Analysis of variants of *Dnmbp*, a Putative Modifier of *Cecr2* and Neural Tube Defect

Susceptibility Gene

by:

Parmveer Singh

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Molecular Biology and Genetics

Department of Biological Sciences

University of Alberta

Abstract

One of the most common groups of birth defects are neural tube defects (NTDs), affecting about 1 out of 1000 established pregnancies. NTDs arise during neurulation in vertebrates, preventing proper formation of the brain and/or spina cord. NTDs are characterized by an unfused neural tube, which leads to neurodegeneration due to exposure of neural tissue to the uterine environment. Both genetic and environmental factors contribute to the formation of NTDs. In mice, hundreds of genes have been associated with NTDs. One of these genes is *Cecr2*, which encodes a protein involved in chromatin remodelling. Mutation to *Cecr2* leads to the lethal cranial NTD exencephaly in mice. However, the penetrance of exencephaly in homozygous *Cecr2* mutants differs between strains. In BALB/cCrl mice, a homozygous hypomorphic *Cecr2* gene-trap mutation causes exencephaly in 54% of embryos while the same mutation on an FVB/N background does not lead to exencephaly. This difference in penetrance is likely due to modifier genes having an effect on the *Cecr2* mutant phenotype. Further genetic analyses revealed a major modifier locus on chromosome 19 thought to contain multiple modifiers of *Cecr2*. In a previous study, coding variants of genes in the modifier region that differed between the two strains of mice were identified along with expression differences. Candidate mouse genes hypothesized to be modifiers of *Cecr2* were then sequenced in a cohort of 156 unrelated human cranial NTD fetuses to uncover human variants of the genes that may lead to modification of *Cecr2*. Human and mouse variants were ranked based on *in silico* analyses, rarity, and predicted function. One of the top candidates was *Dnmbp*, a scaffold protein involved in actin and dynamin dynamics as well as CDC42 activation.

In this thesis, I hypothesized that *Dnmbp* modifies the penetrance of exencephaly in mice through *Cecr2*. Eight human variants of *DNMBP* and one mouse variant underwent functional

testing for the preservation of protein-protein interactions, CDC42 activation, and cellular localization. Each variant was produced through site-directed mutagenesis and cloned in an expression vector with an N-terminus HA-tag. Variants were transiently transfected into cell lines for analyses. VASP and MENA are two proteins that interact with DNMBP and have been associated with NTDs. I was unable to determine if DNMBP variants maintain their interaction with VASP or MENA due to inconsistency with the co-immunoprecipitation experiment. An assay that measured the number of dense-core vesicles in neuroblastoma cells was conducted to test the level of activated CDC42. Over expression of DNMBP should lead to increased dense-core vesicles due to increased activation of CDC42. However, when the mean fluorescence intensity of a dense-core vesicle marker was measured, there was much variation. Thus, this assay did not provide information on the ability of each variant to activate CDC42.

To determine if there was a genetic interaction between Cdc42 and Cecr2, a mouse model was produced on a FVB/N background containing a heterozygous deletion of Cdc42 and homozygous mutation to Cecr2. Exencephaly was not observed in mice with these two mutations, thus, *Dnmbp* may not be modifying Cecr2 through *Cdc42*. Irregularities in localization of DNMBP protein were observed with one of the variants (R1024X) in epithelial cells. Caco-2 cells transfected with R1024X did not show localization of DNMBP to the apical cell membrane, confirming that this variant is damaging and giving evidence that *Dnmbp* is an NTD modifier gene. While other variants did show localization to the apical membrane of epithelial cells, it is possible that they may not localize correctly to specific regions of other cell types. Thus, it would be worth investigating localization patterns of each variant in other cellular regions DNMBP is known to concentrate. To further associate *Dnmbp* as a modifier of *Cecr2* and a NTD susceptibility gene, the best strategy would be to knockout one allele of *Dnmbp* in the

resistant FVB/N strain of mice with a homozygous *Cecr2* mutation. If penetrance of exencephaly

increases, it would be well supported that *Dnmbp* is a modifier of *Cecr2*. Characterizing the

function and localization of *Dnmbp* variants from susceptible mice and human NTD fetuses

could be beneficial in revealing a possible NTD susceptibility gene as well as further elucidating

the complexity of neural system development.

Preface

Sections 1.2-1.4: Parts of sections 1.2-1.4 were taken directly from a published review article for which I was second author (R. Y. M. Leduc, P. Singh, and H. E. McDermid, "Genetic Backgrounds and Modifier Genes of NTD Mouse Models: An Opportunity for Greater Understanding of the Multifactorial Etiology of Neural Tube Defects," Birth Defects Research Part A). These parts were written primarily by myself with some portions being written by Dr. Heather McDermid. Table 1.4.1, which I produced and was edited by Dr. Heather McDermid, was also published in the article.

Section 2.15: The use of mice for this research was approved by the Animal Care and Use Committee of the University of Alberta, University of Alberta AUP 00000094.

Section 3.1: Some of the co-immunoprecipitation experiments were conducted with the help of an undergraduate student, Mckenzie Mitchell.

Acknowledgments

The completion of this thesis definitely would have been much more difficult and lengthy without the help and support of a number of individuals.

Firstly, I would like to express my appreciation for my supervisor, Dr. Heather McDermid, for providing me with the opportunity to conduct this research. I am grateful for her constant support, guidance, and patience. I was tremendously fortunate to have a supervisor who was always there to help. Without a doubt, I would have been have been lost without her.

I would like to thank my committee members, Dr. Frank Nargang and Dr. Andrew Waskiewicz, for taking the time to provide their expertise and thoughtful feedback on my project and encouraging me to progress throughout my program.

Lastly, I would like to thank many of my labmates, both undergraduate and graduate, who were there to help me on a daily basis and made it a joy to come into the lab. I would like to specifically thank Dr. Renee Leduc for setting up the initial stages of the project and for her guidance as I began. I would also like to thank Farshad Niri and Kacie Norton for taking the time to teach me many techniques and guide me through experiments.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

°C	Degrees Celsius
1KG	1000 Genomes Database
3D	Three-dimensional
AA	Amino Acid
Abl	ABL Proto-Oncogene 1, Non-Receptor Tyrosine Kinase
Aldh1a2	Aldehyde Dehydrogenase 1 Family Member A2
aPKC	Atypical Protein Kinase C
Arg	ABL Proto-Oncogene 2, Non-Receptor Tyrosine Kinase
ATP2B2	ATPase Plasma Membrane Ca^{2+} Transporting 2
attL	Attachment L
attP	Attachment P
attR	Attachment R
BAR	Bin/Amphiphysin/Rvs
BET	Bromodomain and Extraterminal Domain
BMP	Bone Morphogenetic Protein
BRD2	Bromodomain Containing 2
BRDT	Bromodomain Testis Associated
BSA	Bovine Serum Albumin
CDC42	Cell Division Cycle 42
Cdh23	Cadherin 23
cDNA	Complementary DNA
Cecr2	Cat Eye Syndrome Chromosome Region, Candidate 2
Celsr	Cadherin, EGF LAG seven-pass G-type receptor 1
CERF	Cecr2-containing Remodeling Factor
Cfl1	Cofilin 1
CgA	Chromogranin A
Churc	Churchill
CR	Conserved Region
CRS	Craniorachischisis
ct	Curly-tail
Cyp26a1	Cytochrome P450 Family 26 Subfamily A Member 1
Dact1	Dishevelled-Binding Antagonist Of Beta-Catenin 1
DAM	Deoxyadenosine methylase
DAPI	4',6-diamidino-2-phenylindole
DDT	DNA binding homeobox and different transcription factor
Del	Deletion
dfw	Deaf waddler
DH	Dbl-homologous
Dlc1	Deleted In Liver Cancer 1
DLHP	Dorsal Lateral Hinge Point
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid

DNMBP	Dynamin Binding Protein
dNTP	Deoxynucleoside Triphosphate
DTT	Dithiothreitol
Dvl	Dishevelled
E. coli	*Escherichia coli*
E-cadherin	Epithelial cadherin
EDTA	Ethylenediaminetetraacetic Acid
Epha7	Ephrin type-A Receptor 7
ESP6500	Exome Sequencing Project 6500
EVL	Enah/Vasp-Like
EX	Exencephaly
F1	Offspring of first cross
F-actin	Filamentous Actin
FC	Facial Cleft
FCD	Folate and Choline Deficient
FGF	Fibroblast Growth Factor
Fpn	Ferroportin
FRAP	Fluorescence Recovery After Photobleaching
FRET	Fluorescence Resonance Energy Transfer
Fzd	Frizzled
G-actin	Globular Actin
GAP	GTPase-Activating Protein
GDI	Guanine Nucleotide Dissociation Inhibitor
GDP	Guanosine diphosphate
GEF	Guanine Nucleotide Exchange Factor
GERP	Genomic Evolutionary Rate Profiling
GM130	Golgi Matrix Protein 130
Gpr161	G Protein-Coupled Receptor 161
Grhl	Grainyhead Like Transcription Factor
Grlf1	Glucocorticoid Receptor DNA Binding Factor 1
GST	Glutathione S-transferase
GT	Gene-trap
GTbic45	*Cecr2* gene-trap mutation
GTP	Guanosine triphosphate
HA	Hemagglutinin
HRP	Horseradish Peroxidase
Hsp70	Heat Shock Protein 70
IF	Immunofluorescence
IgG	Immunoglobulin G
INDEL	Insertion/Deletion
InlC	Internalin C
IP	Immunoprecipitation
IQGAP1	IQ Motif Containing GTPase Activating Protein 1
IRSp53	Insulin Receptor Tyrosine Kinase Substrate Protein of 53 kDa
ISWI	Imitation Switch

jmj	Jumonji
kDa	Kilodalton
L	Liter
LB	Lysogeny Broth
Lmnb1	Lamin B1
Lmo4	LIM Domain Only 4
LOD	Logarithm of the Odds
Lrp6	LDL Receptor Related Protein 6
MAF	Minor Allele Frequency
Marcks	Myristoylated Alanine Rich Protein Kinase C Substrate
MDCK	Madin-Darby Canine Kidney
mdfw	Modifier of the deaf waddler
Men1	Menin 1
MENA	Mammalian Enabled (Enah)
mg	Milligrams
MHP	Medial Hinge Point
ml	Millilitres
mM	Millimoles
MMS19	MMS19 Nucleotide Excision Repair Homolog
Modvl5	Modifier of Vacuolated Lens 5
n	Sample Number
NAP1	Nucleosome Assembly Protein 1
ND	Non-Denaturing
Nf1	Neurofibromatosis Type 1
ng	Nanograms
nm	Nanometers
NT	Non-Transfected
NTD	Neural Tube Defect
N-WASP	Neural Wiskott-Aldrich Syndrome Protein
P	P-value
Palld	Palladin, Cytoskeletal Associated Protein
Pax3	Paired Box 3
PBS	Phosphate Buffered Saline
PBS-T	Phosphate Buffered Saline with Tween-20
PCP	Planar Cell-Polarity
PCR	Polymerase Chain Reaction
Pfn1	Profilin I
PH	Pleckstrin Homology
PIR121	P53-Inducible Protein 121
PP2	PolyPhen-2
PRD	Proline/Arginine-Rich Domain
Ptch1	Protein patched homolog 1
Ptk7	Protein Tyrosine Kinase 7
qRT-PCR	Quantitative Real-Time PCR
RhoGEF	Rho Guanine Nucleotide Exchange Factor

RIPA	Radioimmunoprecipitation Assay
RNA	Ribonucleic Acid
RNAi	RNA interference
RPM	Revolutions Per Minute
Rxn	Reaction
Sall2	Spalt-Like Transcription Factor 2
SB	Spina Bifida
Scrb/Scrib	Scribbled Planar Cell Polarity Protein
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
Sec24B	Secretory Protein 24 Homolog B
SF	Serum-free
Sfrp	Secreted Frizzled-Related Protein
SH3	Src homology-3
Shh	Sonic Hedgehog
Shmt1	Serine Hydroxymethyltransferase 1
Shp2	Protein-Tyrosine Phosphatase 2
shRNA	Short Hairpin RNA
SIFT	Sorting Intolerant from Tolerant
Sip1	Smad-interacting-protein-1
SNF2L	SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily A, Member 1 (SMARCA1)
SNV	Single Nucleotide Variant
SOB	Super Optimal Broth with Glucose
TBST	Tris-Buffered Saline with Tween-20
Tcof1	Treacle Ribosome Biogenesis Factor 1
Tet	Tet Methylcytosine Dioxygenase
Tgif	TGFB induced factor homeobox 1
TJP	Tight Junction Protein
tm1.1Hemc	*Cecr2* Deletion Mutation
Tmem67	Transmembrane Protein 67
Tris-HCl	Tris Base with HCl
TX	Triton X-100
U	Unit
V	Volts
Vangl	VANGL Planar Cell Polarity Protein
VASP	Vasodilator-Stimulated Phosphoprotein
Vcl	Vinculin
vl	Vacuolated Lens
WAVE1	WASP Family Protein Member 1
WB	Western blot
Wnt	Wingless type
Wnt5a	Wnt Family Member 5A
WT	Wildtype
ZO	Zonula occluden
β-Gal	Beta-Galactosidase

μg	Micrograms
μl	Microliters
μm	Micrometer

Chapter 1:

Introduction

1.1: Neurulation

Neurulation is the process by which the central nervous system forms during embryogenesis. It can be split up into two phases – primary and secondary neurulation. Primary neurulation leads to the formation of the precursor of the brain and most of the spinal cord, while secondary neurulation results in formation of the caudal most portion of the spinal cord (*1*).

