGTCTGTTATTAATTTCACAGGTAGTTCTGGTCCATTGGTGAAAGTTTGCGGCTTGCAGAGCACAGAGGCCG
CAGAATGTGCTCTAGATTCCGATGCTGACTTGCTGGGTATTATATGTGTGCCCAATAGAAAGAGAACAATTG
ACCCGGTTATTGCAAGGAAAATTTCAAGTCTTGTAAAAGCATATAAAAATAGTTCAGGCACTCCGAAATACT
TGGTTGGCGTGTTTCGTAATCAACCTAAGGAGGATGTTTTGGCTCTGGTCAATGATTACGGCATTGATATCG
TCCAACTGCATGGAGATGAGTCGTGGCAAGAATACCAAGAGTTCCTCGGTTTGCCAGTTATTAAAAGACTCG
TATTTCCAAAAGACTGCAACATACTACTCAGTGCAGCTTCACAGAAACCTCATTCGTTTATTCCCTTGTTTG
ATTCAGAAGCAGGTGGGACAGGTGAACTTTTGGATTGGAACTCGATTTCTGACTGGGTTGGAAGGCAAGAGA
GCCCCGAAAGCTTACATTTTATGTTAGCTGGTGGACTGACGCCAGAAAATGTTGGTGATGCGCTTAGATTAA
ATGGCGTTATTGGTGTTGATGTAAGCGGAGGTGTGGAGACAAATGGTGTAAAAGACTCTAACAAAATAGCAA
ATTTCGTCAAAAATGCTAAGAAATAGGTTATTACTGAGTAGTATTTATTTAAGTATTGTTTGTGCACTTGCC
GATCTATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGAAATTGTAAGCGTTAA
TATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAA
AATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACT
ATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACC
ATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCG
ATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGC
TAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACA
GGGCGCGTCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGAT

7.3 In silico sequence of AIRE-PHD1 cDNA cloned in pcDNA5/FRT

>pcDNA5/FRT AIRE-PHD1 cDNA

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA
ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCAT
GGAGGAGCAGAAGCTGATCTCAGAGGAGGACCTGCATATGGCCATGGACTCTGGCCCTCCCCAGTGACCCCC
AGCTCCACCAGAAGAATGAGGACGAGTGTGCCGTGTGTCGGGACGGCGGGGAGCTCATCTGCTGTGACGGCT
GCCCTCGGGCCTTCCACCTGGCCTGCCTGTCCCCTCCGCTCCGGGAGATCCCCAGTGGGACCTGGAGGTGCT
CCAGCTGCCTGTAATGAGCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTG
CCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCC
ACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGT
GGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCT
ATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTA
AGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC
GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTA
GGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTACCTAG
AAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCCTTGGCCAAAAAGCCTGAACTCACC
GCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGC
GAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGAT
GGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGAC
ATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTG
CCTGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTT
AGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATA
TGCGCGATTGCTGATCCCCATGTGTATCACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCG
CAGGCTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTC
GGCTCCAACAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGG
GATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGC
TACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTT
GACCAACTCTATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCA
ATCGTCCGATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGAT
GGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCAC
GTACTACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCC
GGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCT
TATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGT
TGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCG
TAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGA
AGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCC
GCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTG
CGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTA
TCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCA
AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCT
GACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG
TTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTT
CTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGC
TCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTT
GAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGG
TATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGT
ATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCT
TTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTA
TCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAG

TAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCA
TCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCT
GCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCC
GAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTA
AGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCG
TTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAA
AAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTT
ATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCA
ACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACC
GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATC
TTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTC
ACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAA
TGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGA
TACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCT
GACGTC

7.4 In silico sequences of the prey cDNA samples isolated from the Y2H screening

> pGADT7-RecAB ATP1β1

TGCATGCCTGCAGGTCGAGATCCGGGATCGAAGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATGAA
GGCAAAAGACAAATATAAGGGTCGAACGAAAAATAAAGTGAAAAGTGTTGATATGATGTATTTGGCTTTGCG
GCGCCGAAAAAACGAGTTTACGCAATTGCACAATCATGCTGACTCTGTGGCGGACCCGCGCTCTTGCCGGCC
CGGCGATAACGCTGGGCGTGAGGCTGTGCCCGGCGGAGTTTTTTGCGCCTGCATTTTCCAAGGTTTACCCTG
CGCTAAGGGGCGAGATTGGAGAAGCAATAAGAATGCCGGTTGGGGTTGCGATGATGACGACCACGACAACTG
GTGTCATTATTTAAGTTGCCGAAAGAACCTGAGTGCATTTGCAACATGAGTATACTAGAAGAATGAGCCAAG
ACTTGCGAGACGCGAGTTTGCCGGTGGTGCGAACAATAGAGCGACCATGACCTTGAAGGTGAGACGCGCATA
ACCGCTAGAGTACTTTGAAGAGGAAACAGCAATAGGGTTGCTACCAGTATAAATAGACAGGTACATACAACA
CTGGAAATGGTTGTCTGTTTGAGTACGCTTTCAATTCATTTGGGTGTGCACTTTATTATGTTACAATATGGA
AGGGAACTTTACACTTCTCCTATGCACATATATTAATTAAAGTCCAATGCTAGTAGAGAAGGGGGGTAACAC
CCCTCCGCGCTCTTTTCCGATTTTTTTCTAAACCGTGGAATATTTCGGATATCCTTTTGTTGTTTCCGGGTG
TACAATATGGACTTCCTCTTTTCTGGCAACCAAACCCATACATCGGGATTCCTATAATACCTTCGTTGGTCT
CCCTAACATGTAGGTGGCGGAGGGGAGATATACAATAGAACAGATACCAGACAAGACATAATGGGCTAAACA
AGACTACACCAATTACACTGCCTCATTGATGGTGGTACATAACGAACTAATACTGTAGCCCTAGACTTGATA
GCCATCATCATATCGAAGTTTCACTACCCTTTTTCCATTTGCCATCTATTGAAGTAATAATAGGCGCATGCA
ACTTCTTTTCTTTTTTTTTCTTTTCTCTCTCCCCCGTTGTTGTCTCACCATATCCGCAATGACAAAAAAATG
ATGGAAGACACTAAAGGAAAAAATTAACGACAAAGACAGCACCAACAGATGTCGTTGTTCCAGAGCTGATGA
GGGGTATCTCGAAGCACACGAAACTTTTTCCTTCCTTCATTCACGCACACTACTCTCTAATGAGCAACGGTA
TACGGCCTTCCTTCCAGTTACTTGAATTTGAAATAAAAAAAAAGTTTGCTGTCTTGCTATCAAGTATAAATAG
ACCTGCAATTATTAATCTTTTGTTTCCTCGTCATTGTTCTCGTTCCCTTTCTTCCTTGTTTCTTTTTCTGCA
CAATATTTCAAGCTATACCAAGCATACAATCAACTCCAAGCTTTGCAAAGATGGATAAAGCGGAATTAATTC
CCGAGCCTCCAAAAAAGAAGAGAAAGGTCGAATTGGGTACCGCCGCCAATTTTAATCAAAGTGGGAATATTG
CTGATAGCTCATTGTCCTTCACTTTCACTAACAGTAGCAACGGTCCGAACCTCATAACAACTCAAACAAATT
CTCAAGCGCTTTCACAACCAATTGCCTCCTCTAACGTTCATGATAACTTCATGAATAATGAAATCACGGCTA
GTAAAATTGATGATGGTAATAATTCAAAACCACTGTCACCTGGTTGGACGGACCAAACTGCGTATAACGCGT
TTGGAATCACTACAGGGATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATTCGATGATGAAG
ATACCCCACCAAACCCAAAAAAAGAGATCTTTAATACGACTCACTATAGGGCGAGCGCCGCC

Primers to clone ATP1β1 in pcDNA5/FRT

ATGGAGTACCCATACGACGTACCAGATTACGCTCATATGAACATGGAGGCCAGTGAATTCCACCCAAGCAGT
GGTATCAACGCAGAGTGGCCATTACGGCCGGGGTGCTCTGGATTAAATGATGAAACTTATGGCTACAAAGAG
GGCAAACCGTGCATTATTATAAAGCTCAACCGAGTTCTAGGCTTCAAACCTAAGCCTCCCAAGAATGAGTCC
TTGGAGACTTACCCAGTGATGAAGTATAACCCAAATGTCCTTCCCGTTCAGTGCACTGGCAAGCGAGATGAA
GATAAGGATAAAGTTGGAAATGTGGAGTATTTTGGACTGGGCAACTCCCCTGGTTTTCCTCTGCAGTATTAT
CCGTACTATGGCAAACTCCTGCAGCCCAAATACCTGCAGCCCCTGCTGGCCGTACAGTTCACCAATCTTACC
ATGGACACTGAAATTCGCATAGAGTGTAAGGCGTACGGTGAGAACATTGGGTACAGTGAGAAAGACCGTTTT
CAGGGACGTTTTGATGTAAAAATTGAAGTTAAGAGCTGAAAAAAAAAAAAAAGGCCGCCTCGGCCTCTAGAG
GGTGGGCATCGATACGGGATCCATCGAGCTCGAGCTGCAGATGAATCGTAGATACTGAAAAACCCCGCAAGT

TCACTTCAACTGTGCATCGTGCACCATCTCAATTTCTTTCATTTATACATCGTTTTGCCTTCTTTTATGTAA
CTATACTCCTCTAAGTTTCAATCTTGGCCATGTAACCTCTGATCTATAGAATTTTTTAAATGACTAGAATTA
ATGCCCATCTTTTTTTTTGGACCTAAATTCTTCATGAAAATATATTACGAGGGCTTATTCAGAAGCTTTGGAC
TTCTTCGCCAGAGGTTTGGTCAAGTCTCCAATCAAGGTTGTCGGCTTGTCTACCTTGCCAGAAATTTACGAA
AAGATGGAAAAGGGTCAAATCGTTGGTAGATACGTTGTTGACACTTCTAAATAAGCGAATTTCTTATGATTT
ATGATTTTTATTATTAAATAAGTTATAAAAAAAATAAGTGTATACAAATTTTAAAGTGACTCTTAGGTTTTA
AAACGAAAATTCTTATTCTTGAGTAACTCTTTCCTGTAGGTCAGGTTGCTTTCTCAGGTATAGCATGAGGTC
GCTCTTATTGACCACACCTCTACCGGCCGGTCGAAATTCCCCTACCCTATGAACATATTCCATTTTGTAATT
TCGTGTCGTTTCTATTATGAATTTCATTTATAAAGTTTATGTACAAATATCATAAAAAAAGAGAATCTTTTT
AAGCAAGGATTTTCTTAACTTCTTCGGCGACAGCATCACCGACTTCGGTGGTACTGTTGGAACCACCTAAAT
CACCAGTTCTGATACCTGCATCCAAAACCTTTTTAACTGCATCTTCAATGGCCTTACCTTCTTCAGGCAAGT
TCAATGACAATTTCAACATCATTGCAGCAGACAAGATAGTGGCGATAGGGTTGACCTTATTCTTTGGCAAAT
CTGGAGCAGAACCGTGGCATGGTTCGTACAAACCAAATGCGGTGTTCTTGTCTGGCAAAGAGGCCAAGGACG
CAGATGGCAACAAACCCAAGGAACCTGGGATAACGGAGGCTTCATCGGAGATGATATCACCAAACATGTTGC
TGGTGATTATAATACCATTTAGGTGGGTTGGGTTCTTAACTAGGATCATGGCGGCAGAATCAATCAATTGAT
GTTGAACCTTCAATGTAGGAAATTCGTTCTTGATGGTTTCCTCCACAGTTTTTCTCCATAATCTTGAAGAGG
CCAAAACATTAGCTTTATCCAAGGACCAAATAGGCAATGGTGGCTCATGTTGTAGGGCCATGAAAGCGGCCA
TTCTTGTGATTCTTTGCACTTCTGGAACGGTGTATTGTTCACTATCCCAAGCGACACCATCACCATCGTCTT
CCTTTCTCTTACCAAAGTAAATACCTCCCACTAATTCTCTGACAACAACGAAGTCAGTACCTTTAGCAAATT
GTGGCTTGATTGGAGATAAGTCTAAAAGAGAGTCGGATGCAAAGTTACATGGTCTTAAGTTGGCGTACAATT
GAAGTTCTTTACGGATTTTTAGTAAACCTTGTTCAGGTCTAACACTACCTGTACCCCATTTAGGACCACCCA
CAGCACCTAACAAAACGGCATCAGCCTTCTTGGAGGCTTCCAGCGCCTCATCTGGAAGTGGGACACCTGTAG
CTTCGATAGCAGCACCACCAATTAAATGATTTTCGAAATCGAACTTGACATTGGAACGAACATCAGAAATAG
CTTTAAGAACCTTAATGGCTTCGGCTGTGATTTCTTGACCAACGTGGTCACCTGGCAAAACGACGATCTTCT
TAGGGGCAGACATTAGAATGGTATATCCTTGAAATATATATATATATTGCTGAAATGTAAAAGGTAAGAAAA
GTTAGAAAGTAAGACGATTGCTAACCACCTATTGGAAAAAACAATAGGTCCTTAAATAATATTGTCAACTTC
AAGTATTGTGATGCAAGCATTTAGTCATGAACGCTTCTCTATTCTATATGAAAAGCCGGTTCCGGCGCTCTC
ACCTTTCCTTTTTCTCCCAATTTTTCAGTTGAAAAAGGTATATGCGTCAGGCGACCTCTGAAATTAACAAAA
AATTTCCAGTCATCGAATTTGATTCTGTGCGATAGCGCCCCTGTGTGTTCTCGTTATGTTGAGGAAAAAAAT
AATGGTTGCTAAGAGATTCGAACTCTTGCATCTTACGATACCTGAGTATTCCCACAGTTGGGGGATCTCGAC
TCTAGCTAGAGGATCAATTCGTAATCATGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTC
CACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAA
TTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGATAACTTCGTATAATGTAT
GCTATACGAAGTTATTAGGTCTGAAGAGGAGTTTACGTCCAGCCAAGCTAGCTTGGCTGCAGGTCGAGCGGC
CGCGATCCGGAACCCTTAATATAACTTCGTATAATGTATGCTATACGAAGTTATCAGCTGCATTAATGAATC
GGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGC
TCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGG
GATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTG
GCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC
CCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTG
CCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGG
TATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGC
TGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC
ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTAC
GGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGT
AGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGC
AGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCA
CGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGT
TTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCT
ATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGG
GAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCA
GCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCT
ATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCT
ACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGA
GTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAG
TTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGA
TGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCT
TGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGT
TCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCC
AACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA
AAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT
TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCG
CGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAT
AGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTC

CCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGT
GTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATAACGCA
TTTAAGCATAAACACGCACTATGCCGTTCTTCTCATGTATATATATATACAGGCAACACGCAGATATAGGTG
CGACGTGAACAGTGAGCTGTATGTGCGCAGCTCGCGTTGCATTTTCGGAAGCGCTCGTTTTCGGAAACGCTT
TGAAGTTCCTATTCCGAAGTTCCTATTCTCTAGCTAGAAAGTATAGGAACTTCAGAGCGCTTTTGAAAACCA
AAAGCGCTCTGAAGACGCACTTTCAAAAAACCAAAAACGCACCGGACTGTAACGAGCTACTAAAATATTGCG
AATACCGCTTCCACAAACATTGCTCAAAAGTATCTCTTTGCTATATATCTCTGTGCTATATCCCTATATAAC
CTACCCATCCACCTTTCGCTCCTTGAACTTGCATCTAAACTCGACCTCTACATTTTTTATGTTTATCTCTAG
TATTACTCTTTAGACAAAAAAATTGTAGTAAGAACTATTCATAGAGTGAATCGAAAACAATACGAAAATGTA
AACATTTCCTATACGTAGTATATAGAGACAAAATAGAAGAAACCGTTCATAATTTTCTGACCAATGAAGAAT
CATCAACGCTATCACTTTCTGTTCACAAAGTATGCGCAATCCACATCGGTATAGAATATAATCGGGGATGCC
TTTATCTTGAAAAAATGCACCCGCAGCTTCGCTAGTAATCAGTAAACGCGGGAAGTGGAGTCAGGCTTTTTT
TATGGAAGAGAAAATAGACACCAAAGTAGCCTTCTTCTAACCTTAACGGACCTACAGTGCAAAAAGTTATCA
AGAGACTGCATTATAGAGCGCACAAAGGAGAAAAAAAGTAATCTAAGATGCTTTGTTAGAAAAATAGCGCTC
TCGGGATGCATTTTTGTAGAACAAAAAAGAAGTATAGATTCTTTGTTGGTAAAATAGCGCTCTCGCGTTGCA
TTTCTGTTCTGTAAAAATGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTT
TACAAAAATGAAGCACAGATTCTTCGTTGGTAAAATAGCGCTTTCGCGTTGCATTTCTGTTCTGTAAAAATG
CAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTCTACAAAATGAAGCACAGAT
GCTTCGTTGCT

> pGADT7-RecAB NPC2

TGCATGCCTGCAGGTCGAGATCCGGGATCGAAGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATGAA
GGCAAAAGACAAATATAAGGGTCGAACGAAAAATAAAGTGAAAAGTGTTGATATGATGTATTTGGCTTTGCG
GCGCCGAAAAAACGAGTTTACGCAATTGCACAATCATGCTGACTCTGTGGCGGACCCGCGCTCTTGCCGGCC
CGGCGATAACGCTGGGCGTGAGGCTGTGCCCGGCGGAGTTTTTTGCGCCTGCATTTTCCAAGGTTTACCCTG
CGCTAAGGGGCGAGATTGGAGAAGCAATAAGAATGCCGGTTGGGGTTGCGATGATGACGACCACGACAACTG
GTGTCATTATTTAAGTTGCCGAAAGAACCTGAGTGCATTTGCAACATGAGTATACTAGAAGAATGAGCCAAG
ACTTGCGAGACGCGAGTTTGCCGGTGGTGCGAACAATAGAGCGACCATGACCTTGAAGGTGAGACGCGCATA
ACCGCTAGAGTACTTTGAAGAGGAAACAGCAATAGGGTTGCTACCAGTATAAATAGACAGGTACATACAACA
CTGGAAATGGTTGTCTGTTTGAGTACGCTTTCAATTCATTTGGGTGTGCACTTTATTATGTTACAATATGGA
AGGGAACTTTACACTTCTCCTATGCACATATATTAATTAAAGTCCAATGCTAGTAGAGAAGGGGGGTAACAC
CCCTCCGCGCTCTTTTCCGATTTTTTTCTAAACCGTGGAATATTTCGGATATCCTTTTGTTGTTTCCGGGTG
TACAATATGGACTTCCTCTTTTCTGGCAACCAAACCCATACATCGGGATTCCTATAATACCTTCGTTGGTCT
CCCTAACATGTAGGTGGCGGAGGGGAGATATACAATAGAACAGATACCAGACAAGACATAATGGGCTAAACA
AGACTACACCAATTACACTGCCTCATTGATGGTGGTACATAACGAACTAATACTGTAGCCCTAGACTTGATA
GCCATCATCATATCGAAGTTTCACTACCCTTTTTCCATTTGCCATCTATTGAAGTAATAATAGGCGCATGCA
ACTTCTTTTCTTTTTTTTTCTTTTCTCTCTCCCCCGTTGTTGTCTCACCATATCCGCAATGACAAAAAAATG
ATGGAAGACACTAAAGGAAAAAATTAACGACAAAGACAGCACCAACAGATGTCGTTGTTCCAGAGCTGATGA
GGGGTATCTCGAAGCACACGAAACTTTTTCCTTCCTTCATTCACGCACACTACTCTCTAATGAGCAACGGTA
TACGGCCTTCCTTCCAGTTACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTTGCTATCAAGTATAAATAG
ACCTGCAATTATTAATCTTTTGTTTCCTCGTCATTGTTCTCGTTCCCTTTCTTCCTTGTTTCTTTTTCTGCA
CAATATTTCAAGCTATACCAAGCATACAATCAACTCCAAGCTTTGCAAAGATGGATAAAGCGGAATTAATTC
CCGAGCCTCCAAAAAAGAAGAGAAAGGTCGAATTGGGTACCGCCGCCAATTTTAATCAAAGTGGGAATATTG
CTGATAGCTCATTGTCCTTCACTTTCACTAACAGTAGCAACGGTCCGAACCTCATAACAACTCAAACAAATT
CTCAAGCGCTTTCACAACCAATTGCCTCCTCTAACGTTCATGATAACTTCATGAATAATGAAATCACGGCTA
GTAAAATTGATGATGGTAATAATTCAAAACCACTGTCACCTGGTTGGACGGACCAAACTGCGTATAACGCGT
TTGGAATCACTACAGGGATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATTCGATGATGAAG
ATACCCCACCAAACCCAAAAAAAGAGATCTTTAATACGACTCACTATAGGCGAGCGCCGCC

Primer to clone NPC2 in pcDNA5/FRT

ATGGAGTACCCATACGACGTACCAGATTACGCTCATATGAACATGGAGGCCAGTGAATTCCACCCAAGCAGT
GGTATCAACGCAGAGTGGCCATTACGGCCGGGGGGGCCTGATGGTTGTAAGAGTGGAATTAACGGCCCTATC
CAAAAAGACAAGACCTATAGCTACCTGAATAAACTACCAGTGAAAAGCGAATATCCCTCTATAAAACTGGTG
GTGGAGTGGCAACTTCAGGATGACAAAAACCAAAGTCTCTTCTGCTGGGAAATCCCAGTACAGATCGTTTCT
CATCTCTAAAAAAAAAAAAAAAAAAAAGGCCGCCTCGGCCTCTAGAGGGTGGGCATCGATACGGGATCCA
TCGAGCTCGAGCTGCAGATGAATCGTAGATACTGAAAAACCCCGCAAGTTCACTTCAACTGTGCATCGTGCA
CCATCTCAATTTCTTTCATTTATACATCGTTTTGCCTTCTTTTATGTAACTATACTCCTCTAAGTTTCAATC
TTGGCCATGTAACCTCTGATCTATAGAATTTTTTAAATGACTAGAATTAATGCCCATCTTTTTTTTGGACCT

AAATTCTTCATGAAAATATATTACGAGGGCTTATTCAGAAGCTTTGGACTTCTTCGCCAGAGGTTTGGTCAA
GTCTCCAATCAAGGTTGTCGGCTTGTCTACCTTGCCAGAAATTTACGAAAAGATGGAAAAGGGTCAAATCGT
TGGTAGATACGTTGTTGACACTTCTAAATAAGCGAATTTCTTATGATTTATGATTTTTATTATTAAATAAGT
TATAAAAAAAAATAAGTGTATACATTAAAGTGACTCTTAGGTTTTAAAACGAAAATTCTTATTCTTGAGTAAC
TCTTTCCTGTAGGTCAGGTTGCTTTCTCAGGTATAGCATGAGGTCGCTCTTATTGACCACACCTCTACCGGC
CGGTCGAAATTCCCCTACCCTATGAACATATTCCATTTTGTAATTTCGTGTCGTTTCTATTATGAATTTCAT
TTATAAAGTTTATGTACAAATATCATAAAAAAAGAGAATCTTTTTAAGCAAGGATTTTCTTAACTTCTTCGG
CGACAGCATCACCGACTTCGGTGGTACTGTTGGAACCACCTAAATCACCAGTTCTGATACCTGCATCCAAAA
CCTTTTTAACTGCATCTTCAATGGCCTTACCTTCTTCAGGCAAGTTCAATGACAATTTCAACATCATTGCAG
CAGACAAGATAGTGGCGATAGGGTTGACCTTATTCTTTGGCAAATCTGGAGCAGAACCGTGGCATGGTTCGT
ACAAACCAAATGCGGTGTTCTTGTCTGGCAAAGAGGCCAAGGACGCAGATGGCAACAAACCCAAGGAACCTG
GGATAACGGAGGCTTCATCGGAGATGATATCACCAAACATGTTGCTGGTGATTATAATACCATTTAGGTGGG
TTGGGTTCTTAACTAGGATCATGGCGGCAGAATCAATCAATTGATGTTGAACCTTCAATGTAGGAAATTCGT
TCTTGATGGTTTCCTCCACAGTTTTTCTCCATAATCTTGAAGAGGCCAAAACATTAGCTTTATCCAAGGACC
AAATAGGCAATGGTGGCTCATGTTGTAGGGCCATGAAAGCGGCCATTCTTGTGATTCTTTGCACTTCTGGAA
CGGTGTATTGTTCACTATCCCAAGCGACACCATCACCATCGTCTTCCTTTCTCTTACCAAAGTAAATACCTC
CCACTAATTCTCTGACAACAACGAAGTCAGTACCTTTAGCAAATTGTGGCTTGATTGGAGATAAGTCTAAAA
GAGAGTCGGATGCAAAGTTACATGGTCTTAAGTTGGCGTACAATTGAAGTTCTTTACGGATTTTTAGTAAAC
CTTGTTCAGGTCTAACACTACCTGTACCCCATTTAGGACCACCCACAGCACCTAACAAAACGGCATCAGCCT
TCTTGGAGGCTTCCAGCGCCTCATCTGGAAGTGGGACACCTGTAGCTTCGATAGCAGCACCACCAATTAAAT
GATTTTCGAAATCGAACTTGACATTGGAACGAACATCAGAAATAGCTTTAAGAACCTTAATGGCTTCGGCTG
TGATTTCTTGACCAACGTGGTCACCTGGCAAAACGACGATCTTCTTAGGGGCAGACATTAGAATGGTATATC
CTTGAAATATATATATATATTGCTGAAATGTAAAAGGTAAGAAAAGTTAGAAAGTAAGACGATTGCTAACCA
CCTATTGGAAAAAACAATAGGTCCTTAAATAATATTGTCAACTTCAAGTATTGTGATGCAAGCATTTAGTCA
TGAACGCTTCTCTATTCTATATGAAAAGCCGGTTCCGGCGCTCTCACCTTTCCTTTTTCTCCCAATTTTTCA
GTTGAAAAAGGTATATGCGTCAGGCGACCTCTGAAATTAACAAAAAATTTCCAGTCATCGAATTTGATTCTG
TGCGATAGCGCCCCTGTGTGTTCTCGTTATGTTGAGGAAAAAAATAATGGTTGCTAAGAGATTCGAACTCTT
GCATCTTACGATACCTGAGTATTCCCACAGTTGGGGGATCTCGACTCTAGCTAGAGGATCAATTCGTAATCA
TGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAA
AGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCC
AGTCGGGAAACCTGTCGTGCCAGCTGATAACTTCGTATAATGTATGCTATACGAAGTTATTAGGTCTGAAGA
GGAGTTTACGTCCAGCCAAGCTAGCTTGGCTGCAGGTCGAGCGGCCGCGATCCGGAACCCTTAATATAACTT
CGTATAATGTATGCTATACGAAGTTATCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTG
CGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTA
TCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCA
AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCT
GACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG
TTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTT
CTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGC
TCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTT
GAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGG
TATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGT
ATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCT
TTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTA
TCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAG
TAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCA
TCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCT
GCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCC
GAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTA
AGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCG
TTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAA
AAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTT
ATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCA
ACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACC
GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATC
TTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTC
ACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAA
TGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGA
TACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCT

GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTC
GCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAA
GCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAAC
TATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATAACGCATTTAAGCATAAACACGCACTATGCCGT
TCTTCTCATGTATATATATATACAGGCAACACGCAGATATAGGTGCGACGTGAACAGTGAGCTGTATGTGCG
CAGCTCGCGTTGCATTTTCGGAAGCGCTCGTTTTCGGAAACGCTTTGAAGTTCCTATTCCGAAGTTCCTATT
CTCTAGCTAGAAAGTATAGGAACTTCAGAGCGCTTTTGAAAACCAAAAGCGCTCTGAAGACGCACTTTCAAA
AAACCAAAAACGCACCGGACTGTAACGAGCTACTAAAATATTGCGAATACCGCTTCCACAAACATTGCTCAA
AAGTATCTCTTTGCTATATATCTCTGTGCTATATCCCTATATAACCTACCCATCCACCTTTCGCTCCTTGAA
CTTGCATCTAAACTCGACCTCTACATTTTTTATGTTTATCTCTAGTATTACTCTTTAGACAAAAAAATTGTA
GTAAGAACTATTCATAGAGTGAATCGAAAACAATACGAAAATGTAAACATTTCCTATACGTAGTATATAGAG
ACAAAATAGAAGAAACCGTTCATAATTTTCTGACCAATGAAGAATCATCAACGCTATCACTTTCTGTTCACA
AAGTATGCGCAATCCACATCGGTATAGAATATAATCGGGGATGCCTTTATCTTGAAAAAATGCACCCGCAGC
TTCGCTAGTAATCAGTAAACGCGGGAAGTGGAGTCAGGCTTTTTTTATGGAAGAGAAAATAGACACCAAAGT
AGCCTTCTTCTAACCTTAACGGACCTACAGTGCAAAAAGTTATCAAGAGACTGCATTATAGAGCGCACAAAG
GAGAAAAAAAAGTAATCTAAGATGCTTTGTTAGAAAAATAGCGCTCTCGGGATGCATTTTTGTAGAACAAAAA
AGAAGTATAGATTCTTTGTTGGTAAAATAGCGCTCTCGCGTTGCATTTCTGTTCTGTAAAAATGCAGCTCAG
ATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTTTACAAAAATGAAGCACAGATTCTTCGT
TGGTAAAATAGCGCTTTCGCGTTGCATTTCTGTTCTGTAAAAATGCAGCTCAGATTCTTTGTTTGAAAAATT
AGCGCTCTCGCGTTGCATTTTTGTTCTACAAAATGAAGCACAGATGCTTCGTTGCT