The first phase of neurulation, primary neurulation, transforms the neural plate into the neural tube. A dorsomedial segment of the ectoderm thickens beginning in the cranial region and moving towards the caudal end of the embryo, forming the neural plate (*2*). The primitive node found in the mesoderm orchestrates this ectoderm to neural tissue transformation known as neural plate induction (*3*). By default, the ectoderm will differentiate into neural tissue (*4–6*). However, neural fate can be inhibited by activating the bone morphogenetic protein (BMP) signaling pathway, which results in ectodermal cells becoming epidermal due to the expression of non-neural transcription factors (reviewed in *7, 8*). Proteins secreted from the primitive node (e.g. noggin, chordin, cerberus, XNr3 and follistatin) prevent BMP activation by binding to BMP ligands, thereby allowing neural plate formation. Newly formed neural ectodermal cells express a variety of neural-specific transcription factors from the Fox, Sox, Zic, and lrx families, some of which prevent the neural cells from reverting to an undifferentiated state due to future BMP signals (*7, 9*). The wingless type (Wnt) pathway has been strongly associated with neural plate induction, though there is evidence supporting both its role in the formation of neural tissue and the inhibition of neural induction (reviewed in *10*). The fibroblast growth factor (FGF) pathway also plays a role in neural plate induction. A downstream effector of the FGF pathway, Churchill (Churc) upregulates neural genes while also downregulating mesodermal genes (*11*). Churc mediates Sip1 (Smad-interacting-protein-1) activity, which inhibits BMP signaling and

mesoderm genes (*7*, *12*). Together, the BMP, Wnt and FGF signalling pathways and signals from the primitive node lead to neural induction in a distinct region of the ectoderm, becoming columnar neural ectodermal cells. Overexpression of neural transcription factors that are normally expressed in neural ectodermal cells results in a wider neural plate, likely due to the broader repression of BMP signalling (*7*). The edges of the neural plate are defined by a neural plate border formed by dynamic expression of BMP, FGF, and Wnt associated genes (*8*). The neural plate borders eventually become neural crest cells.

Shaping of the neural tube begins with neural plate folding. The neural plate is initially elliptical in shape, but is transformed to a more elongated shape with a broad rostral region and a narrow caudal region (*13*). This change occurs through convergent extension cell intercalation within the plate with a net movement of cells medially; the plate becomes narrow mediolaterally and lengthens rostrocaudally. The movement of cells during this process is controlled by the planar cell-polarity (PCP) pathway (*14*).

While convergent-extension of the neural plate is occurring, the plate will begin to bend and fold to form a tube (Figure 1.1.1A). This event begins with the formation of hinge points. There are three hinge points, one found medially that is above the notochord and two dorsolateral that are located close to the lateral edges of the neural plate (*13*). The medial hinge point (MHP) forms in the upper region of the body axis and the midspinal region while the dorsal lateral hinge points (DLHPs) form in the cranial region, midspinal region , and lower spinal regions (*1*). While the mechanism of bending at the hinge points is not fully understood, it has been observed that cells at the MHP have a wedge shaped morphology due to apical constriction of neuroepithelial cells (*1*). Sonic hedgehog (*Shh*) expression from the notochord induces the formation of the MHP, but at the same time, represses DLHP formation (*15*). BMP activation

also represses DLHP formation (*16*). In mouse embryos lacking a notochord, DLHP formation occurs throughout the body axis (*17*). DLHP formation is repressed through inhibition of Noggin and neuralin (BMP antagonists) by *Shh*. This inhibition of Noggin leads to the expression of BMP2 from the dorsal ectoderm, preventing DLHP formation (*17*). When Noggin is not inhibited, it is expressed from the edges of the neural plate, allowing for DLHP formation (*18*). Thus, a complex interplay of *Shh*, Noggin, and BMP signalling results in the varying formation of the medial and dorsolateral hinge points along the body axis.

As the neural plate bends at the hinge points, the tips of the neural folds will move towards each other and eventually meet at the dorsal midline. Adhesion followed by fusion will then occur to complete the neural tube. The two ends of the non-neural ectoderm also meet and fuse to form a continuous layer of non-neural ectoderm above the neural tube, later becoming the dorsal epidermis. From the neural fold tips, cellular protrusions similar to filopodia and lamelipodia form when the tips are adjacent to each other and interdigitate (*1*, *19*). Adhesion is not fully understood, though it is thought that ephrin cell surface receptors play a role (*1*).

Once neural folds adhere, fusion initiates at distinct closure sites in mice: (1) the hindbrain/cervical boundary, (2) the forebrain/midbrain boundary, and (3) the extreme rostral end of forebrain (Figure 1.1.1B) (*13*). Fusions spreads bidirectionally from closures site 1 and 2, and unidirectionally towards the caudal end from closure site 3. In humans, only closure sites 1 and 3 are present. Additionally, closure site 2 in mice is polymorphic in location and is not present in some mouse strains (*13*, *20*). Once fusion of the neural tube is complete, remodelling occurs in which the neural tube separates from the surface ectoderm by a mechanism that has yet to be elucidated. At the end of primary neurulation, the neural tube is completely fused except for a caudal region, known as the posterior neuropore.

Secondary neurulation is the phase of neurulation that forms the caudal portion of the neural tube. Rather than folding, the caudal portion forms using a population of stem cells from the embryo tail bud that condense and become epithelial cells, forming an epithelial rod (*17*). Within the rod of cells, a cavity forms that is continuous with the neural tube, resulting in a closed neural tube and the completion of neurulation.
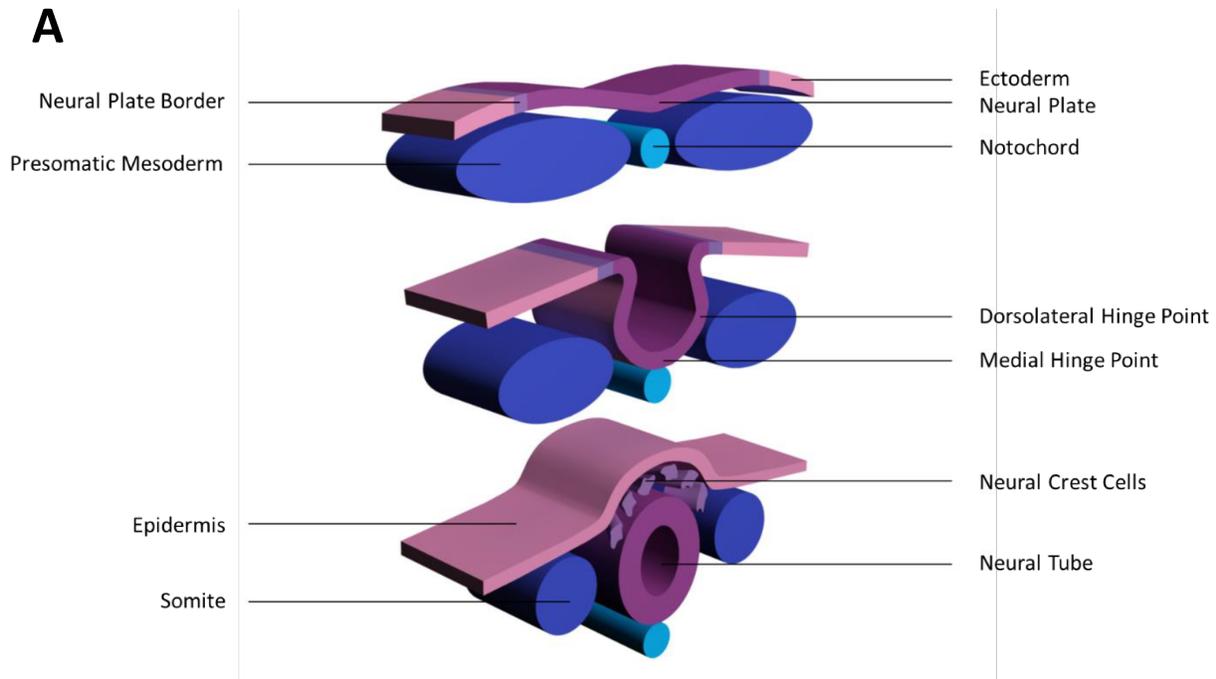
Figure 1.1.1: Formation of the neural tube in vertebrates. Neurulation, the process by which the precursor of the brain and spinal cord are formed, begins shortly after the establishment of the ectoderm, mesoderm and endoderm layers in the embryo. **(A)** A portion of the ectoderm thickens to form the neural plate flanked by the neural plate borders. The neural plate will begin to bend at the medial and dorsolateral hinge points. The edges of the neural fold will meet at the midline and fuse to form the neural tube. Cells making up the neural plate border become neural crest cells. The non-neural ectoderm also meets at the midline and fuses to form the dorsal epidermis. **(B)** Fusion of the neural tube begins at specific closure sites. Closure site 1 (C1) is located at the hindbrain/cervical boundary. In most mice but not humans, a second closure site is found at the forebrain/midbrain boundary (red asterisk). Closure site 3 (C3) is located at the extreme rostral end. Arrows indicate direction of fusion once initiated. Yellow lines indicate region of cranial NTDs such as exencephaly and blue lines indicated regions of spinal NTDs such as spina bifida. The most severe NTD, craniorachischisis results in an open neural tube along the entire body axis.
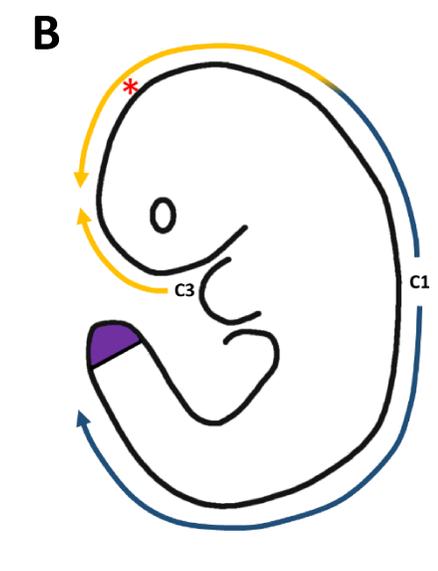
1.2: Neural Tube Defects

When neurulation is disrupted, a neural tube defect (NTD) results. NTDs affect about 1 out of 1000 established pregnancies (*21*). Abnormalities are observed in the neural tube due to degeneration of neural tissue leading to brain and/or spinal cord malformations (*22*). A variety of NTDs exist with varying levels of severity. The development of a certain NTD depends on the location of neurulation disruption along the body axis (Figure 1.1.1B).

NTDs can be classified into two groups: Closed or Open. Closed NTDs, which include spina bifida occulta and spinal dysraphism sequence (tethered cord), are restricted to the spinal region and are covered by skin. Closed NTDs are the least debilitating types of the two (*16*). They are thought to originate from defects in secondary neurulation, though they are not well characterized (*1*). The most well-studied fall under the category of Open NTDs, which arise due to defects in primary neurulation (*16*).

Open NTDs can be sub-classified as cranial or spinal. Cranial NTDs, such as anencephaly (referred to as exencephaly in mice), are lethal and occur when neural folds in the cranial region do not completely fuse. Anencephaly is characterized by craniofacial defects and partial or complete loss of the brain, braincase and skin, leading to degenerated neural tissue protruding from the brain (*16, 23*). Unlike cranial NTDs, spinal NTDs are less severe and usually non-lethal, though individuals are severely handicapped. A common group of spinal NTDs is referred to as spina bifida. Of this group, meningomyelocele (also known as myelomeningocele) is the most common. Meningomyelocele is characterized by the protrusion of the meninges (a membrane that normally covers the spinal cord) containing neural tissue and other spinal components at the site of incomplete neural tube closure (*16*). Myelocele (myeloschisis) is

similar to meningomyelocele, though there is no membrane covering the neural tissue, thus, it is exposed to the uterine environment and is more severe (*23*).

Shh signaling is an example of a genetic pathway that has been associated with exencephaly and spina bifida. Specifically, when continual *Shh* activation is achieved in mice through null mutations to negative regulators of *Shh* (e.g. *Ptch1*) the neural tube does not close in the cranial and spinal region (*24*, *25*). In neurulation, the *Shh* signalling pathway is associated with dorsoventral axis patterning (*25*). Mutations to inhibitors of *Shh* usually result in expansion of ventral markers, followed by exencephaly and sometimes spina bifida, depending on the *Shh* regulator that is mutated (reviewed in *25*). The BMP pathway has also been associated with cranial and spinal NTDs. Mice with a homozygous loss-of-function allele of *Zic2* develop exencephaly and spina bifida at frequencies of 69% and 100%, respectively (*18*). When *Zic2* activity is lost, expression of noggin and neuralin are "undetectable" and there is increased activation of BMP signalling, resulting in repression of DLHP formation. Thus, it is thought that *Zic2* is associated with the expression of BMP antagonists, which would normally allow DLHP formation to proceed. Without the DLHPs, the neural tube is unable to close in certain regions of along the body axis, resulting in spina bifida and exencephaly.

Craniorachischisis is the most severe type of NTD as it results in an open neural tube throughout the cranial and spinal region (*1*). If closure initiation does not occur at the hindbrain-cervical boundary (closure site 1, Figure 1.1.1B), craniorachischisis develops, affecting the midbrain, hindbrain, and spinal cord (*22*). One well studied mechanism that can lead to craniorachischisis when disrupted is convergent extension. As illustrated in mice, abnormalities of convergent extension during neurulation due to defects in the PCP pathway genes such as *Frizzled* (*Fzd*-1 and *Fzd*-2), *Dishevelled* (*Dvl*-1 and *Dvl*-2), *Vangl2*, *Celsr1*, *Scrb1*, and *Ptk7*,

result in abnormally wide neural plates (*1*, *26–28*). This prevents the neural tube from closing likely due to the lack of proximity of the neural folds, leading to craniorachischisis. This notion is further supported by human craniorachischisis fetuses displaying short rostro-caudal body axes (*16*).

In humans, the inheritance of NTDs is complex and multifactorial, involving both environmental and genetic factors (reviewed in *14*). Major environmental factors include folic acid levels early in pregnancy (*29*, *30*), as well as maternal diabetes (*31*), obesity (*32*) and hyperthermia (*33*, *34*). The genetic basis of NTDs in humans involves many genes, with the presence of an affected first degree relative imparting a 3-5% risk of having an NTD in one's offspring (*35*). The recurrence of an NTD in future siblings is 2-5%, a 50-fold increase relative to the general population (*2*). To deal with this complexity of NTDs, the mouse presents a model system more amenable to study. Mice allow easy access to embryos during neurulation, as well as genetic and molecular techniques to manipulate genes to test hypotheses. There are estimated to be more than 300 genes in mice that when mutated result in one or more NTDs (*14*). Of these genes, only a small percentage shows obvious oligogenic inheritance. There are at least 37 gene pairs listed on NTDWiki (https://ntdwiki.wikispaces.com) that show digenic inheritance (*14*). Most of these pairs involve functional redundancy of a pathway, where both genes must be heterozygous or homozygous mutated for an NTD phenotype to arise. There is one trigenic trio of genes involved in the PCP pathway (*Sfrp1*, *Sfrp2* and *Sfrp5*) that all must be knocked out to produce the NTD craniorachischisis (*36*). Environmental factors also interplay with genetic factors in mice in the development of NTDs. For instance, mutations to *Shmt1* in mice on a normal lab diet does not lead to exencephaly (*37*). However, if mice with the same mutation are on a folate and choline deficient diet, there is a significant increase in the penetrance of

exencephaly. Thus, mice have been shown to serve as models for multifactorial inheritance of NTDs (*38*). Still, many of the genes that have been associated with NTDs in mice seem to be single causative genes since a single mutation to a gene can lead to an NTD, making the mouse an unsuitable model of the multigenic human NTDs.