> pGADT7-RecAB PCNA

TGCATGCCTGCAGGTCGAGATCCGGGATCGAAGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATGAA
GGCAAAAGACAAATATAAGGGTCGAACGAAAAATAAAGTGAAAAGTGTTGATATGATGTATTTGGCTTTGCG
GCGCCGAAAAAACGAGTTTACGCAATTGCACAATCATGCTGACTCTGTGGCGGACCCGCGCTCTTGCCGGCC
CGGCGATAACGCTGGGCGTGAGGCTGTGCCCGGCGGAGTTTTTTGCGCCTGCATTTTCCAAGGTTTACCCTG
CGCTAAGGGGCGAGATTGGAGAAGCAATAAGAATGCCGGTTGGGGTTGCGATGATGACGACCACGACAACTG
GTGTCATTATTTAAGTTGCCGAAAGAACCTGAGTGCATTTGCAACATGAGTATACTAGAAGAATGAGCCAAG
ACTTGCGAGACGCGAGTTTGCCGGTGGTGCGAACAATAGAGCGACCATGACCTTGAAGGTGAGACGCGCATA
ACCGCTAGAGTACTTTGAAGAGGAAACAGCAATAGGGTTGCTACCAGTATAAATAGACAGGTACATACAACA
CTGGAAATGGTTGTCTGTTTGAGTACGCTTTCAATTCATTTGGGTGTGCACTTTATTATGTTACAATATGGA
AGGGAACTTTACACTTCTCCTATGCACATATATTAATTAAAGTCCAATGCTAGTAGAGAAGGGGGGTAACAC
CCCTCCGCGCTCTTTTCCGATTTTTTTCTAAACCGTGGAATATTTCGGATATCCTTTTGTTGTTTCCGGGTG
TACAATATGGACTTCCTCTTTTCTGGCAACCAAACCCATACATCGGGATTCCTATAATACCTTCGTTGGTCT
CCCTAACATGTAGGTGGCGGAGGGGAGATATACAATAGAACAGATACCAGACAAGACATAATGGGCTAAACA
AGACTACACCAATTACACTGCCTCATTGATGGTGGTACATAACGAACTAATACTGTAGCCCTAGACTTGATA
GCCATCATCATATCGAAGTTTCACTACCCTTTTTCCATTTGCCATCTATTGAAGTAATAATAGGCGCATGCA
ACTTCTTTTCTTTTTTTTTCTTTTCTCTCTCCCCCGTTGTTGTCTCACCATATCCGCAATGACAAAAAAATG
ATGGAAGACACTAAAGGAAAAAATTAACGACAAAGACAGCACCAACAGATGTCGTTGTTCCAGAGCTGATGA
GGGGTATCTCGAAGCACACGAAACTTTTTCCTTCCTTCATTCACGCACACTACTCTCTAATGAGCAACGGTA
TACGGCCTTCCTTCCAGTTACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTTGCTATCAAGTATAAATAG
ACCTGCAATTATTAATCTTTTGTTTCCTCGTCATTGTTCTCGTTCCCTTTCTTCCTTGTTTCTTTTTCTGCA
CAATATTTCAAGCTATACCAAGCATACAATCAACTCCAAGCTTTGCAAAGATGGATAAAGCGGAATTAATTC
CCGAGCCTCCAAAAAAGAAGAGAAAGGTCGAATTGGGTACCGCCGCCAATTTTAATCAAAGTGGGAATATTG
CTGATAGCTCATTGTCCTTCACTTTCACTAACAGTAGCAACGGTCCGAACCTCATAACAACTCAAACAAATT
CTCAAGCGCTTTCACAACCAATTGCCTCCTCTAACGTTCATGATAACTTCATGAATAATGAAATCACGGCTA
GTAAAATTGATGATGGTAATAATTCAAAACCACTGTCACCTGGTTGGACGGACCAAACTGCGTATAACGCGT
TTGGAATCACTACAGGGATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATTCGATGATGAAG
ATACCCCACCAAACCCAAAAAAAGAGATCTTTAATACGACTCACTATAGGGCGAGCGCCGCC

Primers to clone PCNA in pcDNA5/FRT

ATGGAGTACCCATACGACGTACCAGATTACGCTCATATGAACATGGAGGCCAGTGAATTCCACCCAAGCAGT
GGTATCAACGCAGAGTGGCCATTACGGCCGGGGGATCTCAGCCATATTGGAGATGCTGTTGTAATTTCCTGT
GCAAAAGACGGAGTGAAATTTTCTGCAAGTGGAGAACTTGGAAATGGAAACATTAAATTGTCACAGACAAGT
AATGTCGATAAAGAGGAGGAAGCTGTTACCATAGAGATGAATGAACCAGTTCAACTAACTTTTGCACTGAGG
TACCTGAACTTCTTTACAAAAGCCACTCCACTCTCTTCAACGGTGACACTCAGTATGTCTGCAGATGTACCC
CTTGTTGTAGAGTATAAAATTGCGGATATGGGACACTTAAAATACTACTTGGCTCCCAAGATCGAGGATGAA

GAAGGATCTTAGAAAAAAAAAAAAAAGGCCGCCTCGGCCTCTAGAGGGTGGGCATCGATACGGGATCCATCG
AGCTCGAGCTGCAGATGAATCGTAGATACTGAAAAACCCCGCAAGTTCACTTCAACTGTGCATCGTGCACCA
TCTCAATTTCTTTCATTTATACATCGTTTTGCCTTCTTTTATGTAACTATACTCCTCTAAGTTTCAATCTTG
GCCATGTAACCTCTGATCTATAGAATTTTTTAAATGACTAGAATTAATGCCCATCTTTTTTTTGGACCTAAA
TTCTTCATGAAAATATATTACGAGGGCTTATTCAGAAGCTTTGGACTTCTTCGCCAGAGGTTTGGTCAAGTC
TCCAATCAAGGTTGTCGGCTTGTCTACCTTGCCAGAAATTTACGAAAAGATGGAAAAGGGTCAAATCGTTGG
TAGATACGTTGTTGACACTTCTAAATAAGCGAATTTCTTATGATTTATGATTTTTATTATTAAATAAGTTAT
AAAAAAAATAAGTGTATACAAATTTTAAAGTGACTCTTAGGTTTTAAAACGAAAATTCTTATTCTTGAGTAA
CTCTTTCCTGTAGGTCAGGTTGCTTTCTCAGGTATAGCATGAGGTCGCTCTTATTGACCACACCTCTACCGG
CCGGTCGAAATTCCCCTACCCTATGAACATATTCCATTTTGTAATTTCGTGTCGTTTCTATTATGAATTTCA
TTTATAAAGTTTATGTACAAATATCATAAAAAAAGAGAATCTTTTTAAGCAAGGATTTTCTTAACTTCTTCG
GCGACAGCATCACCGACTTCGGTGGTACTGTTGGAACCACCTAAATCACCAGTTCTGATACCTGCATCCAAA
ACCTTTTTAACTGCATCTTCAATGGCCTTACCTTCTTCAGGCAAGTTCAATGACAATTTCAACATCATTGCA
GCAGACAAGATAGTGGCGATAGGGTTGACCTTATTCTTTGGCAAATCTGGAGCAGAACCGTGGCATGGTTCG
TACAAACCAAATGCGGTGTTCTTGTCTGGCAAAGAGGCCAAGGACGCAGATGGCAACAAACCCAAGGAACCT
GGGATAACGGAGGCTTCATCGGAGATGATATCACCAAACATGTTGCTGGTGATTATAATACCATTTAGGTGG
GTTGGGTTCTTAACTAGGATCATGGCGGCAGAATCAATCAATTGATGTTGAACCTTCAATGTAGGAAATTCG
TTCTTGATGGTTTCCTCCACAGTTTTTCTCCATAATCTTGAAGAGGCCAAAACATTAGCTTTATCCAAGGAC
CAAATAGGCAATGGTGGCTCATGTTGTAGGGCCATGAAAGCGGCCATTCTTGTGATTCTTTGCACTTCTGGA
ACGGTGTATTGTTCACTATCCCAAGCGACACCATCACCATCGTCTTCCTTTCTCTTACCAAAGTAAATACCT
CCCACTAATTCTCTGACAACAACGAAGTCAGTACCTTTAGCAAATTGTGGCTTGATTGGAGATAAGTCTAAA
AGAGAGTCGGATGCAAAGTTACATGGTCTTAAGTTGGCGTACAATTGAAGTTCTTTACGGATTTTTAGTAAA
CCTTGTTCAGGTCTAACACTACCTGTACCCCATTTAGGACCACCCACAGCACCTAACAAAACGGCATCAGCC
TTCTTGGAGGCTTCCAGCGCCTCATCTGGAAGTGGGACACCTGTAGCTTCGATAGCAGCACCACCAATTAAA
TGATTTTCGAAATCGAACTTGACATTGGAACGAACATCAGAAATAGCTTTAAGAACCTTAATGGCTTCGGCT
GTGATTTCTTGACCAACGTGGTCACCTGGCAAAACGACGATCTTCTTAGGGGCAGACATTAGAATGGTATAT
CCTTGAAATATATATATATATTGCTGAAATGTAAAAGGTAAGAAAAGTTAGAAAGTAAGACGATTGCTAACC
ACCTATTGGAAAAAACAATAGGTCCTTAAATAATATTGTCAACTTCAAGTATTGTGATGCAAGCATTTAGTC
ATGAACGCTTCTCTATTCTATATGAAAAGCCGGTTCCGGCGCTCTCACCTTTCCTTTTTCTCCCAATTTTTC
AGTTGAAAAAGGTATATGCGTCAGGCGACCTCTGAAATTAACAAAAAATTTCCAGTCATCGAATTTGATTCT
GTGCGATAGCGCCCCTGTGTGTTCTCGTTATGTTGAGGAAAAAAATAATGGTTGCTAAGAGATTCGAACTCT
TGCATCTTACGATACCTGAGTATTCCCACAGTTGGGGGATCTCGACTCTAGCTAGAGGATCAATTCGTAATC
ATGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATA
AAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTC
CAGTCGGGAAACCTGTCGTGCCAGCTGATAACTTCGTATAATGTATGCTATACGAAGTTATTAGGTCTGAAG
AGGAGTTTACGTCCAGCCAAGCTAGCTTGGCTGCAGGTCGAGCGGCCGCGATCCGGAACCCTTAATATAACT
TCGTATAATGTATGCTATACGAAGTTATCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTT
GCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGT
ATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGC
AAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC
TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGC
GTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTT
TCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCG
CTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCT
TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG
GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGG
TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC
CGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCC
TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATT
ATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGA
GTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTC
ATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC
TGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC
CGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGT
AAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTC
GTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAA
AAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGT
TATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTC
AACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATAC
CGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGAT

CTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTT
CACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAA
ATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGG
ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACC
TGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCT
CGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTA
AGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAA
CTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATAACGCATTTAAGCATAAACACGCACTATGCCG
TTCTTCTCATGTATATATATATACAGGCAACACGCAGATATAGGTGCGACGTGAACAGTGAGCTGTATGTGC
GCAGCTCGCGTTGCATTTTCGGAAGCGCTCGTTTTCGGAAACGCTTTGAAGTTCCTATTCCGAAGTTCCTAT
TCTCTAGCTAGAAAGTATAGGAACTTCAGAGCGCTTTTGAAAACCAAAAGCGCTCTGAAGACGCACTTTCAA
AAAACCAAAAACGCACCGGACTGTAACGAGCTACTAAAATATTGCGAATACCGCTTCCACAAACATTGCTCA
AAAGTATCTCTTTGCTATATATCTCTGTGCTATATCCCTATATAACCTACCCATCCACCTTTCGCTCCTTGA
ACTTGCATCTAAACTCGACCTCTACATTTTTTATGTTTATCTCTAGTATTACTCTTTAGACAAAAAAATTGT
AGTAAGAACTATTCATAGAGTGAATCGAAAACAATACGAAAATGTAAACATTTCCTATACGTAGTATATAGA
GACAAAATAGAAGAAACCGTTCATAATTTTCTGACCAATGAAGAATCATCAACGCTATCACTTTCTGTTCAC
AAAGTATGCGCAATCCACATCGGTATAGAATATAATCGGGGATGCCTTTATCTTGAAAAAATGCACCCGCAG
CTTCGCTAGTAATCAGTAAACGCGGGAAGTGGAGTCAGGCTTTTTTTATGGAAGAGAAAATAGACACCAAAG
TAGCCTTCTTCTAACCTTAACGGACCTACAGTGCAAAAAGTTATCAAGAGACTGCATTATAGAGCGCACAAA
GGAGAAAAAAAAGTAATCTAAGATGCTTTGTTAGAAAAATAGCGCTCTCGGGATGCATTTTTGTAGAACAAAA
AAGAAGTATAGATTCTTTGTTGGTAAAATAGCGCTCTCGCGTTGCATTTCTGTTCTGTAAAAATGCAGCTCA
GATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTTTACAAAAATGAAGCACAGATTCTTCG
TTGGTAAAATAGCGCTTTCGCGTTGCATTTCTGTTCTGTAAAAATGCAGCTCAGATTCTTTGTTTGAAAAAT
TAGCGCTCTCGCGTTGCATTTTTGTTCTACAAAATGAAGCACAGATGCTTCGTTGCT

> pGADT7-RecAB PGM1

TGCATGCCTGCAGGTCGAGATCCGGGATCGAAGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATGAA
GGCAAAAGACAAATATAAGGGTCGAACGAAAAATAAAGTGAAAAGTGTTGATATGATGTATTTGGCTTTGCG
GCGCCGAAAAAACGAGTTTACGCAATTGCACAATCATGCTGACTCTGTGGCGGACCCGCGCTCTTGCCGGCC
CGGCGATAACGCTGGGCGTGAGGCTGTGCCCGGCGGAGTTTTTTGCGCCTGCATTTTCCAAGGTTTACCCTG
CGCTAAGGGGCGAGATTGGAGAAGCAATAAGAATGCCGGTTGGGGTTGCGATGATGACGACCACGACAACTG
GTGTCATTATTTAAGTTGCCGAAAGAACCTGAGTGCATTTGCAACATGAGTATACTAGAAGAATGAGCCAAG
ACTTGCGAGACGCGAGTTTGCCGGTGGTGCGAACAATAGAGCGACCATGACCTTGAAGGTGAGACGCGCATA
ACCGCTAGAGTACTTTGAAGAGGAAACAGCAATAGGGTTGCTACCAGTATAAATAGACAGGTACATACAACA
CTGGAAATGGTTGTCTGTTTGAGTACGCTTTCAATTCATTTGGGTGTGCACTTTATTATGTTACAATATGGA
AGGGAACTTTACACTTCTCCTATGCACATATATTAATTAAAGTCCAATGCTAGTAGAGAAGGGGGGTAACAC
CCCTCCGCGCTCTTTTCCGATTTTTTTCTAAACCGTGGAATATTTCGGATATCCTTTTGTTGTTTCCGGGTG
TACAATATGGACTTCCTCTTTTCTGGCAACCAAACCCATACATCGGGATTCCTATAATACCTTCGTTGGTCT
CCCTAACATGTAGGTGGCGGAGGGGAGATATACAATAGAACAGATACCAGACAAGACATAATGGGCTAAACA
AGACTACACCAATTACACTGCCTCATTGATGGTGGTACATAACGAACTAATACTGTAGCCCTAGACTTGATA
GCCATCATCATATCGAAGTTTCACTACCCTTTTTCCATTTGCCATCTATTGAAGTAATAATAGGCGCATGCA
ACTTCTTTTCTTTTTTTTTCTTTTCTCTCTCCCCCGTTGTTGTCTCACCATATCCGCAATGACAAAAAAATG
ATGGAAGACACTAAAGGAAAAAATTAACGACAAAGACAGCACCAACAGATGTCGTTGTTCCAGAGCTGATGA
GGGGTATCTCGAAGCACACGAAACTTTTTCCTTCCTTCATTCACGCACACTACTCTCTAATGAGCAACGGTA
TACGGCCTTCCTTCCAGTTACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTTGCTATCAAGTATAAATAG
ACCTGCAATTATTAATCTTTTGTTTCCTCGTCATTGTTCTCGTTCCCTTTCTTCCTTGTTTCTTTTTCTGCA
CAATATTTCAAGCTATACCAAGCATACAATCAACTCCAAGCTTTGCAAAGATGGATAAAGCGGAATTAATTC
CCGAGCCTCCAAAAAAGAAGAGAAAGGTCGAATTGGGTACCGCCGCCAATTTTAATCAAAGTGGGAATATTG
CTGATAGCTCATTGTCCTTCACTTTCACTAACAGTAGCAACGGTCCGAACCTCATAACAACTCAAACAAATT
CTCAAGCGCTTTCACAACCAATTGCCTCCTCTAACGTTCATGATAACTTCATGAATAATGAAATCACGGCTA
GTAAAATTGATGATGGTAATAATTCAAAACCACTGTCACCTGGTTGGACGGACCAAACTGCGTATAACGCGT
TTGGAATCACTACAGGGATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATTCGATGATGAAG
ATACCCCACCAAACCCAAAAAAAGAGATCTTTAATACGACTCACTATAGGGCGAGCGCCGCC

Primers to clone PGM1 in pcDNA5/FRT

ATGGAGTACCCATACGACGTACCAGATTACGCTCATATGAACATGGAGGCCAGTGAATTCCACCCAAGCAGT
GGTATCAACGCAGAGTGGCCATTACGGCCGGGGGTGGAGAAGGCCGATAACTTTGAATACAGCGACCCAGTG
GATGGAAGCATTTCAAGAAATCAGGGCTTGCGCCTCATTTTCACAGATGGTTCTCGAATCGTCTTCCGACTG
AGCGGCACTGGGAGTGCCGGGGCCACCATTCGGCTGTACATCGATAGCTATGAGAAGGACGTTGCCAAGATT
AACCAGGACCCCCAGGTCATGTTGGCCCCCCTTATTTCCATTGCTCTGAAAGTGTCCCAGCTGCAGGAGAGG
ACGGGACGCACTGCACCCACTGTCATCACCTAAAAAAAAAAAAAAAAAAAGGCCGCCTCGGCCTCTAGAGGG
TGGGCATCGATACGGGATCCATCGAGCTCGAGCTGCAGATGAATCGTAGATACTGAAAAACCCCGCAAGTTC
ACTTCAACTGTGCATCGTGCACCATCTCAATTTCTTTCATTTATACATCGTTTTGCCTTCTTTTATGTAACT
ATACTCCTCTAAGTTTCAATCTTGGCCATGTAACCTCTGATCTATAGAATTTTTTAAATGACTAGAATTAAT
GCCCATCTTTTTTTTGGACCTAAATTCTTCATGAAAATATATTACGAGGGCTTATTCAGAAGCTTTGGACTT
CTTCGCCAGAGGTTTGGTCAAGTCTCCAATCAAGGTTGTCGGCTTGTCTACCTTGCCAGAAATTTACGAAAA
GATGGAAAAGGGTCAAATCGTTGGTAGATACGTTGTTGACACTTCTAAATAAGCGAATTTCTTATGATTTAT
GATTTTTATTATTAAATAAGTTATAAAAAAAAATAAGTGTATACAAATTTTAAAGTGACTCTTAGGTTTTAAA
ACGAAAATTCTTATTCTTGAGTAACTCTTTCCTGTAGGTCAGGTTGCTTTCTCAGGTATAGCATGAGGTCGC
TCTTATTGACCACACCTCTACCGGCCGGTCGAAATTCCCCTACCCTATGAACATATTCCATTTTGTAATTTC
GTGTCGTTTCTATTATGAATTTCATTTATAAAGTTTATGTACAAATATCATAAAAAAAGAGAATCTTTTTAA
GCAAGGATTTTCTTAACTTCTTCGGCGACAGCATCACCGACTTCGGTGGTACTGTTGGAACCACCTAAATCA
CCAGTTCTGATACCTGCATCCAAAACCTTTTTAACTGCATCTTCAATGGCCTTACCTTCTTCAGGCAAGTTC
AATGACAATTTCAACATCATTGCAGCAGACAAGATAGTGGCGATAGGGTTGACCTTATTCTTTGGCAAATCT
GGAGCAGAACCGTGGCATGGTTCGTACAAACCAAATGCGGTGTTCTTGTCTGGCAAAGAGGCCAAGGACGCA
GATGGCAACAAACCCAAGGAACCTGGGATAACGGAGGCTTCATCGGAGATGATATCACCAAACATGTTGCTG
GTGATTATAATACCATTTAGGTGGGTTGGGTTCTTAACTAGGATCATGGCGGCAGAATCAATCAATTGATGT
TGAACCTTCAATGTAGGAAATTCGTTCTTGATGGTTTCCTCCACAGTTTTTCTCCATAATCTTGAAGAGGCC
AAAACATTAGCTTTATCCAAGGACCAAATAGGCAATGGTGGCTCATGTTGTAGGGCCATGAAAGCGGCCATT
CTTGTGATTCTTTGCACTTCTGGAACGGTGTATTGTTCACTATCCCAAGCGACACCATCACCATCGTCTTCC
TTTCTCTTACCAAAGTAAATACCTCCCACTAATTCTCTGACAACAACGAAGTCAGTACCTTTAGCAAATTGT
GGCTTGATTGGAGATAAGTCTAAAAGAGAGTCGGATGCAAAGTTACATGGTCTTAAGTTGGCGTACAATTGA
AGTTCTTTACGGATTTTTAGTAAACCTTGTTCAGGTCTAACACTACCTGTACCCCATTTAGGACCACCCACA
GCACCTAACAAAACGGCATCAGCCTTCTTGGAGGCTTCCAGCGCCTCATCTGGAAGTGGGACACCTGTAGCT
TCGATAGCAGCACCACCAATTAAATGATTTTCGAAATCGAACTTGACATTGGAACGAACATCAGAAATAGCT
TTAAGAACCTTAATGGCTTCGGCTGTGATTTCTTGACCAACGTGGTCACCTGGCAAAACGACGATCTTCTTA
GGGGCAGACATTAGAATGGTATATCCTTGAAATATATATATATATTGCTGAAATGTAAAAGGTAAGAAAAGT
TAGAAAGTAAGACGATTGCTAACCACCTATTGGAAAAAACAATAGGTCCTTAAATAATATTGTCAACTTCAA
GTATTGTGATGCAAGCATTTAGTCATGAACGCTTCTCTATTCTATATGAAAAGCCGGTTCCGGCGCTCTCAC
CTTTCCTTTTTCTCCCAATTTTTCAGTTGAAAAAGGTATATGCGTCAGGCGACCTCTGAAATTAACAAAAAA
TTTCCAGTCATCGAATTTGATTCTGTGCGATAGCGCCCCTGTGTGTTCTCGTTATGTTGAGGAAAAAAATAA
TGGTTGCTAAGAGATTCGAACTCTTGCATCTTACGATACCTGAGTATTCCCACAGTTGGGGGATCTCGACTC
TAGCTAGAGGATCAATTCGTAATCATGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA
CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATT
GCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGATAACTTCGTATAATGTATGC
TATACGAAGTTATTAGGTCTGAAGAGGAGTTTACGTCCAGCCAAGCTAGCTTGGCTGCAGGTCGAGCGGCCG
CGATCCGGAACCCTTAATATAACTTCGTATAATGTATGCTATACGAAGTTATCAGCTGCATTAATGAATCGG
CCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTC
GGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGA
TAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGC
GTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCC
GACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC
GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTA
TCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG
CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCAC
TGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGG
CTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG
CTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAG
AAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACG
TTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTT
TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTAT
CTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGA
GGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC

AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTAT
TAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTAC
AGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGT
TACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTT
GGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATG
CTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG
CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTC
TTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAA
CTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAA
AAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTA
TCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCG
CACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAG
GCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCC
GGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGT
TGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATAACGCATT
TAAGCATAAACACGCACTATGCCGTTCTTCTCATGTATATATATATACAGGCAACACGCAGATATAGGTGCG
ACGTGAACAGTGAGCTGTATGTGCGCAGCTCGCGTTGCATTTTCGGAAGCGCTCGTTTTCGGAAACGCTTTG
AAGTTCCTATTCCGAAGTTCCTATTCTCTAGCTAGAAAGTATAGGAACTTCAGAGCGCTTTTGAAAACCAAA
AGCGCTCTGAAGACGCACTTTCAAAAAACCAAAAACGCACCGGACTGTAACGAGCTACTAAAATATTGCGAA
TACCGCTTCCACAAACATTGCTCAAAAGTATCTCTTTGCTATATATCTCTGTGCTATATCCCTATATAACCT
ACCCATCCACCTTTCGCTCCTTGAACTTGCATCTAAACTCGACCTCTACATTTTTTATGTTTATCTCTAGTA
TTACTCTTTAGACAAAAAAATTGTAGTAAGAACTATTCATAGAGTGAATCGAAAACAATACGAAAATGTAAA
CATTTCCTATACGTAGTATATAGAGACAAAATAGAAGAAACCGTTCATAATTTTCTGACCAATGAAGAATCA
TCAACGCTATCACTTTCTGTTCACAAAGTATGCGCAATCCACATCGGTATAGAATATAATCGGGGATGCCTT
TATCTTGAAAAAATGCACCCGCAGCTTCGCTAGTAATCAGTAAACGCGGGAAGTGGAGTCAGGCTTTTTTTA
TGGAAGAGAAAATAGACACCAAAGTAGCCTTCTTCTAACCTTAACGGACCTACAGTGCAAAAAGTTATCAAG
AGACTGCATTATAGAGCGCACAAAGGAGAAAAAAAGTAATCTAAGATGCTTTGTTAGAAAAATAGCGCTCTC
GGGATGCATTTTTGTAGAACAAAAAAGAAGTATAGATTCTTTGTTGGTAAAATAGCGCTCTCGCGTTGCATT
TCTGTTCTGTAAAAATGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTTTA
CAAAAATGAAGCACAGATTCTTCGTTGGTAAAATAGCGCTTTCGCGTTGCATTTCTGTTCTGTAAAAATGCA
GCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTCTACAAAATGAAGCACAGATGC
TTCGTTGCT

> pGADT7-RecAB RNF-2

TGCATGCCTGCAGGTCGAGATCCGGGATCGAAGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATGAA
GGCAAAAGACAAATATAAGGGTCGAACGAAAAATAAAGTGAAAAGTGTTGATATGATGTATTTGGCTTTGCG
GCGCCGAAAAAACGAGTTTACGCAATTGCACAATCATGCTGACTCTGTGGCGGACCCGCGCTCTTGCCGGCC
CGGCGATAACGCTGGGCGTGAGGCTGTGCCCGGCGGAGTTTTTTGCGCCTGCATTTTCCAAGGTTTACCCTG
CGCTAAGGGGCGAGATTGGAGAAGCAATAAGAATGCCGGTTGGGGTTGCGATGATGACGACCACGACAACTG
GTGTCATTATTTAAGTTGCCGAAAGAACCTGAGTGCATTTGCAACATGAGTATACTAGAAGAATGAGCCAAG
ACTTGCGAGACGCGAGTTTGCCGGTGGTGCGAACAATAGAGCGACCATGACCTTGAAGGTGAGACGCGCATA
ACCGCTAGAGTACTTTGAAGAGGAAACAGCAATAGGGTTGCTACCAGTATAAATAGACAGGTACATACAACA
CTGGAAATGGTTGTCTGTTTGAGTACGCTTTCAATTCATTTGGGTGTGCACTTTATTATGTTACAATATGGA
AGGGAACTTTACACTTCTCCTATGCACATATATTAATTAAAGTCCAATGCTAGTAGAGAAGGGGGGTAACAC
CCCTCCGCGCTCTTTTCCGATTTTTTTCTAAACCGTGGAATATTTCGGATATCCTTTTGTTGTTTCCGGGTG
TACAATATGGACTTCCTCTTTTCTGGCAACCAAACCCATACATCGGGATTCCTATAATACCTTCGTTGGTCT
CCCTAACATGTAGGTGGCGGAGGGGAGATATACAATAGAACAGATACCAGACAAGACATAATGGGCTAAACA
AGACTACACCAATTACACTGCCTCATTGATGGTGGTACATAACGAACTAATACTGTAGCCCTAGACTTGATA
GCCATCATCATATCGAAGTTTCACTACCCTTTTTCCATTTGCCATCTATTGAAGTAATAATAGGCGCATGCA
ACTTCTTTTCTTTTTTTTTCTTTTCTCTCTCCCCCGTTGTTGTCTCACCATATCCGCAATGACAAAAAAATG
ATGGAAGACACTAAAGGAAAAAATTAACGACAAAGACAGCACCAACAGATGTCGTTGTTCCAGAGCTGATGA
GGGGTATCTCGAAGCACACGAAACTTTTTCCTTCCTTCATTCACGCACACTACTCTCTAATGAGCAACGGTA
TACGGCCTTCCTTCCAGTTACTTGAATTTGAAATAAAAAAAAAGTTTGCTGTCTTGCTATCAAGTATAAATAG
ACCTGCAATTATTAATCTTTTGTTTCCTCGTCATTGTTCTCGTTCCCTTTCTTCCTTGTTTCTTTTTCTGCA
CAATATTTCAAGCTATACCAAGCATACAATCAACTCCAAGCTTTGCAAAGATGGATAAAGCGGAATTAATTC
CCGAGCCTCCAAAAAAGAAGAGAAAGGTCGAATTGGGTACCGCCGCCAATTTTAATCAAAGTGGGAATATTG
CTGATAGCTCATTGTCCTTCACTTTCACTAACAGTAGCAACGGTCCGAACCTCATAACAACTCAAACAAATT
CTCAAGCGCTTTCACAACCAATTGCCTCCTCTAACGTTCATGATAACTTCATGAATAATGAAATCACGGCTA