However, even with apparently single causative genes, there is often an underlying level of complexity that becomes apparent when the same mutations are viewed on different genetic backgrounds, revealing the possible presence of variant modifier genes. In fact, the reason that some genes show "simple" inheritance may be because they have only been studied on a single congenic strain where individual variation has been almost eliminated. Since laboratory mice are maintained in very controlled environments, variation is even lower compared to humans. Studying mutations on various genetic background, in which protein coding of a gene may vary, allows for the discovery of modifier genes that may express a different phenotype due to variation in coding of other genes. The genetic variation between genetic backgrounds would also to a lesser extent mimic the variation seen in human populations. Thus, uncovering a modifier gene in mice that is affecting what was thought to be a single causative gene would produce a multigenic system for studying NTDs and relating it back to humans.

1.3: Modifier Genes

While modifier genes are defined in various ways throughout the literature, the overall concept remains consistent. Simply put, a modifier can be described as a gene that affects the phenotype produced by another gene. Here, a modifier is defined as a variant secondary locus that affects the disease phenotype of a mutated target locus. The target locus itself is necessary to produce a disease phenotype, while the modifier will enhance or suppress the phenotype (*39*).

The modifier locus does not by itself produce the disease phenotype of interest, but may produce other abnormal phenotypes when mutated. A variation on this is gene-gene interactions between genes that both produce the phenotype of interest when mutated. In this case the "target" gene and the "modifier" gene are interchangeable. However, most modifier genes remain unidentified and therefore unclassified, and these two situations may be extremes of an overall spectrum.

Research on variant mouse modifier genes has been increasing steadily over the past 30 years (*40*). The ability to cross mutations onto inbred strains of mice with differing genetic backgrounds allows for less complicated identification of candidate modifier genes relative to humans. The mouse model also provides greater control over environmental factors, such as housing and diet, than with humans, although it is not necessarily clear which environmental factors are important to control. However, even in mice it is no simple task associating a modifier with the target locus disease phenotype. Modifiers and target genes do not necessarily interact directly, rather, they likely work through a complex genetic network. A modifier can have a seemingly unrelated function to that of the target gene, leading to a larger list of possible candidate modifier genes for a disease phenotype. Additionally, multiple modifiers can affect a single target gene, having additive effects on the disease phenotype. Thus, many genes may need to be studied extensively before modifiers can be identified.

The effect of a variant modifier gene can be subtle or substantial on a phenotype. Modification can lead to novel, enhanced, or reduced mutant phenotypes. Specifically, modifier genes can affect the dominance, expressivity, pleiotropy, and/or penetrance of a target gene as described below (*41*).

Even under similar environmental conditions, a trait that is normally dominant can be inherited in a recessive or semi-dominant fashion due to the effect of a variant modifier. For

instance, Noben-Trauth et al. (*42*) have described a modifier of the deaf waddler mutation (*mdfw*) in mice. When homozygous on the BALB/cByJ background, the deaf waddler (*dfw*) mutation leads to profound hearing loss by 3 weeks. A BALB/cByJ heterozygote for *dfw* shows progressive hearing loss beginning at 4 weeks and is deaf by 12 weeks, indicating semidominant inheritance. However, when crossed with the CAST/Ei background, the dfw$^{+/-}$ hybrids have no hearing loss, suggesting a recessive inheritance. The *dfw* mutation is in the ATPase, Ca^{2+} transporting, plasma membrane 2 gene (*ATP2B2*) (*43*) and the *mdfw* mutation is one of several alleles in the cadherin 23 gene (*Cdh23*) (*44*).

A variant modifier can affect the severity of a particular phenotype, thereby affecting expressivity. In terms of NTDs, changes in expressivity can lead to more or less severe NTDs – an increase in severity could mean a combination of different disease phenotypes (i.e. spina bifida and exencephaly) rather than only one disease phenotype. This is illustrated by modifiers of the *Ski* proto-oncogene showing expressivity effects. While an incipient congenic 129P3 strain with *Ski*$^{-/-}$ mutations only shows penetrance of exencephaly, 129P2-Swiss mixed, 129P2-C57Bl/6 mixed, and C57BL/6J congenic backgrounds display a second NTD, facial clefting, at various penetrance and clefting severity, thereby increasing the severity of the *Ski* mutation (*45*).

A target gene can have varying pleiotropic effects that are only observed with the presence of a variant modifier. For instance, a variant modifier can lead to changes to two unrelated disease phenotypes rather than just one in the absence of the variant modifier, as Andersson et al. (*46*) have shown. *Lrp6*$^{-/-}$ mutants results in various phenotypes including exencephaly and heart shape defects. *Wnt5a*$^{-/-}$ mutants results in limb defects and a shorter anteroposterior body axis. Compound homozygous mutations of *Lrp6* and *Wnt5a* result in an

absence of exencephaly, limb defects and anteroposterior body axis defects, while the penetrance

of heart shape defects increases.

Lastly, a variant modifier can affect the penetrance of the target gene phenotype. With

one variant of a modifier gene, we may not see a disease phenotype, but with an alternative

variant, the disease phenotype could may be observed at varying frequencies (*47–54*). Changes

in the penetrance of exencephaly can be observed when *Cecr2* is deficient in mice. On an FVB/N

congenic background, a hypomorphic *Cecr2* mutation does not result in any NTDs (*47*).

However, when the same mutation is on a congenic BALB/cCrl background, exencephaly

increases significantly to about 54% (*55*).

1.4: NTDs and Mouse Modifiers

A list of NTD susceptibility genes under the influence of a modifier is found in Table

1.4.1. Most of these modifier genes are currently unidentified. For those genes with identified

modifiers, *Lmnb1* exemplifies a gene that modifies the effect of *Grhl3* but does not produce an

NTD itself when mutated (*56*). Alternatively, *Pax3-Nf1* is a gene-gene interaction where

mutation of either gene results in an NTD (*57*). There are many such interactions known,

especially in the PCP pathway, which have not been included in Table 1.4.1. The interaction of

Lrp6 and *Wnt5a* could be considered intermediate, since mutation of modifier *Wnt5a* does not

cause NTDs but does produce related defects of body axis (*49*).

For NTD genes, the modifying gene(s) most frequently affect penetrance, followed by

expressivity (Table 1.4.1). In most cases, crossing a mutant gene that results in an NTD onto a

C57BL/6 background leads to a reduction in penetrance (usually compared to a 129/Sv, 129P2 or

129P3 background). Out of 14 studies reviewed that investigated the effects of the C57BL/6

background on NTD penetrance, 12 showed a decline in penetrance (Table 1.4.1). For example, when the proto-oncogene *Ski* contains a homozygous null mutation, the penetrance of NTDs is variable depending on the genetic background (*45*). On a 129P2-Swiss mixed background, the incidence of exencephaly is highest at 88% (n >19). Crossing the mutation onto a 129P2-C57Bl/6 background results in a slightly lower penetrance of 80% (n>19). When the mutation is on a congenic C57BL/6J background (N14), penetrance is reduced to 5% (n >19).

However, two genes reviewed showed an increase of NTDs when crossed onto the C57BL/6J background (Table 1.4.1). A heterozygous mutation of *TCOF1*, a gene encoding a nuclear phosphoprotein, results in complete penetrance of exencephaly on a C57BL/6 background (n=206), while on the BALB/cCrl background exencephaly penetrance is 6.6% (n=76) (*58*). *Tgif*, a homeobox gene, also shows increasing penetrance of exencephaly as the mutation is moved onto a C57BL/6 background from a 129Sv-C57BL/6 mixture: in homozygous mutants, the F1 cross showed 5.1% (2/39), the N2 showed 20.9%, (14/67) and the N3 showed 33.3% (9/27) (*59*). It is interesting to speculate on the general significance of the C57BL/6 background. Does this background contain a small number of key "generic" modifiers that have broad effects on one or more pathways, such as the strain-dependent position of the cranial closure site 2 (*60*) or a general shape change in the embryo that makes closure more difficult? Or does the C57BL/6 background contain many modifiers that are specific for each of the genes involved? The identification of modifier genes is required to answer these questions.

Mutations in single genes have been discovered that lead to enhanced or suppressed modified NTD phenotypes, affecting expressivity in addition to penetrance. Diet also has been seen to act as an environmental modifier. These factors combine in the common NTD mouse model, *Splotch*, which is characterized by almost completely penetrant exencephaly and/or spina

bifida due to a mutation in the *Pax3* gene. In addition to NTDs, homozygous mutant *Splotch* embryos show impairments in de novo thymidylate biosynthesis (*37*). NTD phenotypes can be rescued in *Splotch* mutants through folic acid and thymidine supplementation, presumably through rescuing of *de novo* thymidylate synthesis (*60*). Alternatively, NTD phenotypes can be enhanced in severity by mutations to a modifier gene, *Shmt1* (*37*). *Shmt1*, which encodes the enzyme serine hydroxymethyltransferase, is a part of the *de novo* thymidylate synthesis pathway. Isolated *Shmt1* deficiency in mice leads to impairment of de novo thymidylate biosynthesis, although mice do not develop NTDs when fed commercial rodent diet (*61, 62*). However, when *Shmt1* mutant mice are fed a folate and choline deficient (FCD) diet, exencephaly is seen at a low level (Table 1.4.1) (*37*). While the *Shmt1* mutation does not lead to NTDs in mice on diets supplemented with folic acid, the introduction of *Shmt1* mutations into *Splotch* mutants on a folate-supplemented diet increases the severity of NTD phenotypes (i.e. a combination of spina bifida and exencephaly) (Table 1.4.1). Heterozygous and homozygous *Shmt1* mutations on a *Splotch* background with a *Pax3* homozygous mutation have a significantly ($P = 0.04$ and $P = 0.001$, respectively) higher frequency of severe NTDs relative to embryos with just the homozygous *Pax3* mutation (*37*). Only 22% of $Pax3^{-/-}$ mutant embryos display lesions to both the spinal and cranial regions, whereas 66% of $Pax3^{-/-}$; $Shmt1^{-/-}$ mutants have legions to both regions. Moreover, NTD severity is increases and cases of craniorachischisis are present in double *Pax3*/*Shmt1* mutant embryos from dams on FCD diets (*37*). *Shmt1* deficiency seems to be modifying the expressivity of *Pax3* mutants by further impairing the de novo thymidylate synthesis pathway, leading to enhanced NTD phenotypes.

Conversely, Bryja et al. (*49*) have described a modifier that suppresses an NTD phenotype. Null mutations in *Lrp6*, a receptor in the canonical Wnt/β-catenin pathway, lead to

exencephaly penetrance of about 30% (n=9). When the non-canonical Wnt signaling pathway ligand *Wnt5a* is also deleted, exencephaly penetrance is reduced. An *Lrp6*[-/-]; *Wnt5a*[+/-] double mutant reduces exencephaly incidence to ~5% (n=30), while a double homozygous mutation leads to a complete suppression of exencephaly (n=7) (*49*). *Wnt5a* homozygous mutants do not show NTDs (*46*).

In some cases, further analyses of the modifying effects of different genetic backgrounds have localized specific modifier loci. The *vacuolated lens (vl)* mouse mutant displays NTD phenotypes, such as spina bifida due to a deletion mutation in the orphan G protein-coupled receptor gene, *Gpr161* (*63*). Using the various genetic backgrounds that affect penetrance of NTDs when *Gpr161* is mutated (Table 1.4.1), candidate modifier gene regions were mapped out with quantitative trait locus analyses (*64*). One of the candidate modifier loci mapped on the MOLF background, named *Modvl5*, reduces the incidence and severity of NTDs caused by *Gpr161* when it is crossed onto a penetrant strain, providing evidence for this locus as a modifier of *Gpr161*.

Moreover, in the case of the curly-tail mouse, specific modifier genes have been found that affect the penetrance of NTDs. The curly-tail mouse is characterized by exencephaly (5-10% penetrance) and spina bifida (15-20% penetrance) due to a homozygous hypomorphic mutation in the grainyhead-like 3 gene (*Grhl3*[ct]) (*56*). A homozygous deletion of *Grhl3* results in complete penetrance of spina bifida with some exencephaly (*65*). Crossing the *Grhl3*[ct] mutation onto other genetic backgrounds reduces the penetrance of NTDs (*66*). Through a proteomic analysis, De Castro et al. (*56*) were able to identify the intermediate filament protein gene *lamin B1 (Lmnb1)* as a modifier gene. A two-dimensional protein gel that separated proteins by isoelectric point and molecular mass found that there is a variation in *Lmnb1* charge between wildtype and curly-tail

embryos. Further analyses indicated that a glutamic acid is deleted in the curly-tail mouse from a series of 9 glutamic acid residues. Thus, the *Lmnb1* variant is denoted by 8E while the wildtype variant is denoted by 9E. Crossing the 9E variant onto a curly-tail background with the *Grhl3^{ct}* mutation results in a 3-fold reduction of spina bifida and exencephaly. Examination of the *Lmnb1* 8E variant indicated instability in the nuclear lamina with increased nuclear dysmorphology and reduced proliferation.

Table 1.4.1: Mouse genes and their associated modifying genes, background, or environment leading to NTD phenotypes.
Table from Leduc et al. (55).