GTAAAATTGATGATGGTAATAATTCAAAACCACTGTCACCTGGTTGGACGGACCAAACTGCGTATAACGCGT
TTGGAATCACTACAGGGATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATTCGATGATGAAG
ATACCCCACCAAACCCAAAAAAAGAGATCTTTAATACGACTCACTATAGGGCGAGCGCCGCC

Primers to clone RNF2 in pcDNA5/FRT

ATGGAGTACCCATACGACGTACCAGATTACGCTCATATGAACATGGAGGCCAGTGAATTCCACCCAAGCAGT
GGTATCAACGCAGAGTGGCCATTACGGCCGGGGCAACCATTAAGCAAAACATGGGAACTCAGTTTATATGAG
TTACAACGAACACCTCAGGAGGCAATAACAGATGGCTTGGAAATTGTGGTTTCACCTCGAAGTCTACACAGT
GAATTAATGTGCCCAATTTGTTTGGATATGTTGAAGAACACCATGACTACAAAGGAGTGTTTACATCGTTTT
TGTGCAGACTGCATCATCACAGCCCTTAGAAGTGGCAACAAAGAATGTCCTACCTGTCGGAAAAAACTAGTT
TCCAAAAGATCACTAAGGCCAGACCCAAACTTTGATGCACTCATCAGCAAAATTTATCCAAGTCGTGATGAG
TATGAAGCTCATCAAGAGAGAGTATTAGCCAGGATCAACAAGCACAATAATCAGCAAGCACTCAGTCACAGC
ATTGAGGAAGGACTGAAGATACAGGCCATGAACAGACTGCAGCGAGGCAAGAAACAACAGATTGAAAATGGT
AGTGGAGCAGAAGATAATGGTGACAGTTCACACTGCAGTAATGCATCCACACATAGCAATCAGGAAGCAGGC
CCTAGTAACAAACGGACCAAAACATCTGATGATTCTGGGCTAGAGCTTGATAATAACAATGCAGCAATGGCA
ATTGATCCAGTAATGGATGGTGCTAGTGAAATTGAATTAGTATTCAGGCCTCATCCCACACTTATGGAAAAA
GATGACAGTGCACAGACGANATACATAAAGACTTCTGGTAACGCCACTGTTGATCACTTATCCAAGTATCTG
GCTGTGAGGTTAGCTTTAGAAGAACTTCGAAGCAAAGGTGAATCAAACCAGATGAACCTTGATACAGCCAGT
GAGAAGCAGTATACCATTTATATAGCAACAGCCGGTGGCCAGTTCACTGTATTAAATGGNTCTTTCTTTGGA
ATTGGTCAGTGAAAAAAAAAAAAAAAAAAGGCCGCCTCGGCCTCTAGAGGGTGGGCATCGATACGGGATCC
ATCGAGCTCGAGCTGCAGATGAATCGTAGATACTGAAAAACCCCGCAAGTTCACTTCAACTGTGCATCGTGC
ACCATCTCAATTTCTTTCATTTATACATCGTTTTGCCTTCTTTTATGTAACTATACTCCTCTAAGTTTCAAT
CTTGGCCATGTAACCTCTGATCTATAGAATTTTTTAAATGACTAGAATTAATGCCCATCTTTTTTTTGGACC
TAAATTCTTCATGAAAATATATTACGAGGGCTTATTCAGAAGCTTTGGACTTCTTCGCCAGAGGTTTGGTCA
AGTCTCCAATCAAGGTTGTCGGCTTGTCTACCTTGCCAGAAATTTACGAAAAGATGGAAAAGGGTCAAATCG
TTGGTAGATACGTTGTTGACACTTCTAAATAAGCGAATTTCTTATGATTTATGATTTTTATTATTAAATAAG
TTATAAAAAAAAATAAGTGTATACAAATTTTAAAGTGACTCTTAGGTTTTAAAACGAAAATTCTTATTCTTGA
GTAACTCTTTCCTGTAGGTCAGGTTGCTTTCTCAGGTATAGCATGAGGTCGCTCTTATTGACCACACCTCTA
CCGGCCGGTCGAAATTCCCCTACCCTATGAACATATTCCATTTTGTAATTTCGTGTCGTTTCTATTATGAAT
TTCATTTATAAAGTTTATGTACAAATATCATAAAAAAAGAGAATCTTTTTAAGCAAGGATTTTCTTAACTTC
TTCGGCGACAGCATCACCGACTTCGGTGGTACTGTTGGAACCACCTAAATCACCAGTTCTGATACCTGCATC
CAAAACCTTTTTAACTGCATCTTCAATGGCCTTACCTTCTTCAGGCAAGTTCAATGACAATTTCAACATCAT
TGCAGCAGACAAGATAGTGGCGATAGGGTTGACCTTATTCTTTGGCAAATCTGGAGCAGAACCGTGGCATGG
TTCGTACAAACCAAATGCGGTGTTCTTGTCTGGCAAAGAGGCCAAGGACGCAGATGGCAACAAACCCAAGGA
ACCTGGGATAACGGAGGCTTCATCGGAGATGATATCACCAAACATGTTGCTGGTGATTATAATACCATTTAG
GTGGGTTGGGTTCTTAACTAGGATCATGGCGGCAGAATCAATCAATTGATGTTGAACCTTCAATGTAGGAAA
TTCGTTCTTGATGGTTTCCTCCACAGTTTTTCTCCATAATCTTGAAGAGGCCAAAACATTAGCTTTATCCAA
GGACCAAATAGGCAATGGTGGCTCATGTTGTAGGGCCATGAAAGCGGCCATTCTTGTGATTCTTTGCACTTC
TGGAACGGTGTATTGTTCACTATCCCAAGCGACACCATCACCATCGTCTTCCTTTCTCTTACCAAAGTAAAT
ACCTCCCACTAATTCTCTGACAACAACGAAGTCAGTACCTTTAGCAAATTGTGGCTTGATTGGAGATAAGTC
TAAAAGAGAGTCGGATGCAAAGTTACATGGTCTTAAGTTGGCGTACAATTGAAGTTCTTTACGGATTTTTAG
TAAACCTTGTTCAGGTCTAACACTACCTGTACCCCATTTAGGACCACCCACAGCACCTAACAAAACGGCATC
AGCCTTCTTGGAGGCTTCCAGCGCCTCATCTGGAAGTGGGACACCTGTAGCTTCGATAGCAGCACCACCAAT
TAAATGATTTTCGAAATCGAACTTGACATTGGAACGAACATCAGAAATAGCTTTAAGAACCTTAATGGCTTC
GGCTGTGATTTCTTGACCAACGTGGTCACCTGGCAAAACGACGATCTTCTTAGGGGCAGACATTAGAATGGT
ATATCCTTGAAATATATATATATATTGCTGAAATGTAAAAGGTAAGAAAAGTTAGAAAGTAAGACGATTGCT
AACCACCTATTGGAAAAAACAATAGGTCCTTAAATAATATTGTCAACTTCAAGTATTGTGATGCAAGCATTT
AGTCATGAACGCTTCTCTATTCTATATGAAAAGCCGGTTCCGGCGCTCTCACCTTTCCTTTTTCTCCCAATT
TTTCAGTTGAAAAAGGTATATGCGTCAGGCGACCTCTGAAATTAACAAAAAATTTCCAGTCATCGAATTTGA
TTCTGTGCGATAGCGCCCCTGTGTGTTCTCGTTATGTTGAGGAAAAAAATAATGGTTGCTAAGAGATTCGAA
CTCTTGCATCTTACGATACCTGAGTATTCCCACAGTTGGGGGATCTCGACTCTAGCTAGAGGATCAATTCGT
AATCATGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAG
CATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGC
TTTCCAGTCGGGAAACCTGTCGTGCCAGCTGATAACTTCGTATAATGTATGCTATACGAAGTTATTAGGTCT
GAAGAGGAGTTTACGTCCAGCCAAGCTAGCTTGGCTGCAGGTCGAGCGGCCGCGATCCGGAACCCTTAATAT
AACTTCGTATAATGTATGCTATACGAAGTTATCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCG
GTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAG
CGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT
GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCC

CCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACC
AGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCG
CCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCG
TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC
GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG
CGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTAT
TTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAA
CCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAG
ATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA
GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATAT
ATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTC
GTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCA
GTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAA
GGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA
GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCT
CGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGT
GCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCA
TGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGT
ACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA
ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAA
GGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTA
CTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACAC
GGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGA
GCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC
CACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTC
GTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTC
TGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGC
TTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATAACGCATTTAAGCATAAACACGCACTAT
GCCGTTCTTCTCATGTATATATATATACAGGCAACACGCAGATATAGGTGCGACGTGAACAGTGAGCTGTAT
GTGCGCAGCTCGCGTTGCATTTTCGGAAGCGCTCGTTTTCGGAAACGCTTTGAAGTTCCTATTCCGAAGTTC
CTATTCTCTAGCTAGAAAGTATAGGAACTTCAGAGCGCTTTTGAAAACCAAAAGCGCTCTGAAGACGCACTT
TCAAAAAACCAAAAACGCACCGGACTGTAACGAGCTACTAAAATATTGCGAATACCGCTTCCACAAACATTG
CTCAAAAGTATCTCTTTGCTATATATCTCTGTGCTATATCCCTATATAACCTACCCATCCACCTTTCGCTCC
TTGAACTTGCATCTAAACTCGACCTCTACATTTTTTATGTTTATCTCTAGTATTACTCTTTAGACAAAAAAA
TTGTAGTAAGAACTATTCATAGAGTGAATCGAAAACAATACGAAAATGTAAACATTTCCTATACGTAGTATA
TAGAGACAAAATAGAAGAAACCGTTCATAATTTTCTGACCAATGAAGAATCATCAACGCTATCACTTTCTGT
TCACAAAGTATGCGCAATCCACATCGGTATAGAATATAATCGGGGATGCCTTTATCTTGAAAAAATGCACCC
GCAGCTTCGCTAGTAATCAGTAAACGCGGGAAGTGGAGTCAGGCTTTTTTTATGGAAGAGAAAATAGACACC
AAAGTAGCCTTCTTCTAACCTTAACGGACCTACAGTGCAAAAAGTTATCAAGAGACTGCATTATAGAGCGCA
CAAAGGAGAAAAAAAAGTAATCTAAGATGCTTTGTTAGAAAAATAGCGCTCTCGGGATGCATTTTTGTAGAAC
AAAAAAGAAGTATAGATTCTTTGTTGGTAAAATAGCGCTCTCGCGTTGCATTTCTGTTCTGTAAAAATGCAG
CTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTTTACAAAAATGAAGCACAGATTC
TTCGTTGGTAAAATAGCGCTTTCGCGTTGCATTTCTGTTCTGTAAAAATGCAGCTCAGATTCTTTGTTTGAA
AAATTAGCGCTCTCGCGTTGCATTTTTGTTCTACAAAATGAAGCACAGATGCTTCGTTGCT

> pGADT7-RecAB STX8

TGCATGCCTGCAGGTCGAGATCCGGGATCGAAGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATGAA
GGCAAAAGACAAATATAAGGGTCGAACGAAAAATAAAGTGAAAAGTGTTGATATGATGTATTTGGCTTTGCG
GCGCCGAAAAAACGAGTTTACGCAATTGCACAATCATGCTGACTCTGTGGCGGACCCGCGCTCTTGCCGGCC
CGGCGATAACGCTGGGCGTGAGGCTGTGCCCGGCGGAGTTTTTTGCGCCTGCATTTTCCAAGGTTTACCCTG
CGCTAAGGGGCGAGATTGGAGAAGCAATAAGAATGCCGGTTGGGGTTGCGATGATGACGACCACGACAACTG
GTGTCATTATTTAAGTTGCCGAAAGAACCTGAGTGCATTTGCAACATGAGTATACTAGAAGAATGAGCCAAG
ACTTGCGAGACGCGAGTTTGCCGGTGGTGCGAACAATAGAGCGACCATGACCTTGAAGGTGAGACGCGCATA
ACCGCTAGAGTACTTTGAAGAGGAAACAGCAATAGGGTTGCTACCAGTATAAATAGACAGGTACATACAACA
CTGGAAATGGTTGTCTGTTTGAGTACGCTTTCAATTCATTTGGGTGTGCACTTTATTATGTTACAATATGGA
AGGGAACTTTACACTTCTCCTATGCACATATATTAATTAAAGTCCAATGCTAGTAGAGAAGGGGGGTAACAC
CCCTCCGCGCTCTTTTCCGATTTTTTTCTAAACCGTGGAATATTTCGGATATCCTTTTGTTGTTTCCGGGTG
TACAATATGGACTTCCTCTTTTCTGGCAACCAAACCCATACATCGGGATTCCTATAATACCTTCGTTGGTCT
CCCTAACATGTAGGTGGCGGAGGGGAGATATACAATAGAACAGATACCAGACAAGACATAATGGGCTAAACA

AGACTACACCAATTACACTGCCTCATTGATGGTGGTACATAACGAACTAATACTGTAGCCCTAGACTTGATA
GCCATCATCATATCGAAGTTTCACTACCCTTTTTCCATTTGCCATCTATTGAAGTAATAATAGGCGCATGCA
ACTTCTTTTCTTTTTTTTTCTTTTCTCTCTCCCCCGTTGTTGTCTCACCATATCCGCAATGACAAAAAAATG
ATGGAAGACACTAAAGGAAAAAATTAACGACAAAGACAGCACCAACAGATGTCGTTGTTCCAGAGCTGATGA
GGGGTATCTCGAAGCACACGAAACTTTTTCCTTCCTTCATTCACGCACACTACTCTCTAATGAGCAACGGTA
TACGGCCTTCCTTCCAGTTACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTTGCTATCAAGTATAAATAG
ACCTGCAATTATTAATCTTTTGTTTCCTCGTCATTGTTCTCGTTCCCTTTCTTCCTTGTTTCTTTTTCTGCA
CAATATTTCAAGCTATACCAAGCATACAATCAACTCCAAGCTTTGCAAAGATGGATAAAGCGGAATTAATTC
CCGAGCCTCCAAAAAAGAAGAGAAAGGTCGAATTGGGTACCGCCGCCAATTTTAATCAAAGTGGGAATATTG
CTGATAGCTCATTGTCCTTCACTTTCACTAACAGTAGCAACGGTCCGAACCTCATAACAACTCAAACAAATT
CTCAAGCGCTTTCACAACCAATTGCCTCCTCTAACGTTCATGATAACTTCATGAATAATGAAATCACGGCTA
GTAAAATTGATGATGGTAATAATTCAAAACCACTGTCACCTGGTTGGACGGACCAAACTGCGTATAACGCGT
TTGGAATCACTACAGGGATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATTCGATGATGAAG
ATACCCCACCAAACCCAAAAAAAGAGATCTTTAATACGACTCACTATAGGGCGAGCGCCGCC

Primers to clone STX8 in pcDNA5/FRT

ATGGAGTACCCATACGACGTACCAGATTACGCTCATATGAACATGGAGGCCAGTGAATTCCACCCAAGCAGT
GGTATCAACGCAGAGTGGCCATTACGGCCGGGGGGGGGAGTCTGCAGGATGGCACCGGACCCCTGGTTCTCC
ACATACGATTCTACTTGTCAAATTGCCCAAGAAATTGCTGAGAAAATTCAACAACGAAATCAATATGAACGA
AAAGGTGAAAAGGCACCAAGGCTTACCGTGACAATCAGAGCTTTGTTGCAGAACCTGAAGGAAAAGATCGCC
CTTTTGAAGGACTTATTGCTAAGAGCTGTGTCAACACATCAGATAACACAGCTTGAAGGGGACCGAAGACAG
AACCTCTTGGATGATCTTGTAACTCGAGAGAGACTACTTCTGGCATCCTTTAAGAATGAGGGTGCCGAACCA
GATCTAATCAGGTCCAGCCTGATGAGTGAAGAGGCTAAGCGAGGAGCACCCAACCCTTGGCTCTTTGAGGAG
CCAGAGGAGACCAGAGGCTTGGGTTTTGATGAAATCCGGCAACAGCAGCAGAAAATTATCCAAGAACAGGAT
GCAGGCCTTGATGCCCTTTCCTCTATCATAAGTCGCCAAAAACAAATGGGGCAGGAAATTGGGAATGAATTG
GATGAACAAAATGAGATAATTGACGACCTTGCCAACCTAGTGGAGAACACAGATGAAAAACTTCGCAATGAA
ACCAGGCGGGTAAACATGGTGGACAGAAAGTCAGCCTCTTGTGGGATGATCATGGTGATTTTACTGCTGCTT
GTGGCTATCGTGGTTGTTGCAGTCTGGCCGACCAACTGAAAAAAAAAAAAAAAAAAAGGCCGCCTCGGCCTC
TAGAGGGTGGGCATCGATACGGGATCCATCGAGCTCGAGCTGCAGATGAATCGTAGATACTGAAAAACCCCG
CAAGTTCACTTCAACTGTGCATCGTGCACCATCTCAATTTCTTTCATTTATACATCGTTTTGCCTTCTTTTA
TGTAACTATACTCCTCTAAGTTTCAATCTTGGCCATGTAACCTCTGATCTATAGAATTTTTTAAATGACTAG
AATTAATGCCCATCTTTTTTTTTGGACCTAAATTCTTCATGAAAATATATTACGAGGGCTTATTCAGAAGCTT
TGGACTTCTTCGCCAGAGGTTTGGTCAAGTCTCCAATCAAGGTTGTCGGCTTGTCTACCTTGCCAGAAATTT
ACGAAAAGATGGAAAAGGGTCAAATCGTTGGTAGATACGTTGTTGACACTTCTAAATAAGCGAATTTCTTAT
GATTTATGATTTTTATTATTAAATAAGTTATAAAAAAAAATAAGTGTATACAAATTTTAAAGTGACTCTTAGG
TTTTAAAACGAAAATTCTTATTCTTGAGTAACTCTTTCCTGTAGGTCAGGTTGCTTTCTCAGGTATAGCATG
AGGTCGCTCTTATTGACCACACCTCTACCGGCCGGTCGAAATTCCCCTACCCTATGAACATATTCCATTTTG
TAATTTCGTGTCGTTTCTATTATGAATTTCATTTATAAAGTTTATGTACAAATATCATAAAAAAAGAGAATC
TTTTTAAGCAAGGATTTTCTTAACTTCTTCGGCGACAGCATCACCGACTTCGGTGGTACTGTTGGAACCACC
TAAATCACCAGTTCTGATACCTGCATCCAAAACCTTTTTAACTGCATCTTCAATGGCCTTACCTTCTTCAGG
CAAGTTCAATGACAATTTCAACATCATTGCAGCAGACAAGATAGTGGCGATAGGGTTGACCTTATTCTTTGG
CAAATCTGGAGCAGAACCGTGGCATGGTTCGTACAAACCAAATGCGGTGTTCTTGTCTGGCAAAGAGGCCAA
GGACGCAGATGGCAACAAACCCAAGGAACCTGGGATAACGGAGGCTTCATCGGAGATGATATCACCAAACAT
GTTGCTGGTGATTATAATACCATTTAGGTGGGTTGGGTTCTTAACTAGGATCATGGCGGCAGAATCAATCAA
TTGATGTTGAACCTTCAATGTAGGAAATTCGTTCTTGATGGTTTCCTCCACAGTTTTTCTCCATAATCTTGA
AGAGGCCAAAACATTAGCTTTATCCAAGGACCAAATAGGCAATGGTGGCTCATGTTGTAGGGCCATGAAAGC
GGCCATTCTTGTGATTCTTTGCACTTCTGGAACGGTGTATTGTTCACTATCCCAAGCGACACCATCACCATC
GTCTTCCTTTCTCTTACCAAAGTAAATACCTCCCACTAATTCTCTGACAACAACGAAGTCAGTACCTTTAGC
AAATTGTGGCTTGATTGGAGATAAGTCTAAAAGAGAGTCGGATGCAAAGTTACATGGTCTTAAGTTGGCGTA
CAATTGAAGTTCTTTACGGATTTTTAGTAAACCTTGTTCAGGTCTAACACTACCTGTACCCCATTTAGGACC
ACCCACAGCACCTAACAAAACGGCATCAGCCTTCTTGGAGGCTTCCAGCGCCTCATCTGGAAGTGGGACACC
TGTAGCTTCGATAGCAGCACCACCAATTAAATGATTTTCGAAATCGAACTTGACATTGGAACGAACATCAGA
AATAGCTTTAAGAACCTTAATGGCTTCGGCTGTGATTTCTTGACCAACGTGGTCACCTGGCAAAACGACGAT
CTTCTTAGGGGCAGACATTAGAATGGTATATCCTTGAAATATATATATATATTGCTGAAATGTAAAAGGTAA
GAAAAGTTAGAAAGTAAGACGATTGCTAACCACCTATTGGAAAAAACAATAGGTCCTTAAATAATATTGTCA
ACTTCAAGTATTGTGATGCAAGCATTTAGTCATGAACGCTTCTCTATTCTATATGAAAAGCCGGTTCCGGCG
CTCTCACCTTTCCTTTTTCTCCCAATTTTTCAGTTGAAAAAGGTATATGCGTCAGGCGACCTCTGAAATTAA
CAAAAAATTTCCAGTCATCGAATTTGATTCTGTGCGATAGCGCCCCTGTGTGTTCTCGTTATGTTGAGGAAA
AAAATAATGGTTGCTAAGAGATTCGAACTCTTGCATCTTACGATACCTGAGTATTCCCACAGTTGGGGGATC

TCGACTCTAGCTAGAGGATCAATTCGTAATCATGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCAC
AATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCAC
ATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGATAACTTCGTATAA
TGTATGCTATACGAAGTTATTAGGTCTGAAGAGGAGTTTACGTCCAGCCAAGCTAGCTTGGCTGCAGGTCGA
GCGGCCGCGATCCGGAACCCTTAATATAACTTCGTATAATGTATGCTATACGAAGTTATCAGCTGCATTAAT
GAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGC
TGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAAT
CAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGT
TGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC
GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGA
CCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCT
GTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCG
ACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAG
CAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA
ACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAG
TTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTA
CGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAA
ACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAAT
GAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG
CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGA
TACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATT
TATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCC
AGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCA
TTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA
GGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAA
GTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCG
TAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTT
GCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAA
AACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTG
CACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATG
CCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAA
GCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGG
TTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATA
AAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGC
AGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAG
CGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATA
ACGCATTTAAGCATAAACACGCACTATGCCGTTCTTCTCATGTATATATATATACAGGCAACACGCAGATAT
AGGTGCGACGTGAACAGTGAGCTGTATGTGCGCAGCTCGCGTTGCATTTTCGGAAGCGCTCGTTTTCGGAAA
CGCTTTGAAGTTCCTATTCCGAAGTTCCTATTCTCTAGCTAGAAAGTATAGGAACTTCAGAGCGCTTTTGAA
AACCAAAAGCGCTCTGAAGACGCACTTTCAAAAAACCAAAAACGCACCGGACTGTAACGAGCTACTAAAATA
TTGCGAATACCGCTTCCACAAACATTGCTCAAAAGTATCTCTTTGCTATATATCTCTGTGCTATATCCCTAT
ATAACCTACCCATCCACCTTTCGCTCCTTGAACTTGCATCTAAACTCGACCTCTACATTTTTTATGTTTATC
TCTAGTATTACTCTTTAGACAAAAAAATTGTAGTAAGAACTATTCATAGAGTGAATCGAAAACAATACGAAA
ATGTAAACATTTCCTATACGTAGTATATAGAGACAAAATAGAAGAAACCGTTCATAATTTTCTGACCAATGA
AGAATCATCAACGCTATCACTTTCTGTTCACAAAGTATGCGCAATCCACATCGGTATAGAATATAATCGGGG
ATGCCTTTATCTTGAAAAAATGCACCCGCAGCTTCGCTAGTAATCAGTAAACGCGGGAAGTGGAGTCAGGCT
TTTTTTATGGAAGAGAAAATAGACACCAAAGTAGCCTTCTTCTAACCTTAACGGACCTACAGTGCAAAAAGT
TATCAAGAGACTGCATTATAGAGCGCACAAAGGAGAAAAAAAGTAATCTAAGATGCTTTGTTAGAAAAATAG
CGCTCTCGGGATGCATTTTTGTAGAACAAAAAAGAAGTATAGATTCTTTGTTGGTAAAATAGCGCTCTCGCG
TTGCATTTCTGTTCTGTAAAAATGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTT
TGTTTTACAAAAATGAAGCACAGATTCTTCGTTGGTAAAATAGCGCTTTCGCGTTGCATTTCTGTTCTGTAA
AAATGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTCTACAAAATGAAGCA
CAGATGCTTCGTTGCT

7.5 in silico sequences of prey cDNAs clones in pcDNA5/FRT

>pcDNA5/FRT ATP1β1

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA
ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCAT
GGAGTACCCATACGACGTACCAGATTACGCTCATATGAACATGGAGGCCAGTGAATTCCACCCAAGCAGTGG
TATCAACGCAGAGTGGCCATTACGGCCGGGGTGCTCTGGATTAAATGATGAAACTTATGGCTACAAAGAGGG
CAAACCGTGCATTATTATAAAGCTCAACCGAGTTCTAGGCTTCAAACCTAAGCCTCCCAAGAATGAGTCCTT
GGAGACTTACCCAGTGATGAAGTATAACCCAAATGTCCTTCCCGTTCAGTGCACTGGCAAGCGAGATGAAGA
TAAGGATAAAGTTGGAAATGTGGAGTATTTTGGACTGGGCAACTCCCCTGGTTTTCCTCTGCAGTATTATCC
GTACTATGGCAAACTCCTGCAGCCCAAATACCTGCAGCCCCTGCTGGCCGTACAGTTCACCAATCTTACCAT
GGACACTGAAATTCGCATAGAGTGTAAGGCGTACGGTGAGAACATTGGGTACAGTGAGAAAGACCGTTTTCA
GGGACGTTTTGATGTAAAAATTGAAGTTAAGAGCTGAGCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCC
GCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGA
CCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGT
GTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATG
CTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACG
CGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCG
CCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTC
TAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGG
GTGATGGTTCACGTACCTAGAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCCTTGG
CCAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACC
TGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGC
GGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCC
CGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGG
GTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATG
CGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACA
CTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTATCACTGGCAAACTGTGATGGACGACA
CCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGC
ACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACT
GGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTT
GTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGT
ATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGGG
CGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAA
GCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTC
CGAGGGCAAAGGAATAGCACGTACTACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCG
GAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACC
CCAACTTGTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTT
TTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACC
TCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA
CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATT
GCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGC
GCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTT
CGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCA
GGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC
CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA
CTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACC
GGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGT
TCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTA
TCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC
AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACT
AGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGA
TCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAA

GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGG
ATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA
ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCG
ATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTA
CCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAAC
CAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGT
TGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC
GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGA
TCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCA
GTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCT
GTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCG
TCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG
CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCT
TCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGA
ATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGT
TATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTT
CCCCGAAAAGTGCCACCTGACGTC

>pcDNA5/FRT NPC2

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA
ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCAT
GGAGTACCCATACGACGTACCAGATTACGCTCATATGAACATGGAGGCCAGTGAATTCCACCCAAGCAGTGG
TATCAACGCAGAGTGGCCATTACGGCCGGGGGGGCCTGATGGTTGTAAGAGTGGAATTAACGGCCCTATCCA
AAAAGACAAGACCTATAGCTACCTGAATAAACTACCAGTGAAAAGCGAATATCCCTCTATAAAACTGGTGGT
GGAGTGGCAACTTCAGGATGACAAAAACCAAAGTCTCTTCTGCTGGGAAATCCCAGTACAGATCGTTTCTCA
TCTCTAAGCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTT
GCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTT
CCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGC
AGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTG
AGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGG
GTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCC
CTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGAT
TTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTACCTAGAAGTTCCTAT
TCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCCTTGGCCAAAAAGCCTGAACTCACCGCGACGTCTG
TCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTC
GTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACA
AAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAAT
TCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCG
AACTGCCCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGA
GCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTG
CTGATCCCCATGTGTATCACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCG
ATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACA
ATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAAT
ACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGC
GGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCT
ATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGAT
CCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAG
AAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTACTACGAG
ATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGA
TCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTT
ACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGT
CCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGT

CATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGT
GTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT
CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGC
GCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACT
CAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGC
AAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATC
ACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG
GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGG
GAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGG
GCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACC
CGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG
GTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC
TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCG
GTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTT
CTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGA
TCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGT
CTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG
CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAC
CGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAA
GTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGC
CAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGG
CTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA
GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCAC
TGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAT
TCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATA
GCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGT
TGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT
CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATAC
TCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTG
AATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

>pcDNA5/FRT PCNA

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA
ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCAT
GGAGTACCCATACGACGTACCAGATTACGCTCATATGAACATGGAGGCCAGTGAATTCCACCCAAGCAGTGG
TATCAACGCAGAGTGGCCATTACGGCCGGGGGATCTCAGCCATATTGGAGATGCTGTTGTAATTTCCTGTGC
AAAAGACGGAGTGAAATTTTCTGCAAGTGGAGAACTTGGAAATGGAAACATTAAATTGTCACAGACAAGTAA
TGTCGATAAAGAGGAGGAAGCTGTTACCATAGAGATGAATGAACCAGTTCAACTAACTTTTGCACTGAGGTA
CCTGAACTTCTTTACAAAAGCCACTCCACTCTCTTCAACGGTGACACTCAGTATGTCTGCAGATGTACCCCT
TGTTGTAGAGTATAAAATTGCGGATATGGGACACTTAAAATACTACTTGGCTCCCAAGATCGAGGATGAAGA
AGGATCTTAGGCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTA
GTTGCCGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCT
TTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGG
GCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTC
TGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGC
GGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTT
CCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCG
ATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTACCTAGAAGTTCCT
ATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCCTTGGCCAAAAAGCCTGAACTCACCGCGACGTC
TGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATC
TCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTA
CAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGA

ATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAAC
CGAACTGCCCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGAC
GAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGAT
TGCTGATCCCCATGTGTATCACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCT
CGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAA
CAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCA
ATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGA
GCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACT
CTATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCG
ATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGT
AGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTACTACG
AGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGAT
GATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGG
TTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTT
GTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATG
GTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAA
GTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCA
GTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGG
GCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA
CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCA
GCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCA
TCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCC
TGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTC
GGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT
GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAA
CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGG
CGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGC
TCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAG
CGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT
TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAG
GATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTG
GTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGT
TGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGAT
ACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG
AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTC
GCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTAT
GGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGT
TAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGC
ACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC
ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACA
TAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCT
GTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGT
TTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAAT
ACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT
TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

>pcDNA5/FRT PGM1

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA
ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCAT
GGAGTACCCATACGACGTACCAGATTACGCTCATATGAACATGGAGGCCAGTGAATTCCACCCAAGCAGTGG
TATCAACGCAGAGTGGCCATTACGGCCGGGGGTGGAGAAGGCCGATAACTTTGAATACAGCGACCCAGTGGA
TGGAAGCATTTCAAGAAATCAGGGCTTGCGCCTCATTTTCACAGATGGTTCTCGAATCGTCTTCCGACTGAG

CGGCACTGGGAGTGCCGGGGCCACCATTCGGCTGTACATCGATAGCTATGAGAAGGACGTTGCCAAGATTAA
CCAGGACCCCCAGGTCATGTTGGCCCCCCTTATTTCCATTGCTCTGAAAGTGTCCCAGCTGCAGGAGAGGAC
GGGACGCACTGCACCCACTGTCATCACCTAAGCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGAT
CAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGG
AAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATT
CTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGG
ATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCT
GTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAG
CGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATC
GGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATG
GTTCACGTACCTAGAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCCTTGGCCAAAA
AGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCGACCTGATGCA
GCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAA
TAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCC
GGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCAC
GTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGC
TGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATG
GCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTATCACTGGCAAACTGTGATGGACGACACCGTCAG
TGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGT
GCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGGAGCGA
GGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGTATGGA
GCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATATGCT
CCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGGGCGCAGGG
TCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGCGGC
CGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGC
AAAGGAATAGCACGTACTACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGT
TTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTT
GTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTC
ACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAG
CTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAA
CATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTT
GCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGG
GAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCG
CGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAG
AACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGG
CTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA
AGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATAC
CTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTG
TAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGT
AACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATT
AGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGG
ACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGC
AAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCT
CAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTG
GTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAA
AGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT
CTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCT
GGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCA
GCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGG
GAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTG
TCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCC
ATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTA
TCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACT
GGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATA
CGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAA
CTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCA
TCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGG
GCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGT
CTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGA
AAAGTGCCACCTGACGTC

>pcDNA5/FRT RNF-2

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA
ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCAT
GGAGTACCCATACGACGTACCAGATTACGCTCATATGAACATGGAGGCCAGTGAATTCCACCCAAGCAGTGG
TATCAACGCAGAGTGGCCATTACGGCCGGGGCAACCATTAAGCAAAACATGGGAACTCAGTTTATATGAGTT
ACAACGAACACCTCAGGAGGCAATAACAGATGGCTTGGAAATTGTGGTTTCACCTCGAAGTCTACACAGTGA
ATTAATGTGCCCAATTTGTTTGGATATGTTGAAGAACACCATGACTACAAAGGAGTGTTTACATCGTTTTTG
TGCAGACTGCATCATCACAGCCCTTAGAAGTGGCAACAAAGAATGTCCTACCTGTCGGAAAAAACTAGTTTC
CAAAAGATCACTAAGGCCAGACCCAAACTTTGATGCACTCATCAGCAAAATTTATCCAAGTCGTGATGAGTA
TGAAGCTCATCAAGAGAGAGTATTAGCCAGGATCAACAAGCACAATAATCAGCAAGCACTCAGTCACAGCAT
TGAGGAAGGACTGAAGATACAGGCCATGAACAGACTGCAGCGAGGCAAGAAACAACAGATTGAAAATGGTAG
TGGAGCAGAAGATAATGGTGACAGTTCACACTGCAGTAATGCATCCACACATAGCAATCAGGAAGCAGGCCC
TAGTAACAAACGGACCAAAACATCTGATGATTCTGGGCTAGAGCTTGATAATAACAATGCAGCAATGGCAAT
TGATCCAGTAATGGATGGTGCTAGTGAAATTGAATTAGTATTCAGGCCTCATCCCACACTTATGGAAAAAGA
TGACAGTGCACAGACGANATACATAAAGACTTCTGGTAACGCCACTGTTGATCACTTATCCAAGTATCTGGC
TGTGAGGTTAGCTTTAGAAGAACTTCGAAGCAAAGGTGAATCAAACCAGATGAACCTTGATACAGCCAGTGA
GAAGCAGTATACCATTTATATAGCAACAGCCGGTGGCCAGTTCACTGTATTAAATGGNTCTTTCTTTGGAAT
TGGTCAGTGAGCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTA
GTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCC
TTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGG
GGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTT
CTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGG
CGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCT
TCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCC
GATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTACCTAGAAGTTCC
TATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCCTTGGCCAAAAAGCCTGAACTCACCGCGACGT
CTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAAT
CTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCT
ACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGG
AATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAA
CCGAACTGCCCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGA
CGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGA
TTGCTGATCCCCATGTGTATCACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTC
TCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCA
ACAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCC
AATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCG
AGCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAAC
TCTATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCC
GATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTG
TAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTACTAC
GAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGA
TGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATG
GTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTT
TGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCAT
GGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAA
AGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCC
AGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTG
GGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC
ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCC
AGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGC
ATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCC
CTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTT
CGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC

TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCA
ACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAG
GCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCG
CTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA
GCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT
TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAA
GGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTT
GGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAG
TTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGA
TACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA
GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTT
CGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTA
TGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGG
TTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAG
CACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT
CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCAC
ATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGC
TGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCG
TTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAA
TACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT
TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

>pcDNA5/FRT STX8

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA
ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCAT
GGAGTACCCATACGACGTACCAGATTACGCTCATATGAACATGGAGGCCAGTGAATTCCACCCAAGCAGTGG
TATCAACGCAGAGTGGCCATTACGGCCGGGGGGGGGAGTCTGCAGGATGGCACCGGACCCCTGGTTCTCCAC
ATACGATTCTACTTGTCAAATTGCCCAAGAAATTGCTGAGAAAATTCAACAACGAAATCAATATGAACGAAA
AGGTGAAAAGGCACCAAGGCTTACCGTGACAATCAGAGCTTTGTTGCAGAACCTGAAGGAAAAGATCGCCCT
TTTGAAGGACTTATTGCTAAGAGCTGTGTCAACACATCAGATAACACAGCTTGAAGGGGACCGAAGACAGAA
CCTCTTGGATGATCTTGTAACTCGAGAGAGACTACTTCTGGCATCCTTTAAGAATGAGGGTGCCGAACCAGA
TCTAATCAGGTCCAGCCTGATGAGTGAAGAGGCTAAGCGAGGAGCACCCAACCCTTGGCTCTTTGAGGAGCC
AGAGGAGACCAGAGGCTTGGGTTTTGATGAAATCCGGCAACAGCAGCAGAAAATTATCCAAGAACAGGATGC
AGGCCTTGATGCCCTTTCCTCTATCATAAGTCGCCAAAAACAAATGGGGCAGGAAATTGGGAATGAATTGGA
TGAACAAAATGAGATAATTGACGACCTTGCCAACCTAGTGGAGAACACAGATGAAAAACTTCGCAATGAAAC
CAGGCGGGTAAACATGGTGGACAGAAAGTCAGCCTCTTGTGGGATGATCATGGTGATTTTACTGCTGCTTGT
GGCTATCGTGGTTGTTGCAGTCTGGCCGACCAACTGAGCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCC
GCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGA
CCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGT
GTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATG
CTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACG
CGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCG
CCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTC
TAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGG
GTGATGGTTCACGTACCTAGAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCCTTGG
CCAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACC
TGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGC
GGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCC
CGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGG
GTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATG
CGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACA
CTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTATCACTGGCAAACTGTGATGGACGACA

CCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGC
ACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACT
GGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTT
GTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGT
ATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGGG
CGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAA
GCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTC
CGAGGGCAAAGGAATAGCACGTACTACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCG
GAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACC
CCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCAT
TTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGA
CCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTC
CACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAA
TTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGTGCATTAATGAATCGGCCAACG
CGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGT
TCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC
AGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTT
CCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG
ACTATAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACC
GGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGT
TCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTA
TCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC
AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACT
AGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGA
TCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAA
GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGG
ATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA
ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCG
ATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTA
CCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAAC
CAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGT
TGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC
GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGA
TCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCA
GTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCT
GTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCG
TCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG
CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCT
TCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGA
ATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGT
TATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTT
CCCCGAAAAGTGCCACCTGACGTC

7.6 In silico sequences of wild-type and pathogenic APS-1 mutated AIRE-CARD cDNA cloned in pcDNA/FRT

>pcDNA5/FRT AIRE-CARD cDNA

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA

ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCAT
GGAGGAGCAGAAGCTGATCTCAGAGGAGGACCTGCATATGGCCATGGCGACGGACGCGGCGCTACGCCGGCT
TCTGAGGCTGCACCGCACGGAGATCGCGGTGGCCGTGGACAGCGCCTTCCCACTGCTGCACGCGCTGGCTGA
CCACGACGTGGTCCCCGAGGACAAGTTTCAGGAGACGCTTCATCTGAAGGAAAAGGAGGGCTGCCCCCAGGC
CTTCCACGCCCTCCTGTCCTGGCTGCTGACCCAGGACTCCACAGCCATCCTGGACTTCTGGAGGGTGCTGTT
CAAGGACTACAACCTGGAGCGCTATGGCCGGCTGCAGCCCATCCTGGACAGCTTCTGAGCGGCCGCTCGAGT
CTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCC
CCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTG
CATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATT
GGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGG
GCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCG
TGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCG
CCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCG
ACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTACCTAGAAGTTCCTATTCCGAAGTTCCTATTCTCTAG
AAAGTATAGGAACTTCCTTGGCCAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAA
AGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAG
GAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGC
ACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATT
GCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGC
CGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGAC
CGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTATCACT
GGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCG
AGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCC
GCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCT
TCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAG
GATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCA
ATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGC
GTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAA
ACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTACTACGAGATTTCGATTCCACCGCCGCCT
TCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCA
TGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA
CAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTT
ATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAA
ATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAT
GAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC
TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCA
CTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT
CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA
AGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTC
AGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC
CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATA
GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCG
TTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGC
CACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT
GGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG
GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGC
AGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGT
GGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAA
TCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA
TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGG
CTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCG
CCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACG
TTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCC
AACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG
TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA
TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGC
GACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCA
TCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAC
CCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA
GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAAT
ATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAC
AAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

>pcDNA5/FRT AIRE-T16M cDNA

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA
ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCAT
GGAGGAGCAGAAGCTGATCTCAGAGGAGGACCTGCATATGGCCATGGCGACGGACGCGGCGCTACGCCGGCT
TCTGAGGCTGCACCGCATGGAGATCGCGGTGGCCGTGGACAGCGCCTTCCCACTGCTGCACGCGCTGGCTGA
CCACGACGTGGTCCCCGAGGACAAGTTTCAGGAGACGCTTCATCTGAAGGAAAAGGAGGGCTGCCCCCAGGC
CTTCCACGCCCTCCTGTCCTGGCTGCTGACCCAGGACTCCACAGCCATCCTGGACTTCTGGAGGGTGCTGTT
CAAGGACTACAACCTGGAGCGCTATGGCCGGCTGCAGCCCATCCTGGACAGCTTCTGAGCGGCCGCTCGAGT
CTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCC
CCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTG
CATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATT
GGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGG
GCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCG
TGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCG
CCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCG
ACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTACCTAGAAGTTCCTATTCCGAAGTTCCTATTCTCTAG
AAAGTATAGGAACTTCCTTGGCCAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAA
AGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAG
GAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGC
ACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATT
GCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGC
CGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGAC
CGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTATCACT
GGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCG
AGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCC
GCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCT
TCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAG
GATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCA
ATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGC
GTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAA
ACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTACTACGAGATTTCGATTCCACCGCCGCCT
TCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCA
TGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA
CAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTT
ATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAA
ATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAT
GAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC
TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCA
CTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT
CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA
AGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTC
AGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC
CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATA
GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCG
TTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGC
CACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT
GGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG
GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGC
AGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGT
GGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAA
TCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA
TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGG

CTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCG
CCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACG
TTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCC
AACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG
TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA
TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGC
GACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCA
TCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAC
CCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA
GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAAT
ATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAC
AAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

>pcDNA5/FRT AIRE-L29P cDNA

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA
ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCAT
GGAGGAGCAGAAGCTGATCTCAGAGGAGGACCTGCATATGGCCATGGCGACGGACGCGGCGCTACGCCGGCT
TCTGAGGCTGCACCGCACGGAGATCGCGGTGGCCGTGGACAGCGCCTTCCCACTGCCGCACGCGCTGGCTGA
CCACGACGTGGTCCCCGAGGACAAGTTTCAGGAGACGCTTCATCTGAAGGAAAAGGAGGGCTGCCCCCAGGC
CTTCCACGCCCTCCTGTCCTGGCTGCTGACCCAGGACTCCACAGCCATCCTGGACTTCTGGAGGGTGCTGTT
CAAGGACTACAACCTGGAGCGCTATGGCCGGCTGCAGCCCATCCTGGACAGCTTCTGAGCGGCCGCTCGAGT
CTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCC
CCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTG
CATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATT
GGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGG
GCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCG
TGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCG
CCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCG
ACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTACCTAGAAGTTCCTATTCCGAAGTTCCTATTCTCTAG
AAAGTATAGGAACTTCCTTGGCCAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAA
AGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAG
GAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGC
ACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATT
GCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGC
CGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGAC
CGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTATCACT
GGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCG
AGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCC
GCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCT
TCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAG
GATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCA
ATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGC
GTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAA
ACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTACTACGAGATTTCGATTCCACCGCCGCCT
TCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCA
TGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA

CAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTT
ATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAA
ATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAT
GAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC
TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCA
CTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT
CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA
AGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTC
AGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC
CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATA
GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCG
TTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGC
CACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT
GGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG
GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGC
AGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGT
GGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAA
TCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA
TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGG
CTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCG
CCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACG
TTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCC
AACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG
TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA
TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGC
GACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCA
TCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAC
CCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA
GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAAT
ATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAC
AAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

>pcDNA5/FRT AIRE-L93R cDNA

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA
ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCAT
GGAGGAGCAGAAGCTGATCTCAGAGGAGGACCTGCATATGGCCATGGCGACGGACGCGGCGCTACGCCGGCT
TCTGAGGCTGCACCGCACGGAGATCGCGGTGGCCGTGGACAGCGCCTTCCCACTGCTGCACGCGCTGGCTGA
CCACGACGTGGTCCCCGAGGACAAGTTTCAGGAGACGCTTCATCTGAAGGAAAAGGAGGGCTGCCCCCAGGC
CTTCCACGCCCTCCTGTCCTGGCTGCTGACCCAGGACTCCACAGCCATCCTGGACTTCCGGAGGGTGCTGTT
CAAGGACTACAACCTGGAGCGCTATGGCCGGCGGCAGCCCATCCTGGACAGCTTCTGAGCGGCCGCTCGAGT
CTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCC
CCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTG
CATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATT
GGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGG
GCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCG
TGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCG

CCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCG
ACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTACCTAGAAGTTCCTATTCCGAAGTTCCTATTCTCTAG
AAAGTATAGGAACTTCCTTGGCCAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAA
AGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAG
GAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGC
ACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATT
GCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGC
CGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGAC
CGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTATCACT
GGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCG
AGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCC
GCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCT
TCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAG
GATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCA
ATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGC
GTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAA
ACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTACTACGAGATTTCGATTCCACCGCCGCCT
TCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCA
TGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA
CAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTT
ATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAA
ATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAT
GAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC
TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCA
CTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT
CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA
AGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTC
AGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC
CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATA
GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCG
TTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGC
CACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT
GGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG
GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGC
AGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGT
GGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAA
TCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA
TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGG
CTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCG
CCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACG
TTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCC
AACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG
TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA
TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGC
GACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCA
TCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAC
CCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA
GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAAT
ATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAC
AAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

7.7 In silico sequence of the precursor AIRE-SAND cDNA clone

pGBKT7 AIRE-SAND

CGGTGCGGGCCTCTTCGCTATTACGCCAGATCCTTTTGTTGTTTCCGGGTGTACAATATGGACTTCCTCTTT
TCTGGCAACCAAACCCATACATCGGGATTCCTATAATACCTTCGTTGGTCTCCCTAACATGTAGGTGGCGGA
GGGGAGATATACAATAGAACAGATACCAGACAAGACATAATGGGCTAAACAAGACTACACCAATTACACTGC
CTCATTGATGGTGGTACATAACGAACTAATACTGTAGCCCTAGACTTGATAGCCATCATCATATCGAAGTTT
CACTACCCTTTTTCCATTTGCCATCTATTGAAGTAATAATAGGCGCATGCAACTTCTTTTCTTTTTTTTTCT
TTTCTCTCTCCCCCGTTGTTGTCTCACCATATCCGCAATGACAAAAAAATGATGGAAGACACTAAAGGAAAA
AATTAACGACAAAGACAGCACCAACAGATGTCGTTGTTCCAGAGCTGATGAGGGGTATCTCGAAGCACACGA
AACTTTTTCCTTCCTTCATTCACGCACACTACTCTCTAATGAGCAACGGTATACGGCCTTCCTTCCAGTTAC
TTGAATTTGAAATAAAAAAAAGTTTGCTGTCTTGCTATCAAGTATAAATAGACCTGCAATTATTAATCTTTT
GTTTCCTCGTCATTGTTCTCGTTCCCTTTCTTCCTTGTTTCTTTTTCTGCACAATATTTCAAGCTATACCAA
GCATACAATCAACTCCAAGCTTGAAGCAAGCCTCCTGAAAGATGAAGCTACTGTCTTCTATCGAACAAGCAT
GCGATATTTGCCGACTTAAAAAGCTCAAGTGCTCCAAAGAAAAACCGAAGTGCGCCAAGTGTCTGAAGAACA
ACTGGGAGTGTCGCTACTCTCCCAAAACCAAAAGGTCTCCGCTGACTAGGGCACATCTGACAGAAGTGGAAT
CAAGGCTAGAAAGACTGGAACAGCTATTTCTACTGATTTTTCCTCGAGAAGACCTTGACATGATTTTGAAAA
TGGATTCTTTACAGGATATAAAAGCATTGTTAACAGGATTATTTGTACAAGATAATGTGAATAAAGATGCCG
TCACAGATAGATTGGCTTCAGTGGAGACTGATATGCCTCTAACATTGAGACAGCATAGAATAAGTGCGACAT
CATCATCGGAAGAGAGTAGTAACAAAGGTCAAAGACAGTTGACTGTATCGCCGGAATTTGTAATACGACTCA
CTATAGGGCGAGCGCCATC

Primers to clone AIRE-SAND in pcDNA5/FRT

ATGGAGGAGCAGAAGCTGATCTCAGAGGAGGACCTGCATATGGCCATGGTCCTGGACAGCTTCCCCAAAGAT
GTGGACCTCAGCCAGCCCCGGAAGGGGAGGAAGCCCCCGGCCGTCCCCAAGGCTTTGGTACCGCCACCCAGA
CTCCCCACCAAGAGGAAGGCCTCAGAAGAGGCTCGAGCTGCCGCGCCAGCAGCCCTGACTCCAAGGGGCACC
GCCAGCCCAGGCTCTCAACTGAAGGCCAAGCCCCCCCAAGAAGCCGGAGAGCAGCGCAGAGCAGCAGCGCCTT
CCACTCGGGAACGGGATTCAGACCATGTCAGCTTCAGTCCAGAGAGCTGTGGCCATGTCCTCCGGGGACGTC
CCGGGAGCCCGAGGGGCCGTGGAGGGGATCCTCATCCAGCAGGTGTTTGAGTCAGGCGGCTCCAAGAAGTGC
ATCCAGGTTGGTGGGGAGTTCTACACTCCCAGCAAGTTCGAAGACTCCGGCAGTGGGAAGAACAAGGCCCGC
AGCAGCAGTGGCCCGAAGCCTCTGGTTCGAGCCAAGGGAGCCCAGGGCGCTGCCCCCGGTGGGGGTGAGGCT
AGGCTGGGCCAGCAGGGCAGCGTTCCCTAAGCCCCTCTGGCCCTCCCCAGTGGTCGACCTGCAGCGGCCGCA
TAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCGCGCTTGCAGCCAAGCTAAT
TCCGGGCGAATTTCTTATGATTTATGATTTTTATTATTAAATAAGTTATAAAAAAAATAAGTGTATACAAAT
TTTAAAGTGACTCTTAGGTTTTAAAACGAAAATTCTTATTCTTGAGTAACTCTTTCCTGTAGGTCAGGTTGC
TTTCTCAGGTATAGCATGAGGTCGCTCTTATTGACCACACCTCTACCGGCATGCAAGCTTGGCGTAATCATG
GTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAA
GTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCA
GTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGG
GCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA
CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCA
GCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCA
TCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCC
TGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTC
GGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT
GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAA
CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGG
CGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGC
TCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAG
CGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT
TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAG
GATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACCTGA
GGCTATGGCAGGGCCTGCCGCCCCGACGTTGGCTGCGAGCCCTGGGCCTTCACCCGAACTTGGGGGGTGGGG
TGGGGAAAAGGAAGAAACGCGGGCGTATTGGCCCCAATGGGGTCTCGGTGGGGTATCGACAGAGTGCCAGCC
CTGGGACCGAACCCCGCGTTTATGAACAAACGACCCAACACCGTGCGTTTTATTCTGTCTTTTTATTGCCGT
CATAGCGCGGGTTCCTTCCGGTATTGTCTCCTTCCGTGTTTCAGTTAGCCTCCCCCTAGGGTGGGCGAAGAA
CTCCAGCATGAGATCCCCGCGCTGGAGGATCATCCAGCCGGCGTCCCGGAAAACGATTCCGAAGCCCAACCT

TTCATAGAAGGCGGCGGTGGAATCGAAATCTCGTGATGGCAGGTTGGGCGTCGCTTGGTCGGTCATTTCGAA
CCCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCG
ATACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAAC
GCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCC
ACCATGATATTCGGCAAGCAGGCATCGCCATGAGTCACGACGAGATCCTCGCCGTCGGGCATGCTCGCCTTG
AGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCTTCGTCCAGATCATCCTGATCGACAAGACCG
GCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCGAATGGGCAGGTAGCCGGATCA
AGCGTATGCAGCCGCCGCATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGG
AGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAGCACAGCT
GCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTTGCAGTTCATTCAGGGCACCG
GACAGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAG
CCGATTGTCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGCAAT
CCATCTTGTTCAATCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGC
GGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCA
CCTGAACGAAGCATCTGTGCTTCATTTTGTAGAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACAAA
GAATCTGAGCTGCATTTTTACAGAACAGAAATGCAACGCGAAAGCGCTATTTTACCAACGAAGAATCTGTGC
TTCATTTTTGTAAAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACAAAGAATCTGAGCTGCATTTTT
ACAGAACAGAAATGCAACGCGAGAGCGCTATTTTACCAACAAAGAATCTATACTTCTTTTTTGTTCTACAAA
AATGCATCCCGAGAGCGCTATTTTTCTAACAAAGCATCTTAGATTACTTTTTTTCTCCTTTGTGCGCTCTAT
AATGCAGTCTCTTGATAACTTTTTGCACTGTAGGTCCGTTAAGGTTAGAAGAAGGCTACTTTGGTGTCTATT
TTCTCTTCCATAAAAAAAGCCTGACTCCACTTCCCGCGTTTACTGATTACTAGCGAAGCTGCGGGTGCATTT
TTTCAAGATAAAGGCATCCCCGATTATATTCTATACCGATGTGGATTGCGCATACTTTGTGAACAGAAAGTG
ATAGCGTTGATGATTCTTCATTGGTCAGAAAATTATGAACGGTTTCTTCTATTTTGTCTCTATATACTACGT
ATAGGAAATGTTTACATTTTCGTATTGTTTTCGATTCACTCTATGAATAGTTCTTACTACAATTTTTTTGTC
TAAAGAGTAATACTAGAGATAAACATAAAAAATGTAGAGGTCGAGTTTAGATGCAAGTTCAAGGAGCGAAAG
GTGGATGGGTAGGTTATATAGGGATATAGCACAGAGATATATAGCAAAGAGATACTTTTGAGCAATGTTTGT
GGAAGCGGTATTCGCAATATTTTAGTAGCTCGTTACAGTCCGGTGCGTTTTTGGTTTTTTGAAAGTGCGTCT
TCAGAGCGCTTTTGGTTTTCAAAAGCGCTCTGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGGAATAG
GAACTTCAAAGCGTTTCCGAAAACGAGCGCTTCCGAAAATGCAACGCGAGCTGCGCACATACAGCTCACTGT
TCACGTCGCACCTATATCTGCGTGTTGCCTGTATATATATATACATGAGAAGAACGGCATAGTGCGTGTTTA
TGCTTAAATGCGTACTTATATGCGTCTATTTATGTAGGATGAAAGGTAGTCTAGTACCTCCTGTGATATTAT
CCCATTCCATGCGGGGTATCGTATGCTTCCTTCAGCACTACCCTTTAGCTGTTCTATATGCTGCCACTCCTC
AATTGGATTAGTCTCATCCTTCAATGCTATCATTTCCTTTGATATTGGATCATACTAAGAAACCATTATTAT
CATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGA
AAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGC
CCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGT
ACTGAGAGTGCACCATAGATCAACGACATTACTATATATATAATATAGGAAGCATTTAATAGAACAGCATCG
TAATATATGTGTACTTTGCAGTTATGACGCCAGATGGCAGTAGTGGAAGATATTCTTTATTGAAAAATAGCT
TGTCACCTTACGTACAATCTTGATCCGGAGCTTTTCTTTTTTTGCCGATTAAGAATTAATTCGGTCGAAAAA
AGAAAAGGAGAGGGCCAAGAGGGAGGGCATTGGTGACTATTGAGCACGTGAGTATACGTGATTAAGCACACA
AAGGCAGCTTGGAGTATGTCTGTTATTAATTTCACAGGTAGTTCTGGTCCATTGGTGAAAGTTTGCGGCTTG
CAGAGCACAGAGGCCGCAGAATGTGCTCTAGATTCCGATGCTGACTTGCTGGGTATTATATGTGTGCCCAAT
AGAAAGAGAACAATTGACCCGGTTATTGCAAGGAAAATTTCAAGTCTTGTAAAAGCATATAAAAATAGTTCA
GGCACTCCGAAATACTTGGTTGGCGTGTTTCGTAATCAACCTAAGGAGGATGTTTTGGCTCTGGTCAATGAT
TACGGCATTGATATCGTCCAACTGCATGGAGATGAGTCGTGGCAAGAATACCAAGAGTTCCTCGGTTTGCCA
GTTATTAAAAGACTCGTATTTCCAAAAGACTGCAACATACTACTCAGTGCAGCTTCACAGAAACCTCATTCG
TTTATTCCCTTGTTTGATTCAGAAGCAGGTGGGACAGGTGAACTTTTGGATTGGAACTCGATTTCTGACTGG
GTTGGAAGGCAAGAGAGCCCCGAAAGCTTACATTTTATGTTAGCTGGTGGACTGACGCCAGAAAATGTTGGT
GATGCGCTTAGATTAAATGGCGTTATTGGTGTTGATGTAAGCGGAGGTGTGGAGACAAATGGTGTAAAAGAC
TCTAACAAAATAGCAAATTTCGTCAAAAATGCTAAGAAATAGGTTATTACTGAGTAGTATTTATTTAAGTAT
TGTTTGTGCACTTGCCGATCTATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGG
AAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAAT
AGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTT
GGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATG
GCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACC
CTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG
CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGC
TTAATGCGCCGCTACAGGGCGCGTCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGAT

7.8 In silico sequences of wild-type and mutated AIRE-SAND cloned in pcDNA5/FRT

>pcDNA5/FRT AIRE-SAND cDNA

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA
ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCATGGAGGAGCAGAAGCTGATCTCAGAGGAGGACCTG
CATATGGCCATGGTCCTGGACAGCTTCCCCAAAGATGTGGACCTCAGCCAGCCCCGGAAGGGGAGGAAGCCC
CCGGCCGTCCCCAAGGCTTTGGTACCGCCACCCAGACTCCCCACCAAGAGGAAGGCCTCAGAAGAGGCTCGA
GCTGCCGCGCCAGCAGCCCTGACTCCAAGGGGCACCGCCAGCCCAGGCTCTCAACTGAAGGCCAAGCCCCCC
AAGAAGCCGGAGAGCAGCGCAGAGCAGCAGCGCCTTCCACTCGGGAACGGGATTCAGACCATGTCAGCTTCA
GTCC

Primer to clone AIRE-SAND in pBIND

AGAGAGCTGTGGCCATGTCCTCCGGGGACGTCCCGGGAGCCCGAGGGGCCGTGGAGGGGATCCTCATCCAGC
AGGTGTTTGAGTCAGGCGGCTCCAAGAAGTGCATCCAGGTTGGTGGGGAGTTCTACACTCCCAGCAAGTTCG
AAGACTCCGGCAGTGGGAAGAACAAGGCCCGCAGCAGCAGTGGCCCGAAGCCTCTGGTTCGAGCCAAGGGAG
CCCAGGGCGCTGCCCCCGGTGGGGGTGAGGCTAGGCTGGGCCAGCAGGGCAGCGTTCCCTGAGCGGCCGCTC
GAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTT
TGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAA
ATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAG
GATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGC
TGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGC
AGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACG
TTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCAC
CTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTACCTAGAAGTTCCTATTCCGAAGTTCCTATTCT
CTAGAAAGTATAGGAACTTCCTTGGCCAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATC
GAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGAT
GTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTAT
CGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACC
TATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTG
CAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTC
GGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTAT
CACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGG
GCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAAT
GGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATC
TTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTT
GCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGAC
GGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTC
GGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGT
GGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTACTACGAGATTTCGATTCCACCGCC
GCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGAT
CTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGC
ATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTA
TCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTG
TGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCC
TAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGC
CAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCG
CTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGG
TTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGT

AAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCA
AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGC
TCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT
CATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCC
CCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA
TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTG
AAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC
TTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGC
AAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCT
CAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTT
TTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGC
TTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTG
TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCA
CCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTA
TCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGC
AACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGT
TCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCG
ATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACT
GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATG
CGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTG
CTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATG
TAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACA
GGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTT
CAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAAT
AAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

>pcDNA5/FRT SAND-G228W cDNA

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA
ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCATGGAGGAGCAGAAGCTGATCTCAGAGGAGGACCTG
CATATGGCCATGGTCCTGGACAGCTTCCCCAAAGATGTGGACCTCAGCCAGCCCCGGAAGGGGAGGAAGCCC
CCGGCCGTCCCCAAGGCTTTGGTACCGCCACCCAGACTCCCCACCAAGAGGAAGGCCTCAGAAGAGGCTCGA
GCTGCCGCGCCAGCAGCCCTGACTCCAAGGGGCACCGCCAGCCCAGGCTCTCAACTGAAGGCCAAGCCCCCC
AAGAAGCCGGAGAGCAGCGCAGAGCAGCAGCGCCTTCCACTCGGGAACGGGATTCAGACCATGTCAGCTTCA
GTCC

Primers to clone SAND-G228W in pBIND

AGAGAGCTGTGGCCATGTCCTCCGGGGACGTCCCGGGAGCCCGAGGGGCCGTGGAGGGGATCCTCATCCAGC
AGGTGTTTGAGTCAGGCGGCTCCAAGAAGTGCATCCAGGTTGGTTGGGAGTTCTACACTCCCAGCAAGTTCG
AAGACTCCGGCAGTGGGAAGAACAAGGCCCGCAGCAGCAGTGGCCCGAAGCCTCTGGTTCGAGCCAAGGGAG
CCCAGGGCGCTGCCCCCGGTGGGGGTGAGGCTAGGCTGGGCCAGCAGGGCAGCGTTCCCTGAGCGGCCGCTC
GAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTT
TGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAA
ATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAG
GATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGC
TGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGC
AGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACG
TTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCAC
CTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTACCTAGAAGTTCCTATTCCGAAGTTCCTATTCT
CTAGAAAGTATAGGAACTTCCTTGGCCAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATC
GAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGAT

GTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTAT
CGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACC
TATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTG
CAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTC
GGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTAT
CACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGG
GCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAAT
GGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATC
TTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTT
GCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGAC
GGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTC
GGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGT
GGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTACTACGAGATTTCGATTCCACCGCC
GCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGAT
CTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGC
ATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTA
TCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTG
TGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCC
TAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGC
CAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCG
CTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGG
TTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGT
AAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCA
AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGC
TCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT
CATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCC
CCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA
TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTG
AAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC
TTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGC
AAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCT
CAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTT
TTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGC
TTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTG
TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCA
CCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTA
TCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGC
AACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGT
TCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCG
ATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACT
GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATG
CGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTG
CTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATG
TAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACA
GGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTT
CAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAAT
AAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

>pcDNA5/FRT SAND-P252L cDNA

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTAC
GGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTC
ATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTT
GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTG
GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTA

ATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCATGGAGGAGCAGAAGCTGATCTCAGAGGAGGACCTG
CATATGGCCATGGTCCTGGACAGCTTCCCCAAAGATGTGGACCTCAGCCAGCCCCGGAAGGGGAGGAAGCCC
CCGGCCGTCCCCAAGGCTTTGGTACCGCCACCCAGACTCCCCACCAAGAGGAAGGCCTCAGAAGAGGCTCGA
GCTGCCGCGCCAGCAGCCCTGACTCCAAGGGGCACCGCCAGCCCAGGCTCTCAACTGAAGGCCAAGCCCCCC
AAGAAGCCGGAGAGCAGCGCAGAGCAGCAGCGCCTTCCACTCGGGAACGGGATTCAGACCATGTCAGCTTCA
GTCC

Primers to clone SAND-P252L in pBIND

AGAGAGCTGTGGCCATGTCCTCCGGGGACGTCCCGGGAGCCCGAGGGGCCGTGGAGGGGATCCTCATCCAGC
AGGTGTTTGAGTCAGGCGGCTCCAAGAAGTGCATCCAGGTTGGTGGGGAGTTCTACACTCCCAGCAAGTTCG
AAGACTCCGGCAGTGGGAAGAACAAGGCCCGCAGCAGCAGTGGCCTGAAGCCTCTGGTTCGAGCCAAGGGAG
CCCAGGGCGCTGCCCCCGGTGGGGGTGAGGCTAGGCTGGGCCAGCAGGGCAGCGTTCCCTGAGCGGCCGCTC
GAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTT
TGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAA
ATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAG
GATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGC
TGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGC
AGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACG
TTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCAC
CTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTACCTAGAAGTTCCTATTCCGAAGTTCCTATTCT
CTAGAAAGTATAGGAACTTCCTTGGCCAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATC
GAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGAT
GTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTAT
CGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACC
TATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTG
CAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTC
GGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTAT
CACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGG
GCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAAT
GGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATC
TTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTT
GCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGAC
GGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTC
GGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGT
GGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTACTACGAGATTTCGATTCCACCGCC
GCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGAT
CTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGC
ATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTA
TCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTG
TGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCC
TAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGC
CAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCG
CTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGG
TTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGT
AAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCA
AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGC
TCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT
CATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCC
CCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA
TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTG
AAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC
TTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGC
AAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCT
CAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTT
TTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGC
TTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTG
TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCA
CCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTA
TCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGC

AACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGT
TCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCG
ATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACT
GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATG
CGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTG
CTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATG
TAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACA
GGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTT
CAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAAT
AAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

7.9 In silico sequences of wild-type and mutated AIRE-SAND cloned in pBIND

>pBIND AIRE-SAND

TCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCA
TACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGGCATTG
ATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGT
TACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGAC
GTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGC
CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCC
CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT
CGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATT
TCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG
TCGTAACAACTGCGATCGCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA
AGCAGAGCTCGTTTAGTGAACCGTCAGATCACTAGAAGCTTTATTGCGGTAGTTTATCACAGTTAAATTGCT
AACGCAGTCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCTGCAGTGACTCTCTTAAGGTAGCCTTGCAG
AAGTTGGTCGTGAGGCACTGGGCAGGTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACT
GGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTTGC
CTTTCTCTCCACAGGTGTCCACTCCCAGTTCAATTACAGCTCTTAAGGCTAGAGTACTTAATACGACTCACT
ATAGGCTAGCCAGCTTGAAGCAAGCCTCCTGAAAGATGAAGCTACTGTCTTCTATCGAACAAGCATGCGATA
TTTGCCGACTTAAAAAGCTCAAGTGCTCCAAAGAAAAACCGAAGTGCGCCAAGTGTCTGAAGAACAACTGGG
AGTGTCGCTACTCTCCCAAAACCAAAAGGTCTCCGCTGACTAGGGCACATCTGACAGAAGTGGAATCAAGGC
TAGAAAGACTGGAACAGCTATTTCTACTGATTTTTCCTCGAGAAGACCTTGACATGATTTTGAAAATGGATT
CTTTACAGGATATAAAAGCATTGTTAACAGGATTATTTGTACAAGATAATGTGAATAAAGATGCCGTCACAG
ATAGATTGGCTTCAGTGGAGACTGATATGCCTCTAACATTGAGACAGCATAGAATAAGTGCGACATCATCAT
CGGAAGAGAGTAGTAACAAAGGTCAAAGACAGTTGACTGTATCGCCGGAATTCCCGGGGATCCGTCGACTTG
ACGCGTCAATGAGAGCTGTGGCCATGTCCTCCGGGGACGTCCCGGGAGCCCGAGGGGCCGTGGAGGGGATCC
TCATCCAGCAGGTGTTTGAGTCAGGCGGCTCCAAGAAGTGCATCCAGGTTGGTGGGGAGTTCTACACTCCCA
GCAAGTTCGAAGACTCCGGCAGTGGGAAGAACAAGGCCCGCAGCAGCAGTGGCCCGAAGCCTCTGGTTCGAG
CCAAGGGAGCCCAGGGCGCTGCCCCCGGTGGGGGTGAGGCTAGGCTGGGCCAGCAGGGCAGCGTTCCCTGAG
GTACCTGAATAACTAAGGCCGCTTCCCTTTAGTGAGGGTTAATGCTTCGAGCAGACATGATAAGATACATTG
ATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTG
CTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGG
TTCAGGGGGAGATGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCCGATAAGGAT
CGATCCGGGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGG
CGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA
CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCACGCCACGTTCGCCGGCTTTCCC
CGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAA
CTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAG
TCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTT
GATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCG
AATTTTAACAAAATATTAACGCTTACAATTTCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTT
CACACCGCATACGCGGATCTTCCGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGG
GTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAG
GTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAT
AGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTG

ACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGG
CTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTGATTCTTCTGACACAACAGTCTCGAACTTAAGGCTAGA
ATTCTGGCTACAGGTAAGCGCCCCTAAAATCCCTTTGGGCACAATGTGTCCTGAGGGGAGAGGCAGCGACCT
GTAGATGGGACGGGGGCACTAACCCTCAGGTTTGGGGCTTCTGAATGTGAGTATCGCCATGTAAGCCCAGTA
TTTGGCCAATCTCAGAAAGCTCCTGGTCCCTGGAGGGATGGAGAGAGAAAAACAAACAGCTCCTGGAGCAGG
GAGAGTGCTGGCCTCTTGCTCTCCGGCTCCCTCTGTTGCCCTCTGGTTTCTCCCCAGGCTCCCGGACGTCCT
CTAGCCACCATGACTTCGAAAGTTTATGATCCAGAACAAAGGAAACGGATGATAACTGGTCCGCAGTGGTGG
GCCAGATGTAAACAAATGAATGTTCTTGATTCATTTATTAATTATTATGATTCAGAAAAACATGCAGAAAAT
GCTGTTATTTTTTTACATGGTAACGCGGCCTCTTCTTATTTATGGCGACATGTTGTGCCACATATTGAGCCA
GTAGCGCGGTGTATTATACCAGACCTTATTGGTATGGGCAAATCAGGCAAATCTGGTAATGGTTCTTATAGG
TTACTTGATCATTACAAATATCTTACTGCATGGTTTGAACTTCTTAATTTACCAAAGAAGATCATTTTTGTC
GGCCATGATTGGGGTGCTTGTTTGGCATTTCATTATAGCTATGAGCATCAAGATAAGATCAAAGCAATAGTT
CACGCTGAAAGTGTAGTAGATGTGATTGAATCATGGGATGAATGGCCTGATATTGAAGAAGATATTGCGTTG
ATCAAATCTGAAGAAGGAGAAAAAATGGTTTTGGAGAATAACTTCTTCGTGGAAACCATGTTGCCATCAAAA
ATCATGAGAAAGTTAGAACCAGAAGAATTTGCAGCATATCTTGAACCATTCAAAGAGAAAGGTGAAGTTCGT
CGTCCAACATTATCATGGCCTCGTGAAATCCCGTTAGTAAAAGGTGGTAAACCTGACGTTGTACAAATTGTT
AGGAATTATAATGCTTATCTACGTGCAAGTGATGATTTACCAAAAATGTTTATTGAATCGGACCCAGGATTC
TTTTCCAATGCTATTGTTGAAGGTGCCAAGAAGTTTCCTAATACTGAATTTGTCAAAGTAAAAGGTCTTCAT
TTTTCGCAAGAAGATGCACCTGATGAAATGGGAAAATATATCAAATCGTTCGTTGAGCGAGTTCTCAAAAAT
GAACAATAATTCTAGCCCTGAATAAGTGATAATAAGCGGATGAATGGCAGAAATTCGTCGAAGCGCAATAAA
ATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGATAGCGATAAGGATCGGAAGATCC
GCGTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACAC
CCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGA
GCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTAT
TTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCG
GAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAA
TGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTG
CGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGG
GTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAAC
GTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAG
AGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATC
TTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACT
TACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTC
GCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAG
CAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAG
ACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTG
ATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCC
GTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAG
GTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAAC
TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTG
AGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGC
GCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTAC
CAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGT
AGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGG
CTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGC
GGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACC
TACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA
GGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGT
TTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCA
GCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGGCTCGACAGATCT

>pBIND SAND-G228W

TCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCA
TACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGGCATTG
ATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGT
TACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGAC
GTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGC
CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCC

CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT
CGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATT
TCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG
TCGTAACAACTGCGATCGCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA
AGCAGAGCTCGTTTAGTGAACCGTCAGATCACTAGAAGCTTTATTGCGGTAGTTTATCACAGTTAAATTGCT
AACGCAGTCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCTGCAGTGACTCTCTTAAGGTAGCCTTGCAG
AAGTTGGTCGTGAGGCACTGGGCAGGTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACT
GGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTTGC
CTTTCTCTCCACAGGTGTCCACTCCCAGTTCAATTACAGCTCTTAAGGCTAGAGTACTTAATACGACTCACT
ATAGGCTAGCCAGCTTGAAGCAAGCCTCCTGAAAGATGAAGCTACTGTCTTCTATCGAACAAGCATGCGATA
TTTGCCGACTTAAAAAGCTCAAGTGCTCCAAAGAAAAACCGAAGTGCGCCAAGTGTCTGAAGAACAACTGGG
AGTGTCGCTACTCTCCCAAAACCAAAAGGTCTCCGCTGACTAGGGCACATCTGACAGAAGTGGAATCAAGGC
TAGAAAGACTGGAACAGCTATTTCTACTGATTTTTCCTCGAGAAGACCTTGACATGATTTTGAAAATGGATT
CTTTACAGGATATAAAAGCATTGTTAACAGGATTATTTGTACAAGATAATGTGAATAAAGATGCCGTCACAG
ATAGATTGGCTTCAGTGGAGACTGATATGCCTCTAACATTGAGACAGCATAGAATAAGTGCGACATCATCAT
CGGAAGAGAGTAGTAACAAAGGTCAAAGACAGTTGACTGTATCGCCGGAATTCCCGGGGATCCGTCGACTTG
ACGCGTCAATGAGAGCTGTGGCCATGTCCTCCGGGGACGTCCCGGGAGCCCGAGGGGCCGTGGAGGGGATCC
TCATCCAGCAGGTGTTTGAGTCAGGCGGCTCCAAGAAGTGCATCCAGGTTGGTTGGGAGTTCTACACTCCCA
GCAAGTTCGAAGACTCCGGCAGTGGGAAGAACAAGGCCCGCAGCAGCAGTGGCCCGAAGCCTCTGGTTCGAG
CCAAGGGAGCCCAGGGCGCTGCCCCCGGTGGGGGTGAGGCTAGGCTGGGCCAGCAGGGCAGCGTTCCCTGAG
GTACCTGAATAACTAAGGCCGCTTCCCTTTAGTGAGGGTTAATGCTTCGAGCAGACATGATAAGATACATTG
ATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTG
CTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGG
TTCAGGGGGAGATGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCCGATAAGGAT
CGATCCGGGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGG
CGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA
CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCACGCCACGTTCGCCGGCTTTCCC
CGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAA
CTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAG
TCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTT
GATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCG
AATTTTAACAAAATATTAACGCTTACAATTTCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTT
CACACCGCATACGCGGATCTTCCGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGG
GTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAG
GTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAT
AGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTG
ACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGG
CTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTGATTCTTCTGACACAACAGTCTCGAACTTAAGGCTAGA
ATTCTGGCTACAGGTAAGCGCCCCTAAAATCCCTTTGGGCACAATGTGTCCTGAGGGGAGAGGCAGCGACCT
GTAGATGGGACGGGGGCACTAACCCTCAGGTTTGGGGCTTCTGAATGTGAGTATCGCCATGTAAGCCCAGTA
TTTGGCCAATCTCAGAAAGCTCCTGGTCCCTGGAGGGATGGAGAGAGAAAAACAAACAGCTCCTGGAGCAGG
GAGAGTGCTGGCCTCTTGCTCTCCGGCTCCCTCTGTTGCCCTCTGGTTTCTCCCCAGGCTCCCGGACGTCCT
CTAGCCACCATGACTTCGAAAGTTTATGATCCAGAACAAAGGAAACGGATGATAACTGGTCCGCAGTGGTGG
GCCAGATGTAAACAAATGAATGTTCTTGATTCATTTATTAATTATTATGATTCAGAAAAACATGCAGAAAAT
GCTGTTATTTTTTTACATGGTAACGCGGCCTCTTCTTATTTATGGCGACATGTTGTGCCACATATTGAGCCA
GTAGCGCGGTGTATTATACCAGACCTTATTGGTATGGGCAAATCAGGCAAATCTGGTAATGGTTCTTATAGG
TTACTTGATCATTACAAATATCTTACTGCATGGTTTGAACTTCTTAATTTACCAAAGAAGATCATTTTTGTC
GGCCATGATTGGGGTGCTTGTTTGGCATTTCATTATAGCTATGAGCATCAAGATAAGATCAAAGCAATAGTT
CACGCTGAAAGTGTAGTAGATGTGATTGAATCATGGGATGAATGGCCTGATATTGAAGAAGATATTGCGTTG
ATCAAATCTGAAGAAGGAGAAAAAATGGTTTTGGAGAATAACTTCTTCGTGGAAACCATGTTGCCATCAAAA
ATCATGAGAAAGTTAGAACCAGAAGAATTTGCAGCATATCTTGAACCATTCAAAGAGAAAGGTGAAGTTCGT
CGTCCAACATTATCATGGCCTCGTGAAATCCCGTTAGTAAAAGGTGGTAAACCTGACGTTGTACAAATTGTT
AGGAATTATAATGCTTATCTACGTGCAAGTGATGATTTACCAAAAATGTTTATTGAATCGGACCCAGGATTC
TTTTCCAATGCTATTGTTGAAGGTGCCAAGAAGTTTCCTAATACTGAATTTGTCAAAGTAAAAGGTCTTCAT
TTTTCGCAAGAAGATGCACCTGATGAAATGGGAAAATATATCAAATCGTTCGTTGAGCGAGTTCTCAAAAAT
GAACAATAATTCTAGCCCTGAATAAGTGATAATAAGCGGATGAATGGCAGAAATTCGTCGAAGCGCAATAAA
ATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGATAGCGATAAGGATCGGAAGATCC
GCGTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACAC
CCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGA
GCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTAT

TTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCG
GAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAA
TGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTG
CGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGG
GTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAAC
GTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAG
AGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATC
TTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACT
TACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTC
GCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAG
CAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAG
ACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTG
ATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCC
GTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAG
GTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAAC
TTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTG
AGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGC
GCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTAC
CAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGT
AGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGG
CTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGC
GGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACC
TACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA
GGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGT
TTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCA
GCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGGCTCGACAGATCT

7.10 In silico sequences of wild-type and mutated AIRE cloned in pBIND

>pBIND AIRE

TCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCA
TACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGGCATTG
ATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGT
TACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGAC
GTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGC
CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCC
CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT
CGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATT
TCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG
TCGTAACAACTGCGATCGCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA
AGCAGAGCTCGTTTAGTGAACCGTCAGATCACTAGAAGCTTTATTGCGGTAGTTTATCACAGTTAAATTGCT
AACGCAGTCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCTGCAGTGACTCTCTTAAGGTAGCCTTGCAG
AAGTTGGTCGTGAGGCACTGGGCAGGTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACT
GGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTTGC
CTTTCTCTCCACAGGTGTCCACTCCCAGTTCAATTACAGCTCTTAAGGCTAGAGTACTTAATACGACTCACT
ATAGGCTAGCCAGCTTGAAGCAAGCCTCCTGAAAGATGAAGCTACTGTCTTCTATCGAACAAGCATGCGATA
TTTGCCGACTTAAAAAGCTCAAGTGCTCCAAAGAAAAACCGAAGTGCGCCAAGTGTCTGAAGAACAACTGGG
AGTGTCGCTACTCTCCCAAAACCAAAAGGTCTCCGCTGACTAGGGCACATCTGACAGAAGTGGAATCAAGGC
TAGAAAGACTGGAACAGCTATTTCTACTGATTTTTCCTCGAGAAGACCTTGACATGATTTTGAAAATGGATT
CTTTACAGGATATAAAAGCATTGTTAACAGGATTATTTGTACAAGATAATGTGAATAAAGATGCCGTCACAG
ATAGATTGGCTTCAGTGGAGACTGATATGCCTCTAACATTGAGACAGCATAGAATAAGTGCGACATCATCAT
CGGAAGAGAGTAGTAACAAAGGTCAAAGACAGTTGACTGTATCGCCGGAATTCCCGGGGATCCGTCGACTTG
ACGCGTCTATGGCGACGGACGCGGCGCTACGCCGGCTTCTGAGGCTGCACCGCACGGAGATCGCGGTGGCCG
TGGACAGCGCCTTCCCACTGCTGCACGCGCTGGCTGACCACGACGTGGTCCCCGAGGACAAGTTTCAGGAGA
CGCTTCATCTGAAGGAAAAGGAGGGCTGCCCCCAGGCCTTCCACGCCCTCCTGTCCTGGCTGCTGACCCAGG
ACTCCACAGCCATCCTGGACTTCTGGAGGGTGCTGTTCAAGGACTACAACCTGGAGCGCTATGGCCGGCTGC

AGCCCATCCTGGACAGCTTCCCCAAAGATGTGGACCTCAGCCAGCCCCGGAAGGGGAGGAAGCCCCCGGCCG
TCCCCAAGGCTTTGGTACCGCCACCCAGACTCCCCACCAAGAGGAAGGCCTCAGAAGAGGCTCGAGCTGCCG
CGCCAGCAGCCCTGACTCCAAGGGGCACCGCCAGCCCAGGCTCTCAACTGAAGGCCAAGCCCCCCAAGAAGC
CGGAGAGCAGCGCAGAGCAGCAGCGCCTTCCACTCGGGAACGGGATTCAGACCATGTCAGCTTCAGTCCAGA
GAGCTGTGGCCATGTCCTCCGGGGACGTCCCGGGAGCCCGAGGGGCCGTGGAGGGGATCCTCATCCAGCAGG
TGTTTGAGTCAGGCGGCTCCAAGAAGTGCATCCAGGTTGGCGGGGAGTTCTACACTCCCAGCAAGTTCGAAG
ACTCCGGCAGTGGGAAGAACAAGGCCCGCAGCAGCAGTGGCCCGAAGCCTCTGGTTCGAGCCAAGGGAGCCC
AGGGCGCTGCCCCCGGTGGAGGTGAGGCTAGGCTGGGCCAGCAGGGCAGCGTTCCCGCCCCTCTGGCCCTCC
CCAGTGACCCCCAGCTCCACCAGAAGAATGAGGACGAGTGTGCCGTGTGTCGGGACGGCGGGGAGCTCATCT
GCTGTGACGGCTGCCCTCGGGCCTTCCACCTGGCCTGCCTGTCCCCTCCGCTCCGGGAGATCCCCAGTGGGA
CCTGGAGGTGCTCCAGCTGCCTGCAGGCAACAGTCCAGGAGGTGCAGCCCCGGGCAGAGGAGCCCCGGCCCC
AGGAGCCACCCGTGGAGACCCCGCTCCCCCCGGGGCTTAGGTCGGCGGGAGAGGAGGTAAGAGGTCCACCTG
GGGAACCCCTAGCCGGCATGGACACGACTCTTGTCTACAAGCACCTGCCGGCTCCGCCTTCTGCAGCCCCGC
TGCCAGGGCTGGACTCCTCGGCCCTGCACCCCCTACTGTGTGTGGGTCCTGAGGGTCAGCAGAACCTGGCTC
CTGGTGCGCGTTGCGGGGTGTGCGGAGATGGTACGGACGTGCTGCGGTGTACTCACTGCGCCGCTGCCTTCC
ACTGGCGCTGCCACTTCCCAGCCGGCACCTCCCGGCCCGGGACGGGCCTGCGCTGCAGATCCTGCTCAGGAG
ACGTGACCCCAGCCCCTGTGGAGGGGGTGCTGGCCCCCAGCCCCGCCCGCCTGGCCCCTGGGCCTGCCAAGG
ATGACACTGCCAGTCACGAGCCCGCTCTGCACAGGGATGACCTGGAGTCCCTTCTGAGCGAGCACACCTTCG
ATGGCATCCTGCAGTGGGCCATCCAGAGCATGGCCCGTCCGGCGGCCCCCTTCCCCTCCGGGGGTGGAGGCT
CTGAGCAGAAACTCATCTCAGAAGAGGATCTGTAAGCGGCCGCAGGTACCTGAATAACTAAGGCCGCTTCCC
TTTAGTGAGGGTTAATGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAA
TGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCA
ATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTT
AAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCCGATAAGGATCGATCCGGGCTGGCGTAATAGCGAAGA
GGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCA
TTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCT
TTCGCTTTCTTCCCTTCCTTTCACGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCT
TTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGT
GGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTG
TTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCG
GCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACA
ATTTCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATACGCGGATCTTCCGTAC
CTTCTGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGC
AGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGC
AGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCC
GCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGC
CGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAA
AAAGCTTGATTCTTCTGACACAACAGTCTCGAACTTAAGGCTAGAATTCTGGCTACAGGTAAGCGCCCCTAA
AATCCCTTTGGGCACAATGTGTCCTGAGGGGAGAGGCAGCGACCTGTAGATGGGACGGGGGCACTAACCCTC
AGGTTTGGGGCTTCTGAATGTGAGTATCGCCATGTAAGCCCAGTATTTGGCCAATCTCAGAAAGCTCCTGGT
CCCTGGAGGGATGGAGAGAGAAAAACAAACAGCTCCTGGAGCAGGGAGAGTGCTGGCCTCTTGCTCTCCGGC
TCCCTCTGTTGCCCTCTGGTTTCTCCCCAGGCTCCCGGACGTCCTCTAGCCACCATGACTTCGAAAGTTTAT
GATCCAGAACAAAGGAAACGGATGATAACTGGTCCGCAGTGGTGGGCCAGATGTAAACAAATGAATGTTCTT
GATTCATTTATTAATTATTATGATTCAGAAAAACATGCAGAAAATGCTGTTATTTTTTTACATGGTAACGCG
GCCTCTTCTTATTTATGGCGACATGTTGTGCCACATATTGAGCCAGTAGCGCGGTGTATTATACCAGACCTT
ATTGGTATGGGCAAATCAGGCAAATCTGGTAATGGTTCTTATAGGTTACTTGATCATTACAAATATCTTACT
GCATGGTTTGAACTTCTTAATTTACCAAAGAAGATCATTTTTGTCGGCCATGATTGGGGTGCTTGTTTGGCA
TTTCATTATAGCTATGAGCATCAAGATAAGATCAAAGCAATAGTTCACGCTGAAAGTGTAGTAGATGTGATT
GAATCATGGGATGAATGGCCTGATATTGAAGAAGATATTGCGTTGATCAAATCTGAAGAAGGAGAAAAAATG
GTTTTGGAGAATAACTTCTTCGTGGAAACCATGTTGCCATCAAAAATCATGAGAAAGTTAGAACCAGAAGAA
TTTGCAGCATATCTTGAACCATTCAAAGAGAAAGGTGAAGTTCGTCGTCCAACATTATCATGGCCTCGTGAA
ATCCCGTTAGTAAAAGGTGGTAAACCTGACGTTGTACAAATTGTTAGGAATTATAATGCTTATCTACGTGCA
AGTGATGATTTACCAAAAATGTTTATTGAATCGGACCCAGGATTCTTTTCCAATGCTATTGTTGAAGGTGCC
AAGAAGTTTCCTAATACTGAATTTGTCAAAGTAAAAGGTCTTCATTTTTCGCAAGAAGATGCACCTGATGAA
ATGGGAAAATATATCAAATCGTTCGTTGAGCGAGTTCTCAAAAATGAACAATAATTCTAGCCCTGAATAAGT
GATAATAAGCGGATGAATGGCAGAAATTCGTCGAAGCGCAATAAAATATCTTTATTTTCATTACATCTGTGT
GTTGGTTTTTTGTGTGAATCGATAGCGATAAGGATCGGAAGATCCGCGTATGGTGCACTCTCAGTACAATCT
GCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTC
TGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGT
CATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAA

TGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAA
TACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGA
GTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTC
ACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGG
ATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAG
TTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATT
CTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT
TATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGA
AGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGA
ATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTAT
TAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAG
GACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGT
CTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGA
GTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAAC
TGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGG
TGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACC
CCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAA
AACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCT
TCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTG
TAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTC
TTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCA
CACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA
CGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGG
AGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGAT
TTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGG
CCTTTTGCTGGCCTTTTGCTCACATGGCTCGACAGATCT

>pBIND AIRE-G228W

TCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCA
TACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGGCATTG
ATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGT
TACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGAC
GTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGC
CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCC
CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT
CGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATT
TCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG
TCGTAACAACTGCGATCGCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA
AGCAGAGCTCGTTTAGTGAACCGTCAGATCACTAGAAGCTTTATTGCGGTAGTTTATCACAGTTAAATTGCT
AACGCAGTCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCTGCAGTGACTCTCTTAAGGTAGCCTTGCAG
AAGTTGGTCGTGAGGCACTGGGCAGGTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACT
GGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTTGC
CTTTCTCTCCACAGGTGTCCACTCCCAGTTCAATTACAGCTCTTAAGGCTAGAGTACTTAATACGACTCACT
ATAGGCTAGCCAGCTTGAAGCAAGCCTCCTGAAAGATGAAGCTACTGTCTTCTATCGAACAAGCATGCGATA
TTTGCCGACTTAAAAAGCTCAAGTGCTCCAAAGAAAAACCGAAGTGCGCCAAGTGTCTGAAGAACAACTGGG
AGTGTCGCTACTCTCCCAAAACCAAAAGGTCTCCGCTGACTAGGGCACATCTGACAGAAGTGGAATCAAGGC
TAGAAAGACTGGAACAGCTATTTCTACTGATTTTTCCTCGAGAAGACCTTGACATGATTTTGAAAATGGATT
CTTTACAGGATATAAAAGCATTGTTAACAGGATTATTTGTACAAGATAATGTGAATAAAGATGCCGTCACAG
ATAGATTGGCTTCAGTGGAGACTGATATGCCTCTAACATTGAGACAGCATAGAATAAGTGCGACATCATCAT
CGGAAGAGAGTAGTAACAAAGGTCAAAGACAGTTGACTGTATCGCCGGAATTCCCGGGGATCCGTCGACTTG
ACGCGTCTATGGCGACGGACGCGGCGCTACGCCGGCTTCTGAGGCTGCACCGCACGGAGATCGCGGTGGCCG
TGGACAGCGCCTTCCCACTGCTGCACGCGCTGGCTGACCACGACGTGGTCCCCGAGGACAAGTTTCAGGAGA
CGCTTCATCTGAAGGAAAAGGAGGGCTGCCCCCAGGCCTTCCACGCCCTCCTGTCCTGGCTGCTGACCCAGG
ACTCCACAGCCATCCTGGACTTCTGGAGGGTGCTGTTCAAGGACTACAACCTGGAGCGCTATGGCCGGCTGC
AGCCCATCCTGGACAGCTTCCCCAAAGATGTGGACCTCAGCCAGCCCCGGAAGGGGAGGAAGCCCCCGGCCG
TCCCCAAGGCTTTGGTACCGCCACCCAGACTCCCCACCAAGAGGAAGGCCTCAGAAGAGGCTCGAGCTGCCG

CGCCAGCAGCCCTGACTCCAAGGGGCACCGCCAGCCCAGGCTCTCAACTGAAGGCCAAGCCCCCCAAGAAGC
CGGAGAGCAGCGCAGAGCAGCAGCGCCTTCCACTCGGGAACGGGATTCAGACCATGTCAGCTTCAGTCCAGA
GAGCTGTGGCCATGTCCTCCGGGGACGTCCCGGGAGCCCGAGGGGCCGTGGAGGGGATCCTCATCCAGCAGG
TGTTTGAGTCAGGCGGCTCCAAGAAGTGCATCCAGGTTGGCTTGGGAGTTCTACACTCCCAGCAAGTTCGAAG
ACTCCGGCAGTGGGAAGAACAAGGCCCGCAGCAGCAGTGGCCCGAAGCCTCTGGTTCGAGCCAAGGGAGCCC
AGGGCGCTGCCCCCGGTGGAGGTGAGGCTAGGCTGGGCCAGCAGGGCAGCGTTCCCGCCCCTCTGGCCCTCC
CCAGTGACCCCCAGCTCCACCAGAAGAATGAGGACGAGTGTGCCGTGTGTCGGGACGGCGGGGAGCTCATCT
GCTGTGACGGCTGCCCTCGGGCCTTCCACCTGGCCTGCCTGTCCCCTCCGCTCCGGGAGATCCCCAGTGGGA
CCTGGAGGTGCTCCAGCTGCCTGCAGGCAACAGTCCAGGAGGTGCAGCCCCGGGCAGAGGAGCCCCGGCCCC
AGGAGCCACCCGTGGAGACCCCGCTCCCCCCGGGGCTTAGGTCGGCGGGAGAGGAGGTAAGAGGTCCACCTG
GGGAACCCCTAGCCGGCATGGACACGACTCTTGTCTACAAGCACCTGCCGGCTCCGCCTTCTGCAGCCCCGC
TGCCAGGGCTGGACTCCTCGGCCCTGCACCCCCTACTGTGTGTGGGTCCTGAGGGTCAGCAGAACCTGGCTC
CTGGTGCGCGTTGCGGGGTGTGCGGAGATGGTACGGACGTGCTGCGGTGTACTCACTGCGCCGCTGCCTTCC
ACTGGCGCTGCCACTTCCCAGCCGGCACCTCCCGGCCCGGGACGGGCCTGCGCTGCAGATCCTGCTCAGGAG
ACGTGACCCCAGCCCCTGTGGAGGGGGTGCTGGCCCCCAGCCCCGCCCGCCTGGCCCCTGGGCCTGCCAAGG
ATGACACTGCCAGTCACGAGCCCGCTCTGCACAGGGATGACCTGGAGTCCCTTCTGAGCGAGCACACCTTCG
ATGGCATCCTGCAGTGGGCCATCCAGAGCATGGCCCGTCCGGCGGCCCCCTTCCCCTCCGGGGGTGGAGGCT
CTGAGCAGAAACTCATCTCAGAAGAGGATCTGTAAGCGGCCGCAGGTACCTGAATAACTAAGGCCGCTTCCC
TTTAGTGAGGGTTAATGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAA
TGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCA
ATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTT
AAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCCGATAAGGATCGATCCGGGCTGGCGTAATAGCGAAGA
GGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCA
TTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCT
TTCGCTTTCTTCCCTTCCTTTCACGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCT
TTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGT
GGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTG
TTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCG
GCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACA
ATTTCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATACGCGGATCTTCCGTAC
CTTCTGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGC
AGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGC
AGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCC
GCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGC
CGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAA
AAAGCTTGATTCTTCTGACACAACAGTCTCGAACTTAAGGCTAGAATTCTGGCTACAGGTAAGCGCCCCTAA
AATCCCTTTGGGCACAATGTGTCCTGAGGGGAGAGGCAGCGACCTGTAGATGGGACGGGGGCACTAACCCTC
AGGTTTGGGGCTTCTGAATGTGAGTATCGCCATGTAAGCCCAGTATTTGGCCAATCTCAGAAAGCTCCTGGT
CCCTGGAGGGATGGAGAGAGAAAAACAAACAGCTCCTGGAGCAGGGAGAGTGCTGGCCTCTTGCTCTCCGGC
TCCCTCTGTTGCCCTCTGGTTTCTCCCCAGGCTCCCGGACGTCCTCTAGCCACCATGACTTCGAAAGTTTAT
GATCCAGAACAAAGGAAACGGATGATAACTGGTCCGCAGTGGTGGGCCAGATGTAAACAAATGAATGTTCTT
GATTCATTTATTAATTATTATGATTCAGAAAAACATGCAGAAAATGCTGTTATTTTTTTACATGGTAACGCG
GCCTCTTCTTATTTATGGCGACATGTTGTGCCACATATTGAGCCAGTAGCGCGGTGTATTATACCAGACCTT
ATTGGTATGGGCAAATCAGGCAAATCTGGTAATGGTTCTTATAGGTTACTTGATCATTACAAATATCTTACT
GCATGGTTTGAACTTCTTAATTTACCAAAGAAGATCATTTTTGTCGGCCATGATTGGGGTGCTTGTTTGGCA
TTTCATTATAGCTATGAGCATCAAGATAAGATCAAAGCAATAGTTCACGCTGAAAGTGTAGTAGATGTGATT
GAATCATGGGATGAATGGCCTGATATTGAAGAAGATATTGCGTTGATCAAATCTGAAGAAGGAGAAAAAATG
GTTTTGGAGAATAACTTCTTCGTGGAAACCATGTTGCCATCAAAAATCATGAGAAAGTTAGAACCAGAAGAA
TTTGCAGCATATCTTGAACCATTCAAAGAGAAAGGTGAAGTTCGTCGTCCAACATTATCATGGCCTCGTGAA
ATCCCGTTAGTAAAAGGTGGTAAACCTGACGTTGTACAAATTGTTAGGAATTATAATGCTTATCTACGTGCA
AGTGATGATTTACCAAAAATGTTTATTGAATCGGACCCAGGATTCTTTTCCAATGCTATTGTTGAAGGTGCC
AAGAAGTTTCCTAATACTGAATTTGTCAAAGTAAAAGGTCTTCATTTTTCGCAAGAAGATGCACCTGATGAA
ATGGGAAAATATATCAAATCGTTCGTTGAGCGAGTTCTCAAAAATGAACAATAATTCTAGCCCTGAATAAGT
GATAATAAGCGGATGAATGGCAGAAATTCGTCGAAGCGCAATAAAATATCTTTATTTTCATTACATCTGTGT
GTTGGTTTTTTGTGTGAATCGATAGCGATAAGGATCGGAAGATCCGCGTATGGTGCACTCTCAGTACAATCT
GCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTC
TGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGT
CATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAA
TGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAA
TACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGA

GTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTC
ACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGG
ATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAG
TTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATT
CTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT
TATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGA
AGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGA
ATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTAT
TAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAG
GACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGT
CTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGA
GTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAAC
TGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGG
TGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACC
CCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAA
AACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCT
TCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTG
TAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTC
TTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCA
CACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA
CGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGG
AGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGAT
TTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGG
CCTTTTGCTGGCCTTTTGCTCACATGGCTCGACAGATCT

7.11 In silico sequence of the precursor p63 cDNA clone

>pCMV3-HA p63

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGAGTACA
TTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTG
GCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGA
CTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATA
TGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCT
TACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAA
TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCCTCACTC
TCTTCCGCATCGCTGTCTGCGAGGGCCAGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCC
AGTACTCTTGGATCGGAAACCCGTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACCTGAGCGAGTCCGCA
TCGACCGGATCGGAAAACCTCTCGAGAAAGGCGTCTAACCAGTCACAGTCGCAAGGTAGGCTGAGCACCGTG
GCGGGCGGCAGCGGGTGGCGGTCGGGGTTGTTTCTGGCGGAGGTGCTGCTGATGATGTAATTAAAGTAGGCG
GTCTTGAGACGGCGGATGGTCGAGGTGAGGTGTGGCAGGCTTGAGATCCAGCTGTTGGGGTGAGTACTCCCT
CTCAAAAGCGGGCATTACTTCTGCGCTAAGATTGTCAGTTTCCAAAAACGAGGAGGATTTGATATTCACCTG
GCCCGATCTGGCCATACACTTGAGTGACAATGACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCA
GGTCCAAGTTTAAACTTTAATACGACTCACTATAGGGGCCGCCACCAAGCTT

Primers to clone Fp63 in pACT

ATGAATTTTGAAACTTCACGGTGTGCCACCCTACAGTACTGCCCTGACCCTTACATCCAGCGTTTCGTAGAA
ACCCCAGCTCATTTCTCTTGGAAAGAAAGTTATTACCGATCCACCATGTCCCAGAGCACACAGACAAATGAA
TTCCTCAGTCCAGAGGTTTTC

Primers to clone Tp63 in pACT

ATGCAGCATATCTGGGATTTTCTGGAACAGCCTATATGTTCAGTTCAGCCCATTGACTTGAACTTTGTGGAT
GAACCATCAGAAGATGGTGCGACAAACAAGATTGAGATTAGCATGGACTGTATCCGCATGCAGGACTCGGAC
CTGAGTGACCCCATGTGGCCACAGTACACGAACCTGGGGCTCCTGAACAGCATGGACCAGCAGATTCAGAAC

GGCTCCTCGTCCACCAGTCCCTATAACACAGACCACGCGCAGAACAGCGTCACGGCGCCCTCGCCCTACGCA
CAGCCCAGCTCCACCTTCGATGCTCTCTCTCCATCACCCGCCATCCCCTCCAACACCGACTACCCAGGCCCG
CACAGTTTCGACGTGTCCTTCCAGCAGTCGAGCACCGCCAAGTCGGCCACCTGGACGTATTCCACTGAACTG
AAGAAACTCTACTGCCAAATTGCAAAGACATGCCCCATCCAGATCAAGGTGATGACCCCACCTCCTCAGGGA
GCTGTTATCCGCGCCATGCCTGTCTACAAAAAAGCTGAGCACGTCACGGAGGTGGTGAAGCGGTGCCCCAAC
CATGAGCTGAGCCGTGAATTCAACGAGGGACAGATTGCCCCTCCTAGTCATTTGATTCGAGTAGAGGGGAAC
AGCCATGCCCAGTATGTAGAAGATCCCATCACAGGAAGACAGAGTGTGCTGGTACCTTATGAGCCACCCCAG
GTTGGCACTGAATTCACGACAGTCTTGTACAATTTCATGTGTAACAGCAGTTGTGTTGGAGGGATGAACCGC
CGTCCAATTTTAATCATTGTTACTCTGGAAACCAGAGATGGGCAAGTCCTGGGCCGACGCTGCTTTGAGGCC
CGGATCTGTGCTTGCCCAGGAAGAGACAGGAAGGCGGATGAAGATAGCATCAGAAAGCAGCAAGTTTCGGAC
AGTACAAAGAACGGTGATGGTACGAAGCGCCCGTTTCGTCAGAACACACATGGTATCCAGATGACATCCATC
AAGAAACGAAGATCCCCAGATGATGAACTGTTATACTTACCAGTGAGGGGCCGTGAGACTTATGAAATGCTG
TTGAAGATCAAAGAGTCCCTGGAACTCATGCAGTACCTTCCTCAGCACACAATTGAAACGTACAGGCAACAG
CAACAGCAGCAGCACCAGCACTTACTTCAGAAACAGACCTCAATACAGTCTCCATCTTCATATGGTAACAGC
TCCCCACCTCTGAACAAAATGAACAGCATGAACAAGCTGCCTTCTGTGAGCCAGCTTATCAACCCTCAGCAG
CGCAACGCCCTCACTCCTACAACCATTCCTGATGGCATGGGAGCCAACATTCCCATGATGGGCACCCACATG
CCAATGGCTGGAGACATGAATGGACTCAGCCCCACCCAGGCACTCCCTCCCCCACTCTCCATGCCATCCACC
TCCCACTGCACACCCCCACCTCCGTATCCCACAGATTGCAGCATTGTCAGTTTCTTAGCGAGGTTGGGCTGT
TCATCATGTCTGGACTATTTCACGACCCAGGGGCTGACCACCATCTATCAGATTGAGCATTACTCCATGGAT
GATCTGGCAAGTCTGAAAATCCCTGAGCAATTTCGACATGCGATCTGGAAGGGCATCCTGGACCACCGGCAG
CTCCACGAATTCTCCTCCCCTTCTCATCTCCTGCGGACCCCAAGCAGTGCCTCTACAGTCAGTGTGGGCTCC
AGTGAGACCCGGGGTGAGCGTGTTATTGATGCTGTGCGATTCACCCTCCGCCAGACCATCTCTTTCCCACCC
CGAGATGAGTGGAATGACTTCAACTTTGACATGGATGCTCGCCGCAATAAGCAACAGCGCATCAAAGAGGAG
GGGGAGGGGGGTGGAGGCTCTTATCCTTACGACGTGCCTGACTACGCCTAAGCGGCCGCCGAATTCGGGCCC
GTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTG
CCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGT
CTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAAT
AGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGG
TATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA
CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCAGGCTTTCCC
CGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAA
CTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAG
TCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTT
GATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCG
AATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGC
AAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGC
AAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGC
CCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTG
CCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCTCGGGAG
CTTGTATATCCATTTTCGGATCTGATCAGCACGTGATGAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGA
AGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTT
TCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATC
GTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCG
AGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGC
CCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGT
TCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATC
CCCATGTGTATCACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGC
TGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCC
TGACGGACAATGGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGG
TCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGC
ATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGA
GCTTGGTTGACGGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAG
CCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTAC
TCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTGCTACGAGATTTCG
ATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCTGGCTGGATGATCCTCC
AGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAAT
AAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAAC
TCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGC
TGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAG
CCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAA
ACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTT
CCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGG
CGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGG
CCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAA

ATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCT
CCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCG
TGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAA
GACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTA
CAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGA
AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTT
TTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGG
GGTCTGAGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAAT
TCTTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTT
TGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTAT
CGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCA
AGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACT
TGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGAT
TGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGG
CGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCT
GTTTTCCCAGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGA
AGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTG
CCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCG
ACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTAGAGCAA
GACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTT
CATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAATCCGCGCACATTT
CCCCGAAAAGTGCCACCTGACGTC

7.12 In silico sequences of the full-length and truncated p63 cloned in pACT

>pACT Fp63

TCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCA
TACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGGCATTG
ATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGT
TACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGAC
GTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGC
CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCC
CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT
CGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATT
TCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG
TCGTAACAACTGCGATCGCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA
AGCAGAGCTCGTTTAGTGAACCGTCAGATCACTAGAAGCTTTATTGCGGTAGTTTATCACAGTTAAATTGCT
AACGCAGTCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCTGCAGTGACTCTCTTAAGGTAGCCTTGCAG
AAGTTGGTCGTGAGGCACTGGGCAGGTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACT
GGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTTGC
CTTTCTCTCCACAGGTGTCCACTCCCAGTTCAATTACAGCTCTTAAGGCTAGAGTACTTAATACGACTCACT
ATAGGCTAGCCAGCTTGAAGCAAGCCTCCTGAAAGATGAAGCTACTGTCTTCTATCGAACAAGCATGCCCAA
AAAAGAAGAGAAAGGTAGATGAATTCCCGGGGATCTCGACGGCCCCCCCGACCGATGTCAGCCTGGGGGACG
AGCTCCACTTAGACGGCGAGGACGTGGCGATGGCGCATGCCGACGCGCTAGACGATTTCGATCTGGACATGT
TGGGGGACGGGGATTCCCCGGGTCCGGGATCGCCAGGGATCCCTATGAATTTTGAAACTTCACGGTGTGCCA
CCCTACAGTACTGCCCTGACCCTTACATCCAGCGTTTCGTAGAAACCCCAGCTCATTTCTCTTGGAAAGAAA
GTTATTACCGATCCACCATGTCCCAGAGCACACAGACAAATGAATTCCTCAGTCCAGAGGTTTTCATGCAGC
ATATCTGGGATTTTCTGGAACAGCCTATATGTTCAGTTCAGCCCATTGACTTGAACTTTGTGGATGAACCAT
CAGAAGATGGTGCGACAAACAAGATTGAGATTAGCATGGACTGTATCCGCATGCAGGACTCGGACCTGAGTG
ACCCCATGTGGCCACAGTACACGAACCTGGGGCTCCTGAACAGCATGGACCAGCAGATTCAGAACGGCTCCT
CGTCCACCAGTCCCTATAACACAGACCACGCGCAGAACAGCGTCACGGCGCCCTCGCCCTACGCACAGCCCA
GCTCCACCTTCGATGCTCTCTCTCCATCACCCGCCATCCCCTCCAACACCGACTACCCAGGCCCGCACAGTT
TCGACGTGTCCTTCCAGCAGTCGAGCACCGCCAAGTCGGCCACCTGGACGTATTCCACTGAACTGAAGAAAC
TCTACTGCCAAATTGCAAAGACATGCCCCATCCAGATCAAGGTGATGACCCCACCTCCTCAGGGAGCTGTTA
TCCGCGCCATGCCTGTCTACAAAAAAGCTGAGCACGTCACGGAGGTGGTGAAGCGGTGCCCCAACCATGAGC
TGAGCCGTGAATTCAACGAGGGACAGATTGCCCCTCCTAGTCATTTGATTCGAGTAGAGGGGAACAGCCATG
CCCAGTATGTAGAAGATCCCATCACAGGAAGACAGAGTGTGCTGGTACCTTATGAGCCACCCCAGGTTGGCA

CTGAATTCACGACAGTCTTGTACAATTTCATGTGTAACAGCAGTTGTGTTGGAGGGATGAACCGCCGTCCAA
TTTTAATCATTGTTACTCTGGAAACCAGAGATGGGCAAGTCCTGGGCCGACGCTGCTTTGAGGCCCGGATCT
GTGCTTGCCCAGGAAGAGACAGGAAGGCGGATGAAGATAGCATCAGAAAGCAGCAAGTTTCGGACAGTACAA
AGAACGGTGATGGTACGAAGCGCCCGTTTCGTCAGAACACACATGGTATCCAGATGACATCCATCAAGAAAC
GAAGATCCCCAGATGATGAACTGTTATACTTACCAGTGAGGGGCCGTGAGACTTATGAAATGCTGTTGAAGA
TCAAAGAGTCCCTGGAACTCATGCAGTACCTTCCTCAGCACACAATTGAAACGTACAGGCAACAGCAACAGC
AGCAGCACCAGCACTTACTTCAGAAACAGACCTCAATACAGTCTCCATCTTCATATGGTAACAGCTCCCCAC
CTCTGAACAAAATGAACAGCATGAACAAGCTGCCTTCTGTGAGCCAGCTTATCAACCCTCAGCAGCGCAACG
CCCTCACTCCTACAACCATTCCTGATGGCATGGGAGCCAACATTCCCATGATGGGCACCCACATGCCAATGG
CTGGAGACATGAATGGACTCAGCCCCACCCAGGCACTCCCTCCCCCACTCTCCATGCCATCCACCTCCCACT
GCACACCCCCACCTCCGTATCCCACAGATTGCAGCATTGTCAGTTTCTTAGCGAGGTTGGGCTGTTCATCAT
GTCTGGACTATTTCACGACCCAGGGGCTGACCACCATCTATCAGATTGAGCATTACTCCATGGATGATCTGG
CAAGTCTGAAAATCCCTGAGCAATTTCGACATGCGATCTGGAAGGGCATCCTGGACCACCGGCAGCTCCACG
AATTCTCCTCCCCTTCTCATCTCCTGCGGACCCCAAGCAGTGCCTCTACAGTCAGTGTGGGCTCCAGTGAGA
CCCGGGGGTGAGCGTGTTATTGATGCTGTGCGATTCACCCTCCGCCAGACCATCTCTTTCCCACCCCGAGATG
AGTGGAATGACTTCAACTTTGACATGGATGCTCGCCGCAATAAGCAACAGCGCATCAAAGAGGAGGGGGAGT
GAGCGGCCGCAGGTACCTGAATAACTAAGGCCGCTTCCCTTTAGTGAGGGTTAATGCTTCGAGCAGACATGA
TAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTT
GTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATT
TTATGTTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAAAA
TCCGATAAGGATCGATTCCGGAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGG
GTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCC
CTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGAT
TTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCT
GATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAA
CAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAA
AAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTCGCCTGTGT
ACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCA
GCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCA
GCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATC
CCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAG
GCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGC
AAAAAGCTTGATTCTTCTGACACAACAGTCTCGAACTTAAGGCTAGAGCCACCATGATTGAACAAGATGGAT
TGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCT
GCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCG
GTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAG
CTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCC
TGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTG
ATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCG
GTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCA
AGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGG
AAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGT
TGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCG
CCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTT
CGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGATGGCCGCAATAAAATATCTTTATTTTCATTAC
ATCTGTGTGTTGGTTTTTTGTGTGAAGATCCGCGTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCA
TAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCC
GCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGC
GCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACG
TCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATG
TATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAA
CATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTG
GTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGT
AAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC
GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTG
GTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCC
ATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCT
TTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCA
AACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA
CTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGC

TCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATT
GCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATG
GATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTT
TACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTT
GATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATC
AAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCA
GCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAG
ATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACA
TACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGAC
TCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTG
GAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGG
AGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGA
AACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCG
TCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCT
TTTGCTCACATGGCTCGACAGATCT