1. Background Modifications

Target Gene	Genotype	Modifying Background[1]	Modifying Effect (Penetrance)[2]	Reference
Cart1	*Cart1-/-*	C57BL/6-129 Mixed 129/SvEv	Meroanencephaly (67%, 32/48) Meroanencephaly (100%, 13/13)	(67)
*Cecr2**	*Cecr2-/-*	BALB/cCrl (>N10) FVB/N (N5/6)	EX (54%)[3] NTDs Absent (0/45)	(47)[3]
Fpn	*Fpn-/-*	129/SvJ-C57BL/6 Mixed C3H/HeJ-C57BL/6 Mixed	NTDs Absent NTDs Present	(54)
Gcn5	*Gcn5*-flox (neo) (partially functional)	C57BL6 (N4) 129SvEv (N4)	NTDs (19%, 3/16) NTD (100%, 12/12)	(68)
Gpr161	*Gpr161-/-*	B6/C3H MOLF/C3H C3H CAST/C3H Modvl5 Region (MOLF)	SB (54.3%, 57/105) SB (0/126) SB (41.1%, 23/56), Other spinal NTDS (58.9%, 33/56) SB (44.7%, 42/94) NTDs and severity[4] reduced, SB (6.67%, 1/15)	(63, 64, 69)
Grhl3 (curly-tail)	*Grhl3ct/ct*	DBA/2J/ct	NTDs reduced to (2.2%, 5/225) relative to curly-tail background	(66)
jumonji	jmj-/-	BALB/cA-129/Ola Mixed C3H/He (N6-10) BALB/Ca (N11) C57BL/6J (N6) DBA/2J (N6)	NTDs (~40%) NTDs (100%) NTDs absent NTDs absent NTDs absent	(53)
Lmo4	*Lmo4-/-*	C57BL/6 C57/cd-1/129 Mixed	EX (<10%) EX (51.1%, 23/45)	(70)
Men1	*Men1-/-*	C57BL/6 (N5-9) 129S6/SvEV (N5-9)	EX Absent (0/39) EX (21%, 4/19)	(50)

18

Table 1.4.1 Continued

Noggin	Noggin-/-	129/Sv 129/Sv-C57BL/6J Mixed	NTDs ("almost always") NTDs absent ("most often")	(52)
p300	p300+/-	129/Sv 129/Sv-C57BL/6 Mixed	EX (19%, 16/79) EX (6%, 6/101)	(71)
p53	p53-/-	129/Sv C57BL/6-129/Sv Mixed	EX (16.4%, 10/61) EX (8.4%, 14/167)	(72)
Pax3 (Splotch)	Pax3-/-	Splotch Background DBA/Sp (N4) NZW/SP (N4)	SB (25%, 4/16), EX (12.5%, 2/16), SB+EX (62.5%, 10/16) SB (64.3%, 27/42), EX (14.3%, 6/42), SB+ EX (21.4%, 9/42) SB (21.4%, 9/35), EX (17.1%, 6/35), SB+EX (57.1%, 20/35)	(60)
Sall2	Sall2-/-	129Sv/J 129Sv/J-NZW Mixed 129Sv/J-CD1 Mixed 129Sv/DBA Mixed 129Sv/J-C57BL/6 Mixed	EX (11.5%, 24/208) NTDs (17.1%, 6/35) NTDs (12.9%, 8/62) NTDs Absent (0/43) NTDs Absent (0/105)	(48)
Ski	Ski-/-	129P2-Swiss Mixed 129P3 (N4-6) 129P2-C57BL/6 Mixed C57BL/6J (N14)	EX (88%, n>19), FC (12%, n>19) EX (83%, n=12), FC (0%, n=12) EX (80%, n>19), FC (15%, n >19) EX (5%, n>19), FC (93%, n>19)	(45)
Tcof1	Tcof1+/-	C57BL/6 CBA/Ca C3H DBA/1 BALB/c	EX (100%, 206/206) EX (93.9%, 31/33) EX (89.7%, 26/29) EX (12.5%, 6/48) EX (6.6%, 5/76)	(58)
Tgif	Tgif-/-	C57BL/6-129Sv (50:50) C57BL/6-129Sv (75:25) C57BL/6-129Sv (87.5:12.5) C57BL/6	EX (5.1%, 2/39) EX (20.9%, 14/67) EX (33.3%, 9/27) EX (21%, 14/67)	(59)
Tmem67	Tmem67-/-	C57BL/6J-129P2 (N6-10) C57BL/6J	EX (23.5%, 4/17) EX Absent (0/19)	(73)

Table 1.4.1 Continued

2. Identified Modifiers

Target Gene	Genotype	**Modifying Gene**	**Modifying Effect (Penetrance)[2]**	Reference
Grhl3 (curly-tail)	*Grhl3+/+*	*Lmnb1* variant 9E/9E (ct background)	NTDs absent (SB 0/113, EX 0/226)	(56)
	Grhl3+/+	*Lmnb1* variant 8E/8E (ct background)	SB Absent (0/105), EX (2.6%, 4/151)	
	Grhl3ct/ct	*Lmnb1* variant 9E/9E (ct background)	SB (5.8%, 13/224), EX (3.0%, 11/367)	
	Grhl3ct/ct	*Lmnb1* variant 8E/8E (ct background)	SB (15.5%, 23/149), EX (8.2%, 15/183)	

3. Gene-Environment Interaction Modifier[8]

Target Gene	Genotype	**Modifying Environment**	**Modifying Effect (Penetrance)[2]**	Reference
Shmt1	*Shmt1+/-* or -/-[5]	Normal diet (129SvEv)	NTDs Absent (0/39)	(37)[6]
	Shmt1+/+	FCD diet (129SvEv)	NTDs Absent (0/89)	
	Shmt1+/-	FCD diet (129SvEv)	EX (3.2%, 6/185)	
	Shmt1-/-	FCD diet (129SvEv)	EX (14.1%, 13/92)	
	Shmt1+/- or -/-[5]	Normal diet (C57Bl/6)	NTDs Absent (0/78)	
	Shmt1+/+	FCD diet (C57Bl/6)	NTDs Absent (0/167)	
	Shmt1+/-	FCD diet (C57Bl/6)	EX (0.9%, 3/334)	
	Shmt1-/-	FCD diet (C57Bl/6)	EX (4.3%, 5/117)	

4. Interacting Gene Modifications[8]

"Target" Gene[7]	Genotype	**"Modifying" Gene[7]**	**Modifying Effect (Penetrance)[2]**	Reference
Cyp26a1	*Cyp26a1-/-*	*Aldh1a2+/+*	SB Present	(74)
	Cyp26a1-/-	*Aldh1a2+/-*	NTDs Absent	
Lrp6	*Lrp6-/-*	*Wnt-5a+/+*	EX (~30%, n=9)	(49)
	Lrp6-/-	*Wnt-5a+/-*	EX (~5%, n=30)	
	Lrp6-/-	*Wnt-5a-/-*	EX Absent (n=7)	

Table 1.4.1 Continued

Pax3 (Splotch)	Pax3+/+	Nf1+/+ (129/Sv)	EX Absent (0/73)	(57)
	Pax3+/+	Nf1+/- (129/Sv)	EX Absent (0/153)	
	Pax3+/+	Nf1-/- (129/Sv)	EX (12.5%, 6/48)	
	Pax3+/+	Nf1+/+ (C57BL/6-129/Sv)	SB/EX Absent (0/9)	
	Pax3+/+	Nf1+/- (C57BL/6-129/Sv)	SB/EX Absent (0/17)	
	Pax3+/+	Nf1-/- (C57BL/6-129/Sv)	SB/EX Absent (0/7)	
	Pax3+/-	Nf1+/+ (C57BL/6-129/Sv)	SB/EX Absent (0/19)	
	Pax3+/-	Nf1+/- (C57BL/6-129/Sv)	SB (1.6%, 1/64), EX (1.6%, 1/64), SB+EX (3.1%, 2/64)	
	Pax3+/-	Nf1-/- (C57BL/6-129/Sv)	SB Absent (0/17), EX (5.9%, 1/17)	
	Pax3-/-	Nf1+/+ (C57BL/6-129/Sv)	SB (45.4%, 5/11), SB+EX (54.5%, 6/11)	
	Pax3-/-	Nf1+/- (C57BL/6-129/Sv)	SB (88.2%, 15/17), SB+EX (11.8% 2/17)	
	Pax3-/-	Nf1-/- (C57BL/6-129/Sv)	SB+EX (100%, 5/5)	