>pACT Tp63

TCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCA
TACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGGCATTG
ATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGT
TACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGAC
GTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGC
CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCC
CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT
CGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATT
TCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG
TCGTAACAACTGCGATCGCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA
AGCAGAGCTCGTTTAGTGAACCGTCAGATCACTAGAAGCTTTATTGCGGTAGTTTATCACAGTTAAATTGCT
AACGCAGTCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCTGCAGTGACTCTCTTAAGGTAGCCTTGCAG
AAGTTGGTCGTGAGGCACTGGGCAGGTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACT
GGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTTGC
CTTTCTCTCCACAGGTGTCCACTCCCAGTTCAATTACAGCTCTTAAGGCTAGAGTACTTAATACGACTCACT
ATAGGCTAGCCAGCTTGAAGCAAGCCTCCTGAAAGATGAAGCTACTGTCTTCTATCGAACAAGCATGCCCAA
AAAAGAAGAGAAAGGTAGATGAATTCCCGGGGATCTCGACGGCCCCCCCGACCGATGTCAGCCTGGGGGACG
AGCTCCACTTAGACGGCGAGGACGTGGCGATGGCGCATGCCGACGCGCTAGACGATTTCGATCTGGACATGT
TGGGGGACGGGGATTCCCCGGGTCCGGGATCGCCAGGGATCCCAATGCAGCATATCTGGGATTTTCTGGAAC
AGCCTATATGTTCAGTTCAGCCCATTGACTTGAACTTTGTGGATGAACCATCAGAAGATGGTGCGACAAACA
AGATTGAGATTAGCATGGACTGTATCCGCATGCAGGACTCGGACCTGAGTGACCCCATGTGGCCACAGTACA
CGAACCTGGGGCTCCTGAACAGCATGGACCAGCAGATTCAGAACGGCTCCTCGTCCACCAGTCCCTATAACA
CAGACCACGCGCAGAACAGCGTCACGGCGCCCTCGCCCTACGCACAGCCCAGCTCCACCTTCGATGCTCTCT
CTCCATCACCCGCCATCCCCTCCAACACCGACTACCCAGGCCCGCACAGTTTCGACGTGTCCTTCCAGCAGT
CGAGCACCGCCAAGTCGGCCACCTGGACGTATTCCACTGAACTGAAGAAACTCTACTGCCAAATTGCAAAGA
CATGCCCCATCCAGATCAAGGTGATGACCCCACCTCCTCAGGGAGCTGTTATCCGCGCCATGCCTGTCTACA
AAAAAGCTGAGCACGTCACGGAGGTGGTGAAGCGGTGCCCCAACCATGAGCTGAGCCGTGAATTCAACGAGG
GACAGATTGCCCCTCCTAGTCATTTGATTCGAGTAGAGGGGAACAGCCATGCCCAGTATGTAGAAGATCCCA
TCACAGGAAGACAGAGTGTGCTGGTACCTTATGAGCCACCCCAGGTTGGCACTGAATTCACGACAGTCTTGT
ACAATTTCATGTGTAACAGCAGTTGTGTTGGAGGGATGAACCGCCGTCCAATTTTAATCATTGTTACTCTGG
AAACCAGAGATGGGCAAGTCCTGGGCCGACGCTGCTTTGAGGCCCGGATCTGTGCTTGCCCAGGAAGAGACA
GGAAGGCGGATGAAGATAGCATCAGAAAGCAGCAAGTTTCGGACAGTACAAAGAACGGTGATGGTACGAAGC
GCCCGTTTCGTCAGAACACACATGGTATCCAGATGACATCCATCAAGAAACGAAGATCCCCAGATGATGAAC
TGTTATACTTACCAGTGAGGGGCCGTGAGACTTATGAAATGCTGTTGAAGATCAAAGAGTCCCTGGAACTCA
TGCAGTACCTTCCTCAGCACACAATTGAAACGTACAGGCAACAGCAACAGCAGCAGCACCAGCACTTACTTC
AGAAACAGACCTCAATACAGTCTCCATCTTCATATGGTAACAGCTCCCCACCTCTGAACAAAATGAACAGCA
TGAACAAGCTGCCTTCTGTGAGCCAGCTTATCAACCCTCAGCAGCGCAACGCCCTCACTCCTACAACCATTC
CTGATGGCATGGGAGCCAACATTCCCATGATGGGCACCCACATGCCAATGGCTGGAGACATGAATGGACTCA
GCCCCACCCAGGCACTCCCTCCCCCACTCTCCATGCCATCCACCTCCCACTGCACACCCCCACCTCCGTATC
CCACAGATTGCAGCATTGTCAGTTTCTTAGCGAGGTTGGGCTGTTCATCATGTCTGGACTATTTCACGACCC
AGGGGCTGACCACCATCTATCAGATTGAGCATTACTCCATGGATGATCTGGCAAGTCTGAAAATCCCTGAGC

AATTTCGACATGCGATCTGGAAGGGCATCCTGGACCACCGGCAGCTCCACGAATTCTCCTCCCCTTCTCATC
TCCTGCGGACCCCAAGCAGTGCCTCTACAGTCAGTGTGGGCTCCAGTGAGACCCGGGGTGAGCGTGTTATTG
ATGCTGTGCGATTCACCCTCCGCCAGACCATCTCTTTCCCACCCCGAGATGAGTGGAATGACTTCAACTTTG
ACATGGATGCTCGCCGCAATAAGCAACAGCGCATCAAAGAGGAGGGGGAGTGAGCGGCCGCAGGTACCTGAA
TAACTAAGGCCGCTTCCCTTTAGTGAGGGTTAATGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTG
GACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG
TAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGG
AGATGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCCGATAAGGATCGATTCCGG
AGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCG
TGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCG
CCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCG
ACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTT
TGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGG
TCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAA
AATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTCGCCTGTGTACCTTCTGAGGCGGAAAGAAC
CAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGC
ATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGC
ATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGT
TCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCT
GAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTGATTCTTCTGAC
ACAACAGTCTCGAACTTAAGGCTAGAGCCACCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCG
CTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCC
GGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGG
ACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTG
AAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTG
CCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCG
ACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATC
TGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCG
AGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGAT
TCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTG
AAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCA
TCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGAC
GCCCAACCTGCCATCACGATGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGT
GTGAAGATCCGCGTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACAC
CCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACC
GTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTG
ATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGA
AATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAA
CCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATT
CCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAA
GATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGC
CCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGAC
GCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACA
GAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACT
GCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGAT
CATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACG
ATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAA
CAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGG
TTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGT
AAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATC
GCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATT
GATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATC
CCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCT
TTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGAT
CAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTA
GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTG
TTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGAT
AAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAA
CTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCG
GTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGT

CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGG
AAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGGCTCGACAG
ATCT

7.13 In silico sequences of ready-made cDNA clones

>pCMV3-HA BRD4

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGAGTACA
TTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTG
GCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGA
CTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATA
TGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCT
TACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAA
TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCCTCACTC
TCTTCCGCATCGCTGTCTGCGAGGGCCAGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCC
AGTACTCTTGGATCGGAAACCCGTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACCTGAGCGAGTCCGCA
TCGACCGGATCGGAAAACCTCTCGAGAAAGGCGTCTAACCAGTCACAGTCGCAAGGTAGGCTGAGCACCGTG
GCGGGCGGCAGCGGGTGGCGGTCGGGGTTGTTTCTGGCGGAGGTGCTGCTGATGATGTAATTAAAGTAGGCG
GTCTTGAGACGGCGGATGGTCGAGGTGAGGTGTGGCAGGCTTGAGATCCAGCTGTTGGGGTGAGTACTCCCT
CTCAAAAGCGGGCATTACTTCTGCGCTAAGATTGTCAGTTTCCAAAAACGAGGAGGATTTGATATTCACCTG
GCCCGATCTGGCCATACACTTGAGTGACAATGACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCA
GGTCCAAGTTTAAACTTTAATACGACTCACTATAGGGGCCGCCACCAAGCTTATGTCTGCGGAGAGCGGCCC
TGGGACGAGATTGAGAAATCTGCCAGTAATGGGGGATGGACTAGAAACTTCCCAAATGTCTACAACACAGGC
CCAGGCCCAACCCCAGCCAGCCAACGCAGCCAGCACCAACCCCCCGCCCCCAGAGACCTCCAACCCTAACAA
GCCCAAGAGGCAGACCAACCAACTGCAATACCTGCTCAGAGTGGTGCTCAAGACACTATGGAAACACCAGTT
TGCATGGCCTTTCCAGCAGCCTGTGGATGCCGTCAAGCTGAACCTCCCTGATTACTATAAGATCATTAAAAC
GCCTATGGATATGGGAACAATAAAGAAGCGCTTGGAAAACAACTATTACTGGAATGCTCAGGAATGTATCCA
GGACTTCAACACTATGTTTACAAATTGTTACATCTACAACAAGCCTGGAGATGACATAGTCTTAATGGCAGA
AGCTCTGGAAAAGCTCTTCTTGCAAAAAATAAATGAGCTACCCACAGAAGAAACCGAGATCATGATAGTCCA
GGCAAAAGGAAGAGGACGTGGGAGGAAAGAAACAGGGACAGCAAAACCTGGCGTTTCCACGGTACCAAACAC
AACTCAAGCATCGACTCCTCCGCAGACCCAGACCCCTCAGCCGAATCCTCCTCCTGTGCAGGCCACGCCTCA
CCCCTTCCCTGCCGTCACCCCGGACCTCATCGTCCAGACCCCTGTCATGACAGTGGTGCCTCCCCAGCCACT
GCAGACGCCCCCGCCAGTGCCCCCCCAGCCACAACCCCCACCCGCTCCAGCTCCCCAGCCCGTACAGAGCCA
CCCACCCATCATCGCGGCCACCCCACAGCCTGTGAAGACAAAGAAGGGAGTGAAGAGGAAAGCAGACACCAC
CACCCCCACCACCATTGACCCCATTCACGAGCCACCCTCGCTGCCCCCGGAGCCCAAGACCACCAAGCTGGG
CCAGCGGCGGGAGAGCAGCCGGCCTGTGAAACCTCCAAAGAAGGACGTGCCCGACTCTCAGCAGCACCCAGC
ACCAGAGAAGAGCAGCAAGGTCTCGGAGCAGCTCAAGTGCTGCAGCGGCATCCTCAAGGAGATGTTTGCCAA
GAAGCACGCCGCCTACGCCTGGCCCTTCTACAAGCCTGTGGACGTGGAGGCACTGGGCCTACACGACTACTG
TGACATCATCAAGCACCCCATGGACATGAGCACAATCAAGTCTAAACTGGAGGCCCGTGAGTACCGTGATGC
TCAGGAGTTTGGTGCTGACGTCCGATTGATGTTCTCCAACTGCTATAAGTACAACCCTCCTGACCATGAGGT
GGTGGCCATGGCCCGCAAGCTCCAGGATGTGTTCGAAATGCGCTTTGCCAAGATGCCGGACGAGCCTGAGGA
GCCAGTGGTGGCCGTGTCCTCCCCGGCAGTGCCCCCTCCCACCAAGGTTGTGGCCCCGCCCTCATCCAGCGA
CAGCAGCAGCGATAGCTCCTCGGACAGTGACAGTTCGACTGATGACTCTGAGGAGGAGCGAGCCCAGCGGCT
GGCTGAGCTCCAGGAGCAGCTCAAAGCCGTGCACGAGCAGCTTGCAGCCCTCTCTCAGCCCCAGCAGAACAA
ACCAAAGAAAAAGGAGAAAGACAAGAAGGAAAAGAAAAAAGAAAAGCACAAAAGGAAAGAGGAAGTGGAAGA
GAATAAAAAAAGCAAAGCCAAGGAACCTCCTCCTAAAAAGACGAAGAAAAATAATAGCAGCAACAGCAATGT
GAGCAAGAAGGAGCCAGCGCCCATGAAGAGCAAGCCCCCTCCCACGTATGAGTCGGAGGAAGAGGACAAGTG
CAAGCCTATGTCCTATGAGGAGAAGCGGCAGCTCAGCTTGGACATCAACAAGCTCCCCGGCGAGAAGCTGGG
CCGCGTGGTGCACATCATCCAGTCACGGGAGCCCTCCCTGAAGAATTCCAACCCCGACGAGATTGAAATCGA
CTTTGAGACCCTGAAGCCGTCCACACTGCGTGAGCTGGAGCGCTATGTCACCTCCTGTTTGCGGAAGAAAAG

GAAACCTCAAGCTGAGAAAGTTGATGTGATTGCCGGCTCCTCCAAGATGAAGGGCTTCTCGTCCTCAGAGTC
GGAGAGCTCCAGTGAGTCCAGCTCCTCTGACAGCGAAGACTCCGAAACAGGTCCTGCCGGGGGTGGAGGCTC
TTATCCTTACGACGTGCCTGACTACGCCTAAGCGGCCGCCGAATTCGGGCCCGTTTAAACCCGCTGATCAGC
CTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGG
TGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTAT
TCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC
GGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAG
CGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCC
CGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCAGGCTTTCCCCGTCAAGCTCTAAATCGGGG
GCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTC
ACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGG
ACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCC
GATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTG
TGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAG
TCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAG
TCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCG
CCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAG
TAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCTCGGGAGCTTGTATATCCATTTTCGGA
TCTGATCAGCACGTGATGAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTC
GACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGG
CGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTT
GCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATC
TCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTC
GCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAA
GGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTATCACTGGCAA
ACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCGAGGAC
TGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCCGCATA
ACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGG
AGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCG
CCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAATTTC
GATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGCGTACA
CAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGA
CGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTGCTACGAGATTTCGATTCCACCGCCGCCTTCTAT
GAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCTGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTG
GAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAAT
TTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCAT
GTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGT
TATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTG
AGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCAT
TAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGAC
TCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACA
GAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCC
GCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGG
TGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTT
CCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCA
CGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAG
CCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG
CCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAA
AGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAG
ATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGAGCGCGGAACCCCT
ATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATCGAG
CATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAA
TGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCG
TCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGT
GACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATT
ACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAA
TACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGC
ATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCAGGGATCGCAGT
GGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAG

CCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTC
TGGCGCATCGGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTT
ATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATG
GCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGACCAAAATCCCTTAAC
GTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAATCCGCGCACATTTCCCCGAAAAGTGCCACCTGA
CGTC

>pCMV3-HA PIAS1

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGAGTACA
TTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTG
GCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGA
CTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATA
TGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCT
TACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAA
TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCCTCACTC
TCTTCCGCATCGCTGTCTGCGAGGGCCAGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCC
AGTACTCTTGGATCGGAAACCCGTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACCTGAGCGAGTCCGCA
TCGACCGGATCGGAAAACCTCTCGAGAAAGGCGTCTAACCAGTCACAGTCGCAAGGTAGGCTGAGCACCGTG
GCGGGCGGCAGCGGGTGGCGGTCGGGGTTGTTTCTGGCGGAGGTGCTGCTGATGATGTAATTAAAGTAGGCG
GTCTTGAGACGGCGGATGGTCGAGGTGAGGTGTGGCAGGCTTGAGATCCAGCTGTTGGGGTGAGTACTCCCT
CTCAAAAGCGGGCATTACTTCTGCGCTAAGATTGTCAGTTTCCAAAAACGAGGAGGATTTGATATTCACCTG
GCCCGATCTGGCCATACACTTGAGTGACAATGACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCA
GGTCCAAGTTTAAACTTTAATACGACTCACTATAGGGGCCGCCACCAAGCTTATGGCGGACAGTGCGGAACT
AAAGCAAATGGTTATGAGCCTTAGAGTTTCTGAACTCCAAGTACTGTTGGGCTACGCCGGGAGAAACAAGCA
CGGACGCAAACACGAACTTCTCACAAAAGCCCTGCATTTGCTAAAGGCTGGCTGTAGTCCTGCTGTGCAAAT
GAAAATTAAGGAACTCTATAGGCGGCGGTTCCCACAGAAAATCATGACGCCTGCAGACTTGTCCATCCCCAA
CGTACATTCAAGTCCTATGCCAGCAACTTTGTCTCCATCTACCATTCCACAACTCACTTACGATGGTCACCC
TGCATCATCGCCATTACTCCCTGTTTCTCTTCTGGGACCTAAACATGAACTGGAACTCCCACATCTTACATC
AGCTCTTCACCCAGTCCATCCGGATATAAAACTTCAAAAATTACCATTTTATGATTTACTGGATGAACTGAT
AAAACCCACCAGTCTAGCATCAGACAACAGTCAGCGCTTTCGAGAAACCTGTTTTGCATTTGCCTTGACACC
ACAACAAGTGCAGCAAATCAGTAGTTCCATGGATATTTCTGGGACCAAATGTGACTTCACAGTACAGGTCCA
GTTAAGGTTTTGTTTATCAGAAACCAGTTGTCCACAAGAAGATCACTTCCCACCCAATCTTTGTGTGAAAGT
GAATACAAAACCTTGCAGCCTTCCAGGTTACCTTCCACCTACAAAAAATGGCGTGGAACCAAAGCGACCCAG
CCGACCAATTAATATCACCTCACTTGTCCGACTGTCCACAACAGTACCAAACACGATTGTTGTTTCTTGGAC
TGCAGAAATTGGAAGAAACTATTCCATGGCAGTATATCTTGTAAAACAGTTGTCCTCAACAGTTCTTCTTCA
GAGGTTACGAGCAAAGGGAATAAGGAATCCGGATCATTCTAGAGCTTTAATTAAAGAGAAGTTGACTGCGGA
TCCAGACAGTGAAATAGCTACAACCAGCCTAAGGGTTTCTCTACTATGTCCACTTGGTAAAATGCGGCTGAC
AATTCCGTGTCGGGCCCTTACATGTTCTCATCTACAATGTTTTGACGCAACTCTTTACATTCAGATGAATGA
GAAAAAACCAACCTGGGTTTGTCCTGTCTGTGATAAGAAGGCTCCATATGAACACCTTATTATTGATGGCTT
GTTTATGGAAATCCTAAAGTACTGTACAGACTGTGATGAAATACAATTTAAGGAGGATGGCACTTGGGCACC
GATGAGATCAAAAAAGGAAGTACAGGAAGTTTCTGCCTCTTACAATGGAGTCGATGGATGCTTGAGCTCCAC
ATTGGAGCATCAGGTAGCGTCTCACCACCAGTCCTCAAATAAAAACAAGAAAGTAGAAGTGATTGACCTAAC
CATAGACAGTTCATCTGATGAAGAGGAAGAAGAGCCATCTGCCAAGAGGACCTGTCCTTCCCTATCTCCCAC
ATCACCACTAAATAATAAAGGCATTTTAAGTCTTCCACATCAAGCATCTCCAGTATCCCGCACCCCAAGCCT
TCCTGCTGTAGACACAAGCTACATTAATACCTCCCTCATCCAAGACTATAGGCATCCTTTCCACATGACACC
CATGCCTTACGACTTACAAGGATTAGATTTCTTTCCTTTCTTATCAGGAGACAATCAGCATTACAACACCTC
CTTGCTTGCCGCTGCAGCAGCAGCAGTTTCAGATGATCAAGACCTCCTACACTCGTCTCGGTTTTTCCCGTA
TACCTCCTCACAGATGTTTCTTGATCAGTTAAGTGCAGGAGGCAGTACTTCTCTGCCAACCACCAATGGAAG
CAGTAGTGGCAGTAACAGCAGCCTGGTTTCTTCCAACAGCCTAAGGGAAAGCCATAGCCACACCGTCACAAA
CAGGAGCAGCACGGACACGGCATCCATCTTTGGCATCATACCAGACATTATTTCATTGGACGGGGGTGGAGG
CTCTTATCCTTACGACGTGCCTGACTACGCCTAAGCGGCCGCCGAATTCGGGCCCGTTTAAACCCGCTGATC
AGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGA
AGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTC

TATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGA
TGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTG
TAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGC
GCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCAGGCTTTCCCCGTCAAGCTCTAAATCG
GGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGG
TTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAG
TGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTT
GCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAAT
GTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAAT
TAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAAT
TAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCT
CCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAG
AAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCTCGGGAGCTTGTATATCCATTTTC
GGATCTGATCAGCACGTGATGAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAG
TTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGA
GGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCAC
TTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTGC
ATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGCCG
GTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCG
CAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTATCACTGG
CAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCGAG
GACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCCGC
ATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTC
TGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAGGA
TCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAAT
TTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGCGT
ACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAAC
CGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAGCACGTGCTACGAGATTTCGATTCCACCGCCGCCTTC
TATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCTGGCTGGATGATCCTCCAGCGCGGGGATCTCATG
CTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACA
AATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTAT
CATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAAT
TGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGA
GTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTG
CATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACT
GACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCC
ACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAG
GCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAG
AGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCT
GTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC
TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT
CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCA
CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG
TGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA
AAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAG
CAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGAGCGCGGAACC
CCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATC
GAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTG
TAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGAC
TCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATG
AGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCC
ATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACG
AAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAG
CGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCAGGGATCGC
AGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGT
CAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAA
CTCTGGCGCATCGGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCA
TTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAAT
ATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGACCAAAATCCCTT

AACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAATCCGCGCACATTTCCCCGAAAAGTGCCACC
TGACGTC

>pCMV3-HA DAXX

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAA
GGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGAGTACA
TTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTG
GCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGA
CTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATA
TGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCT
TACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG
GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAA
TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCCTCACTC
TCTTCCGCATCGCTGTCTGCGAGGGCCAGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCC
AGTACTCTTGGATCGGAAACCCGTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACCTGAGCGAGTCCGCA
TCGACCGGATCGGAAAACCTCTCGAGAAAGGCGTCTAACCAGTCACAGTCGCAAGGTAGGCTGAGCACCGTG
GCGGGCGGCAGCGGGTGGCGGTCGGGGTTGTTTCTGGCGGAGGTGCTGCTGATGATGTAATTAAAGTAGGCG
GTCTTGAGACGGCGGATGGTCGAGGTGAGGTGTGGCAGGCTTGAGATCCAGCTGTTGGGGTGAGTACTCCCT
CTCAAAAGCGGGCATTACTTCTGCGCTAAGATTGTCAGTTTCCAAAAACGAGGAGGATTTGATATTCACCTG
GCCCGATCTGGCCATACACTTGAGTGACAATGACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCA
GGTCCAAGTTTAAACTTTAATACGACTCACTATAGGGGCCGCCACCAAGCTTATGGCCACCGCTAACAGCAT
CATCGTGCTGGATGATGATGACGAAGATGAAGCAGCTGCTCAGCCAGGGCCCTCCCACCCACTCCCCAATGC
GGCCTCACCTGGGGCAGAAGCCCCTAGCTCCTCTGAGCCTCATGGGGCCAGAGGAAGCAGTAGTTCGGGCGG
CAAGAAATGCTACAAGCTGGAGAATGAGAAGCTGTTCGAAGAGTTCCTTGAACTTTGTAAGATGCAGACAGC
AGACCACCCTGAGGTGGTCCCATTCCTCTATAACCGGCAGCAACGTGCCCACTCTCTGTTTTTGGCCTCGGC
GGAGTTCTGCAACATCCTCTCTAGGGTCCTGTCTCGGGCCCGGAGCCGGCCAGCCAAGCTCTATGTCTACAT
CAATGAGCTCTGCACTGTTCTCAAGGCCCACTCAGCCAAAAAGAAGCTGAACTTGGCCCCTGCCGCCACCAC
CTCCAATGAGCCCTCTGGGAATAACCCTCCCACACACCTCTCCTTGGACCCCACAAATGCTGAAAACACTGC
CTCTCAGTCTCCAAGGACCCGTGGTTCCCGGCGGCAGATCCAGCGTTTGGAGCAGCTGCTGGCGCTCTATGT
GGCAGAGATCCGGCGGCTGCAGGAAAAGGAGTTGGATCTCTCAGAATTGGATGACCCAGACTCCGCATACCT
GCAGGAGGCACGGTTGAAGCGTAAGCTGATCCGCCTCTTTGGGCGACTATGTGAGCTGAAAGACTGCTCTTC
ACTGACCGGCCGTGTCATAGAGCAGCGCATCCCCTACCGTGGCACCCGCTACCCAGAGGTTAACAGGCGCAT
TGAGCGGCTCATCAACAAGCCAGGGCCTGATACCTTCCCTGACTATGGGGATGTGCTTCGGGCTGTAGAGAA
GGCAGCTGCCCGACACAGCCTTGGCCTCCCCCGACAGCAGCTCCAGCTCATGGCTCAGGATGCCTTCCGAGA
TGTGGGCATCAGGTTACAGGAGCGACGTCACCTCGATCTCATCTACAACTTTGGCTGCCACCTCACAGATGA
CTATAGGCCAGGCGTTGACCCTGCACTATCAGATCCTGTGTTGGCCCGGCGCCTTCGGGAAAACCGGAGTTT
GGCCATGAGTCGGCTGGATGAGGTCATCTCCAAATATGCAATGTTGCAAGACAAAAGTGAGGAGGGCGAGAG
AAAAAAGAGAAGAGCTCGGCTCCAAGGCACCTCTTCCCACTCTGCAGACACCCCCGAAGCCTCCTTGGATTC
TGGTGAGGGCCCTAGTGGAATGGCATCCCAGGGGTGCCCTTCTGCCTCCAGAGCTGAGACAGATGACGAAGA
CGATGAGGAGAGTGATGAGGAAGAGGAGGAGGAGGAGGAAGAAGAAGAGGAGGAGGCCACAGATTCTGAAGA
GGAGGAGGATCTGGAACAGATGCAGGAGGGTCAGGAGGATGATGAAGAGGAGGACGAAGAGGAAGAAGCAGC
AGCAGGTAAAGATGGAGACAAGAGCCCCATGTCCTCACTACAGATCTCCAATGAAAAGAACCTGGAACCTGG
CAAACAGATCAGCAGATCTTCAGGGGAGCAGCAAAACAAAGGACGCATAGTGTCACCATCGTTACTGTCAGA
AGAACCCCTGGCCCCCTCCAGCATAGATGCTGAAAGCAATGGAGAACAGCCTGAGGAGCTGACCCTGGAGGA
AGAAAGCCCTGTGTCTCAGCTCTTTGAGCTAGAGATTGAAGCTTTGCCCCTGGATACCCCTTCCTCTGTGGA
GACGGACATTTCCTCTTCCAGGAAGCAATCAGAGGAGCCCTTCACCACTGTCTTAGAGAATGGAGCAGGCAT
GGTCTCTTCTACTTCCTTCAATGGAGGCGTCTCTCCTCACAACTGGGGAGATTCTGGTCCCCCCTGCAAAAA
ATCTCGGAAGGAGAAGAAGCAAACAGGATCAGGGCCATTAGGAAACAGCTATGTGGAAAGGCAAAGGTCAGT
GCATGAGAAGAATGGGAAAAAGATATGTACCCTGCCCAGCCCACCTTCCCCCTTGGCTTCCTTGGCCCCAGT
TGCTGATTCCTCCACGAGGGTGGACTCTCCCAGCCATGGCCTGGTGACCAGCTCCCTCTGCATCCCTTCTCC
AGCCCGGCTGTCCCAAACCCCCCATTCACAGCCTCCTCGGCCTGGTACTTGCAAGACAAGTGTGGCCACACA
ATGCGATCCAGAAGAGATCATCGTGCTCTCAGACTCTGATGGGGGTGGAGGCTCTTATCCTTACGACGTGCC
TGACTACGCCTAAGCGGCCGCCGAATTCGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAG
TTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCT
TTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGG

GCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTC
TGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGC
GGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTT
CCCTTCCTTTCTCGCCACGTTCGCAGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCG
ATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCC
CTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGG
AACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTT
AAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGA
AAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGA
AAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCG
CCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATT
TTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTT
GGAGGCCTAGGCTTTTGCAAAAAGCTCTCGGGAGCTTGTATATCCATTTTCGGATCTGATCAGCACGTGATG
AAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTG
ATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGG
GTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCG
ATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGT
GTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCG
ATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACT
ACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTATCACTGGCAAACTGTGATGGACGACACC
GTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCAC
CTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGG
AGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGT
ATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTAT
ATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGGGCG
CAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGC
GCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCG
AGGGCAAAGGAATAGCACGTGCTACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGA
ATCGTTTTCCGGGACGCTGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCC
AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTT
TTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACC
TCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA
CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATT
GCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGC
GCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTT
CGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCA
GGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC
CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA
CTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACC
GGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGT
TCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTA
TCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC
AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACT
AGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGA
TCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAA
GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGAGCGCGGAACCCCTATTTGTTTATTTTTCTAA
ATACATTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATCGAGCATCAAATGAAACTGCAA
TTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACC
GAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACC
TATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGA
GAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATC
ACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAA
AGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACC
TGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCAGGGATCGCAGTGGTGAGTAACCATGCATC
ATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCAT
CTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCC
ATACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGC
ATCCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGT
ATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACT
GAGCGTCAGACCCCGTAGAAATCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

7.14 In silico sequence of pCMV GFP used to test transfection efficiency

CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAAT
GACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAAC
TGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATG
GCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGT
CATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGG
ATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAA
ATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG
AGCTGGTTTAGTGAACCGTCAGATCCGCTAGCGCTACCGGACTCAGATCTCGAGCTCAAGCTTCGAATTCTG
CAGTCGACGGTACCGCGGGCCCGGGATCCACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACC
GGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGC
GAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGG
CCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAC
GACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAAC
TACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGAC
TTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATG
GCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTG
AGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACC
GCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAGCGGCCGCACTCCTCAGGTGCAGGCTGCC
TATCAGAAGGTGGTGGCTGGTGTGGCCAATGCCCTGGCTCACAAATACCACTGAGATCTTTTTCCCTCTGCC
AAAAATTATGGGGACATCATGAAGCCCCTTGAGCATCTGACTTCTGGCTAATAAAGGAAATTTATTTTCATT
GCAATAGTGTGTTGGAATTTTTTGTGTCTCTCACTCGGAAGGACATATGGGAGGGCAAATCATTTAAAACAT
CAGAATGAGTATTTGGTTTAGAGTTTGGCAACATATGCCCATATGCTGGCTGCCATGAACAAAGGTTGGCTA
TAAAGAGGTCATCAGTATATGAAACAGCCCCCTGCTGTCCATTCCTTATTCCATAGAAAAGCCTTGACTTGA
GGTTAGATTTTTTTTATATTTTGTTTTGTGTTATTTTTTTCTTTAACATCCCTAAAATTTTCCTTACATGTT
TTACTAGCCAGATTTTTCCTCCTCTCCTGACTACTCCCAGTCATAGCTGTCCCTCTTCTCTTATGGAGATCC
CTCGACCTGCAGCCCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCAC
AATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCAC
ATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCGGATCCGCATCTCAAT
TAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCT
CCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAG
AAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTAACTTGTTTATTGCAGCTTATAATG
GTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTT
TGTCCAAACTCATCAATGTATCTTATCATGTCTGGATCCGCTGCATTAATGAATCGGCCAACGCGCGGGGAG
AGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCG
GCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAA
CATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCT
CCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAG
ATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCT
GTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTA
GGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAA
CTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAG
CAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAAC
AGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA
ACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCA
AGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGT
CATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAG
TATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCT
ATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGG
CCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGC
CGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGA
AGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTC
ACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCAT
GTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATC
ACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGG

TGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACG
GGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACT
CTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATC
TTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGC
GACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCT
CATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAA
AGTGCCACCTGGGTCGACATTGATTATTGACTAGTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCC
ATATATGGAGTTCCG

Green：cMyc tag

Grey : HA tag

Yellow：Restriction enzyme

Turquoise: primers

Red : Insert

7.15 Plasmid maps

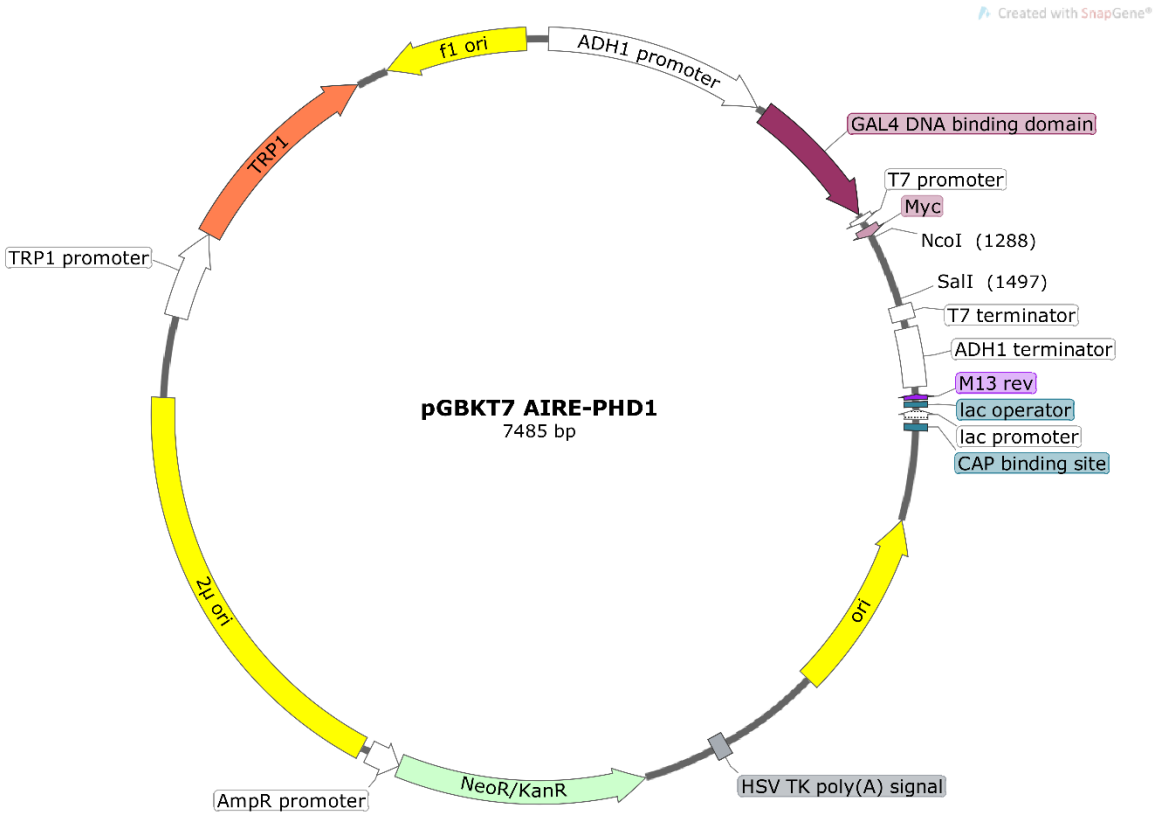


Figure 28 : pGBKT7 AIRE-PHD1 plasmid map.

AIRE-PHD1 was inserted in the multiple cloning region at the *Nco1* and *Sal1* restriction sites, is fused to the upstream c-myc and Gal4 DBD, and is the under the control of the ADH1 promoter. The Plasmid is resistance to kanamycin and is allowed to grow on Tryptophan-deprived medium in yeast. The origin of replication is shown in yellow. This plasmid is generated using SnapGene 5.0.8 software.

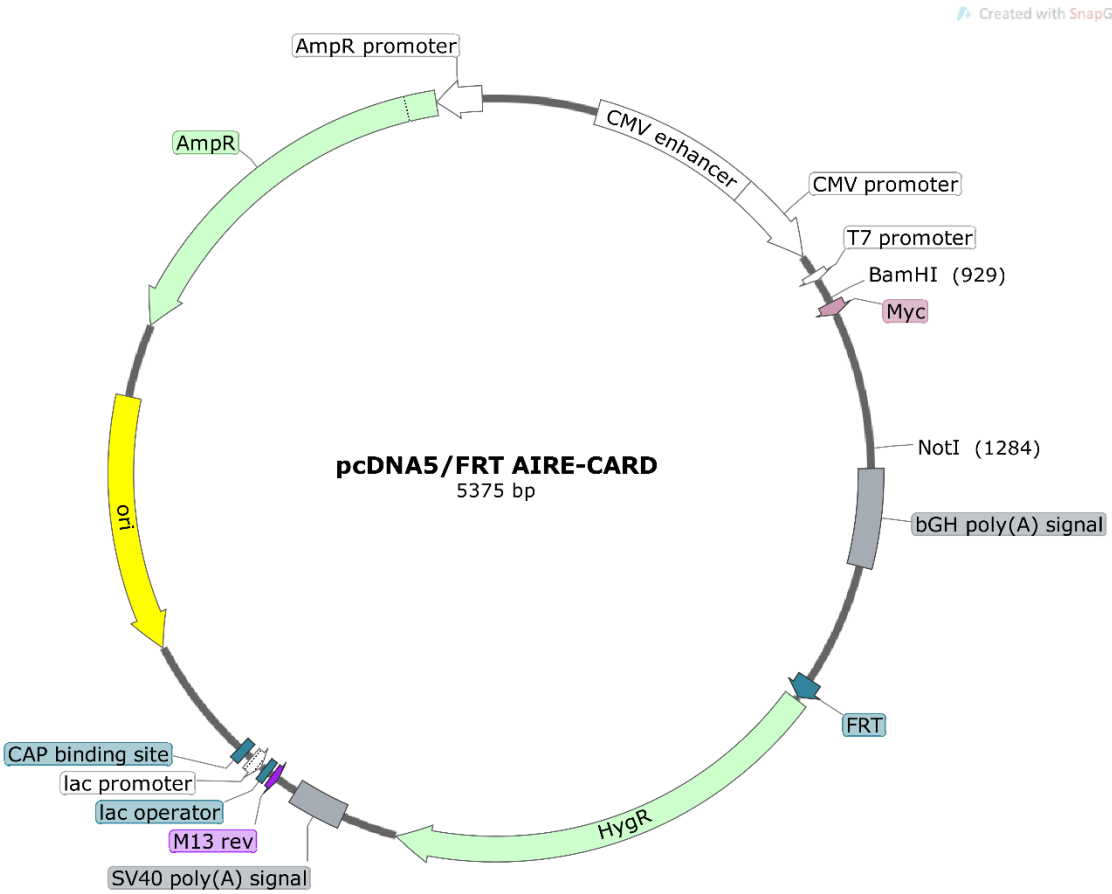


Figure 29 : pcDNA5/FRT AIRE-CARD plasmid map.

AIRE-CARD including the c-cmyc tag was inserted in the multiple cloning site at the *BamH1* and *Not1* restriction sites and and is under the control of the CMV promoter. The Plasmid is resistance to ampicillin and the origin of replication is shown in yellow. This plasmid is generated using SnapGene 5.0.8 software.

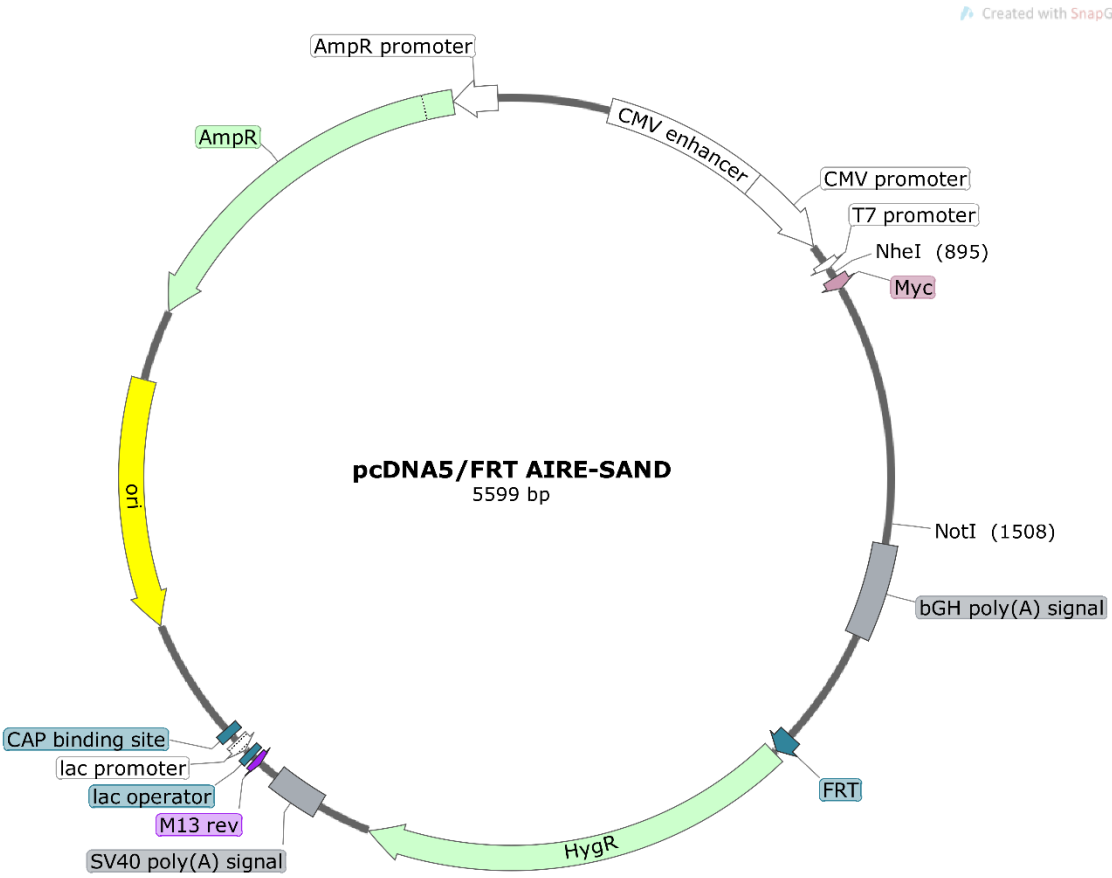


Figure 30 : pcDNA5/FRT SAND plasmid map.

AIRE-SAND including the c-cmyc was inserted in the multiple cloning site at the *Nhe1* and *Not1* restriction sites and and is under the control of the CMV promoter. The Plasmid is resistance to ampicillin and the origin of replication is shown in yellow. This plasmid is generated using SnapGene 5.0.8 software.

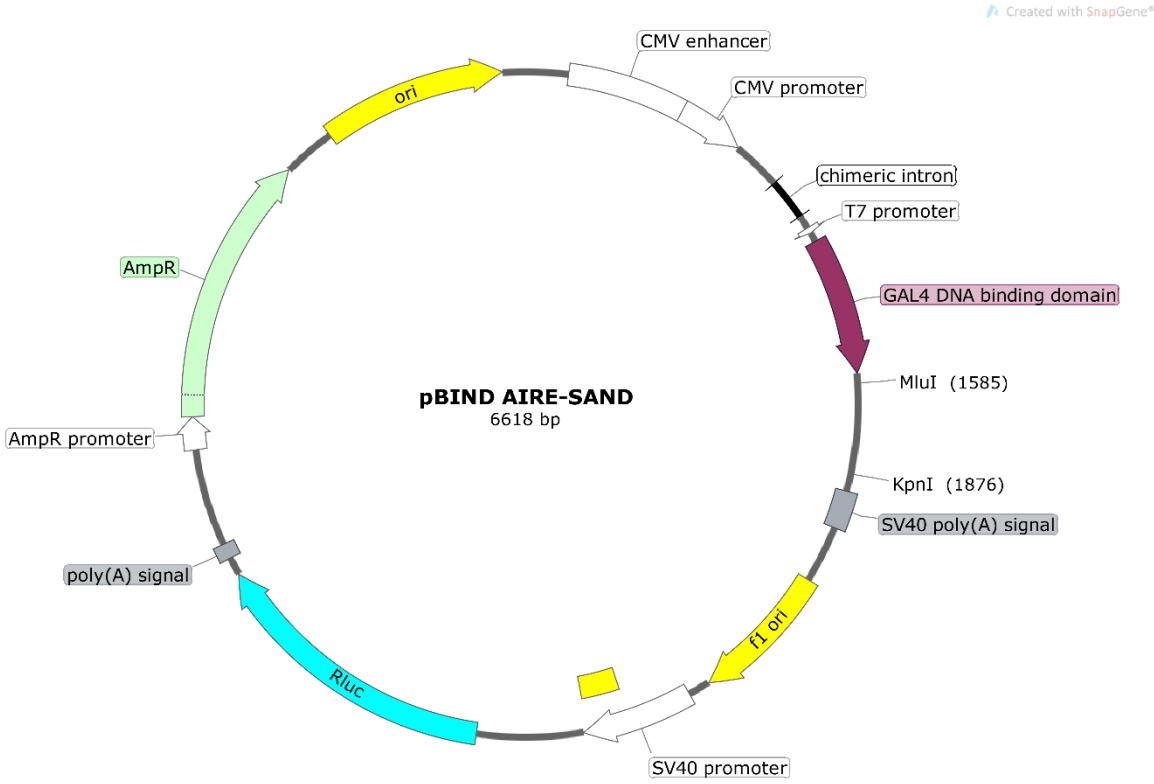


Figure 31 : pBIND AIRE-SAND plasmid map.

AIRE-SAND was inserted in the multiple cloning site at the *Mlu1* and *Kpn1* restriction sites, expressed to the upstream Gal4 DBD,and is under the control of the CMV promoter. The Renilla luciferase gene coding region is shown in turquoise and the origin of replication is displayed in yellow. The plasmid is resistance to ampicillin. This plasmid is generated using SnapGene 5.0.8 software

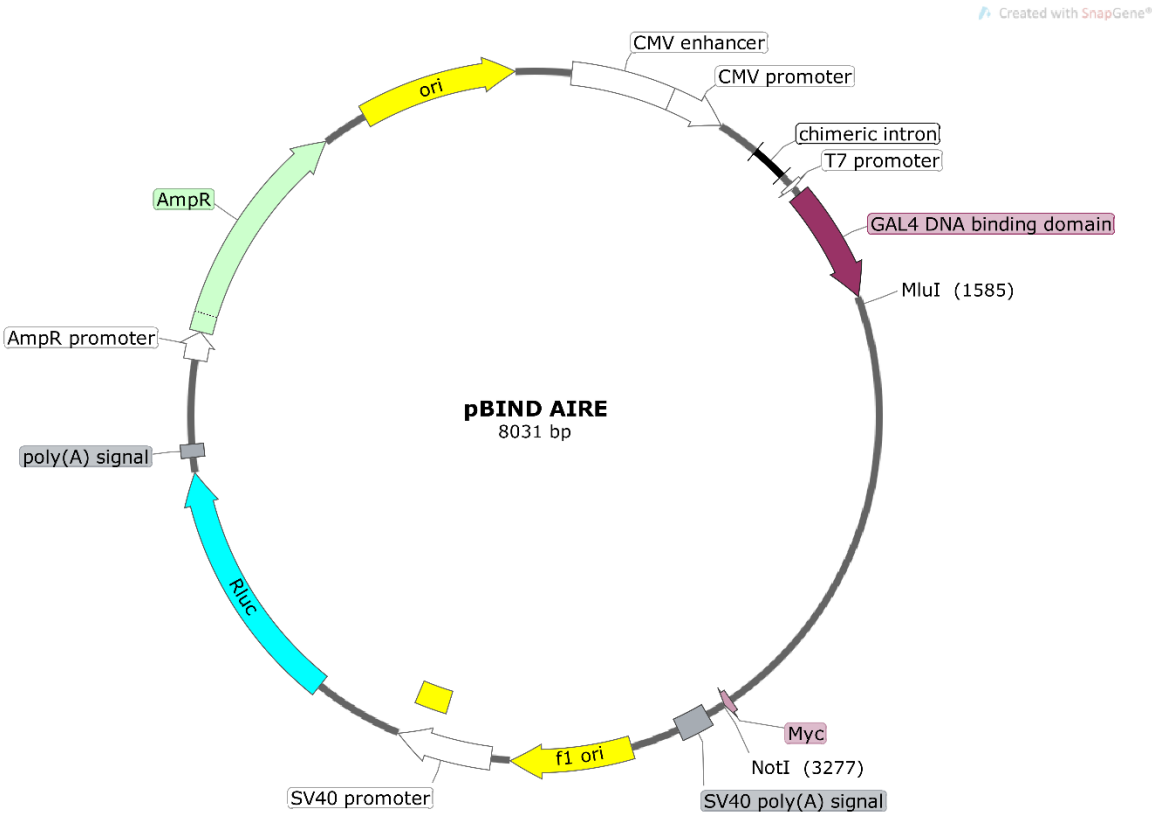


Figure 32 : pBIND AIRE plasmid map.

AIRE was inserted in the multiple cloning site at the *Mlu1* and *Not1* restriction sites, expressed to the upstream Gal4 DBD,and is under the control of the CMV promoter. The Renilla luciferase gene coding region is shown in turquoise and the origin of replication is displayed in yellow. The plasmid is resistance to ampicillin. This plasmid is generated using SnapGene 5.0.8 software

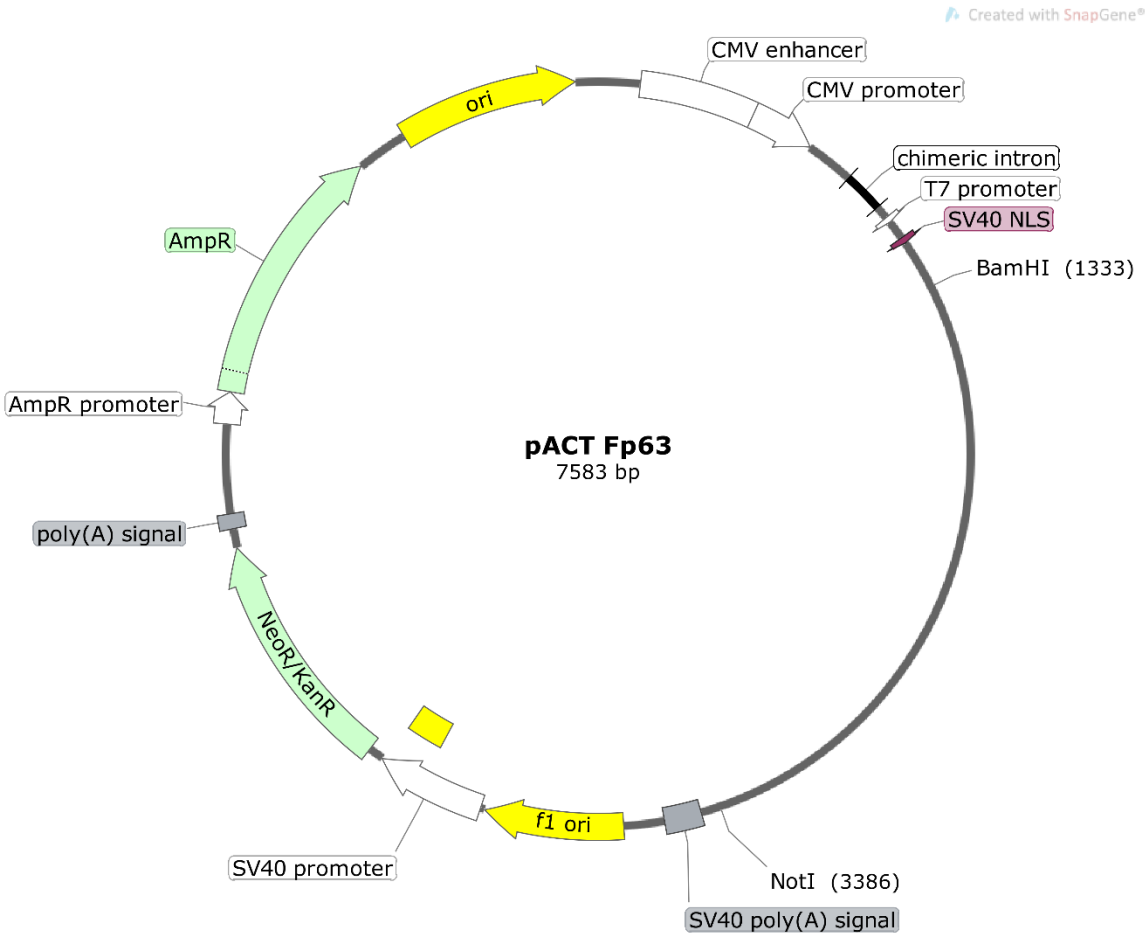


Figure 33 : pACT Fp63 plasmid map.

AIRE was inserted in the multiple cloning site at the *Bamh1* and *Not1* restriction sites, expressed to the upstream nuclear localisation signal (NLS),and is under the control of the CMV promoter. The origin of replication is displayed in yellow. This plasmid is generated using SnapGene 5.0.8 software

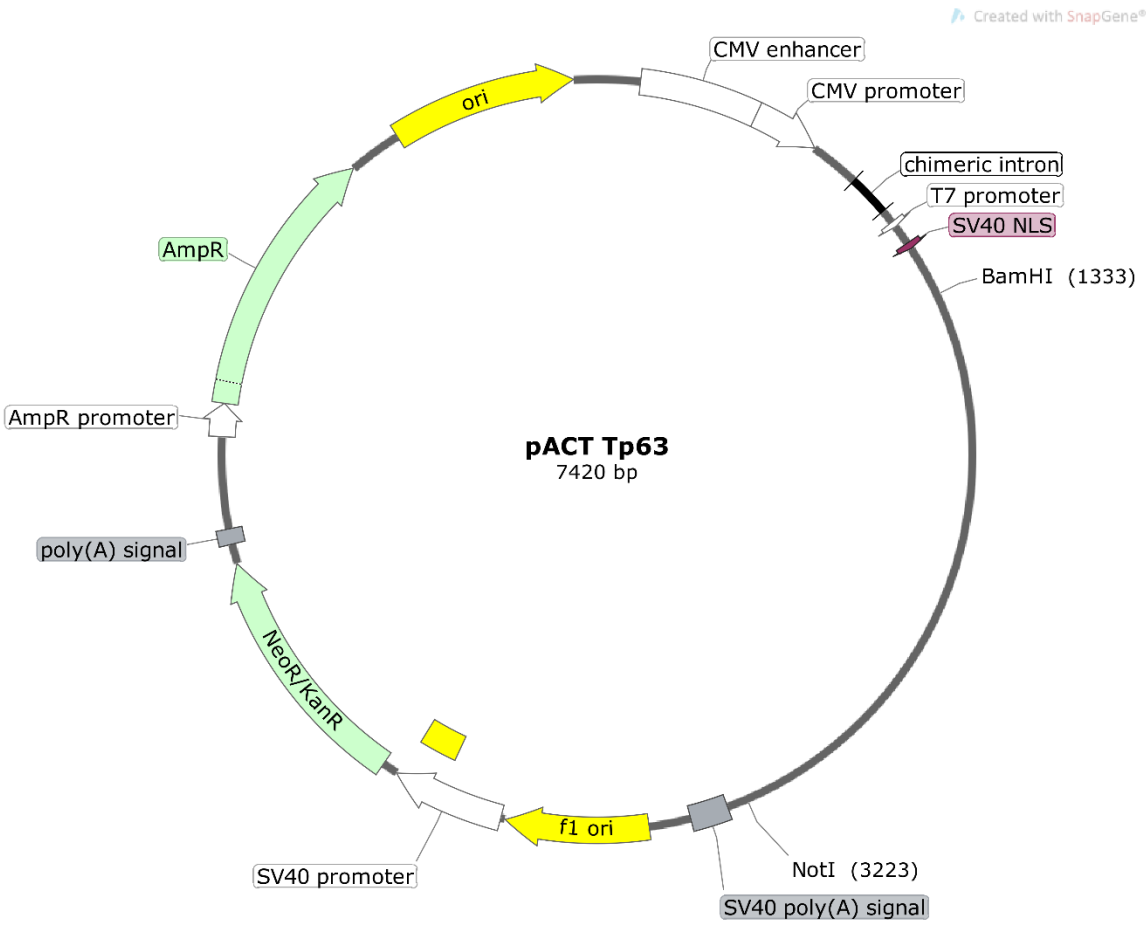


Figure 34 : pACT Tp63 plasmid map.

AIRE was inserted in the multiple cloning site at the *Bamh1* and *Not1* restriction sites, expressed to the upstream nuclear localisation signal (NLS),and is under the control of the CMV promoter. This plasmid is generated using SnapGene 5.0.8 software

7.16 Blots confirming the co-IP of AIRE-SAND against P63.

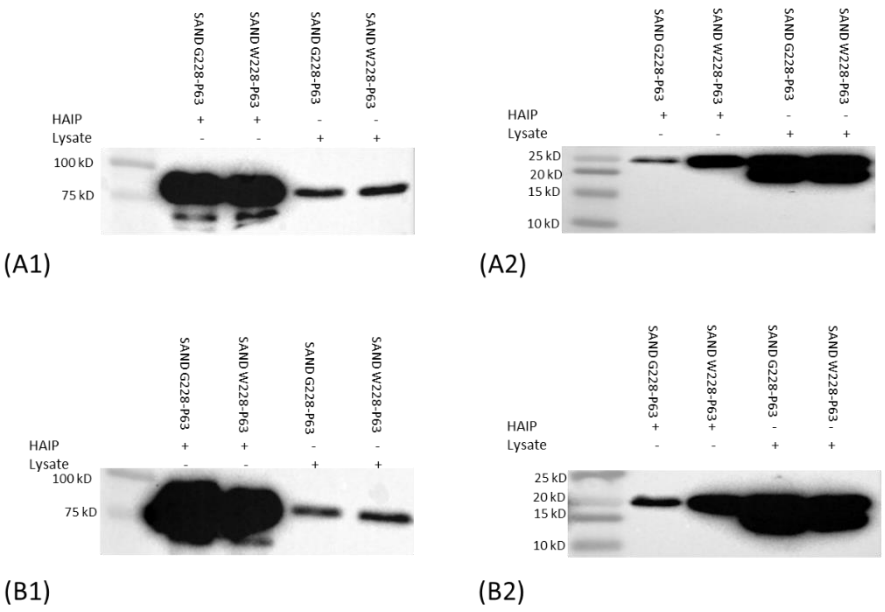


Figure 35: Blots confirming the co-IP of AIRE-SAND wild type and mutant against P63. Co-IP was carried out using HA-tagged beads. Interaction of cmyc-tagged AIRE-SAND mutants with the HA-tagged prey protein P63 was validated via co-immunoprecipitation and analysed via western blotting. HA-tagged beads were used to immunoprecipitate the interacting protein. For blots A1 and B1, the immunoprecipitated proteins were separated by SDS-PAGE in 7% (w/v) polyacrylamide gels and transferred to iBlot PVDF blotting membrane. For blots A2 and B2, the immunoprecipitated proteins were separated by SDS-PAGE in a 15% (w/v) polyacrylamide gel. The results are shown in blots A1 and B1: direct immunoprecipitation of P63 using HA-tagged beads (positive control) (size ~ 80 kD); overexpression of P63 total protein lysates (size ~ 80 kD). Blots A2 and B2 represent coimmunoprecipitation of AIRE-SAND and AIRE-G228W which were immunoblotted using cmyc antibody (size ~ 25 kD): overexpression of total protein lysate AIRE-SAND and AIRE-G228W gave a bans size of (size 20kD and 25 kD).