Pax3 (Splotch)	(C57/129SvEv/BALB/c mix for examples below)			(37)[6]
	Pax3+/+	Shmt1+/+	NTDs Absent (0/56)	
	Pax3+/+	Shmt1+/-	NTDs (4.2%, 2/48)	
	Pax3+/+	Shmt1-/-	NTDs Absent (0/33)	
	Pax3+/-	Shmt1+/+	NTDs Absent (0/117)	
	Pax3+/-	Shmt1+/-	NTDs (5.81%, 5/86)	
	Pax3+/-	Shmt1-/-	NTDs (4.2%, 2/48)	
	Pax3-/-	Shmt1+/+	NTDs (98%, 49/50, 22% SB+EX)□	
	Pax3-/-	Shmt1+/-	NTDs (97.1%, 33/34, 46% SB+EX)	
	Pax3-/-	Shmt1-/-	NTDs (100%, 29/29, 66% SB+EX)	

Notes:

1. Backgrounds are congenic unless otherwise stated. N: number of backcrosses.
2. Only NTD related modifying effects are recorded. Percentage values represent penetrance, provided where available. Abbreviations – SB: spina bifida, EX: exencephaly, CRS: craniorachischisis, FC: facial cleft, SB+EX: combination of spina bifida and exencephaly n: sample size, FCD: folate and choline deficient.
3. Penetrance on BALB/cCrl background has changed since it was recorded in Banting et al, 2005. This is the updated, unpublished value. See text.
4. Severe NTD: multiple NTDs observed in a single embryo (i.e. spina bifida and exencephaly)
5. Genotype of embryos not specified, only parents.
6. Data for embryos with the same genotype but differing maternal genotype are combined for this table.
7. Since the modifier and target genes have not been clearly identified, the terms are interchangeable in these examples.
8. The genes in section 3 and 4 are meant to illustrate each category and are not meant to be an exhaustive list.

1.5: Cecr2 and Neural Tube Defects

A gene that is thought to be under the influence of modifiers is *Cecr2*. CECR2 is a part of the ATP-dependent CERF (CECR2-containing Remodeling Factor) chromatin remodeling complex, which also includes the catalytic ISWI protein SNF2L (*47, 75*). *Cecr2* is found on chromosome 22q11 of humans (chromosome 6 of mice) and encodes a protein roughly 164kDa (1484 amino acids, 19 exons) in size that has been associated with transcriptional regulation (*47, 76*). *Cecr2* is also implicated with double strand break repair in HEK293 cells (*77*). The CECR2 protein is made up of a DDT domain, an AT-hook, and a bromodomain (Figure 1.5.1). The bromodomain in CECR2 is involved in chromatin remodelling through the recognition of acetylated lysine residues. Post-translational modifications of histones allow for changes in chromatin structure and gene expression. This includes methylation, acetylation, ubiquitination, and phosphorylation (*78*). The protruding N-terminal and C-terminal of core histones are targets for modification (*79*). Acetylation of histones occurs on lysine residues of the N-terminal, causing the chromatin to become less condensed due to a change in charge. Acetylation often leads to the expression of proliferation-promoting genes and the inactivation of tumor suppressors (*78*). Acetylated lysine residues are recognized by only proteins containing a bromodomain interaction module, such as CECR2 (*80*). The AT-hook domain binds to DNA while the DDT domain binds to the ISWI component of the complex (*81–83*). A nuclear localization sequence is also found in exon 5 of CECR2 (*47*). During development, expression of CECR2 has been observed throughout the nervous system including the neural tube, limbs, eyes, and nasal epithelium (*47, 84*). Similar expression patterns have been observed in chick embryos (*85*).

A *Cecr2* mutant mouse line on a mixed 129P2/BALB background was produced by inserting a pGT1 gene-trap mutation between exon 7 and 8 (Figure 1.5.1) (*47*). Analysis of the mutation on a congenic (F6) BALB/cCrl background, indicated that 74% of homozygous *Cecr2* (*Cecr2^Gt45Bic/Gt45Bic*) mutants pups develop exencephaly (*47*). Heterozygous mutants do not develop exencephaly. A more recent penetrance analysis indicated that 54% of BALB/cCrl embryos with the homozygous gene-trap mutation develop exencephaly (*55*). In embryos with exencephaly, there is a high expression of *Cecr2^Gt45Bic/Gt45Bic* in neural folds that did not close.

A presumptive null allele *Cecr2* mutant was also produced in mice, termed *Cecr2^tm1.1Hemc* (*76*). The *Cecr2^tm1.1Hemc* mutation was produced through deletion of exon 1 and about 1000 base pairs upstream using Cre-Lox recombination (Figure 1.5.1). Exencephaly is observed at a rate of 96% in homozygous *Cecr2^tm1.1Hemc* BALB/cCrl mutants. Additionally, the null mutation leads to midline facial clefts and forebrain encephalocele in some embryos.

CECR2 deficiency in mice has also been linked to NTDs through mutation of the *Tet1* gene (*86*). TET proteins are involved in the enzymatic demethylation of 5-methylcytosine, thereby allowing transcription to proceed. The majority of mice deficient in TET1 protein develop normally, likely due to compensation from TET2 protein. Double knockouts of *Tet1* and *Tet2* in mice lead to a variety of developmental defects, including 13% of embryos having exencephaly (*87*). However, a truncating mutation to just *Tet1* (*Tet1^tuft*) that disrupts the catalytic domain leads to exencephaly and other cranial defects in mice (*86*, *88*). It has been proposed that the truncated TET1^tuft protein is still binding to genes involved in neurulation, though it is causing abnormal expression due to its inability to demethylate the DNA. If TET1^tuft is still binding, TET2 will not have the ability to bind to TET1 targets and compensate for the loss of function. Interestingly, mice with this *Tet1^tuft* mutation also show decreases in *Cecr2* and *Epha7*

expression (*86*). It had been previously shown that *Epha7* expression decreases due to an absence of *Cecr2* in mice (*76*). Other genes that have been associated with NTDs, such as *Grhl2*, also show changes in expression in *Tet1*[tuft] mutants (*86, 89*). Furthermore, a recombinant *Tet1*[tuft] gene was produced and transfected into mouse IMCD-3 cells. Relative to control cells transfected with wildtype *Tet1*, there is a significant decrease in *Grhl2* expression in TET1[tuft] transfected cells, providing support for gene regulation abnormalities (*86*). *Cecr2* expression was not measured in cells.

	Cecr2^GT45bic/Cecr2^GT45bic	Cecr2^tm1.1Hemc/Cecr2^tm1.1Hemc
Balb/cCrl	54%	96%
FVB/N	0%	12.3%
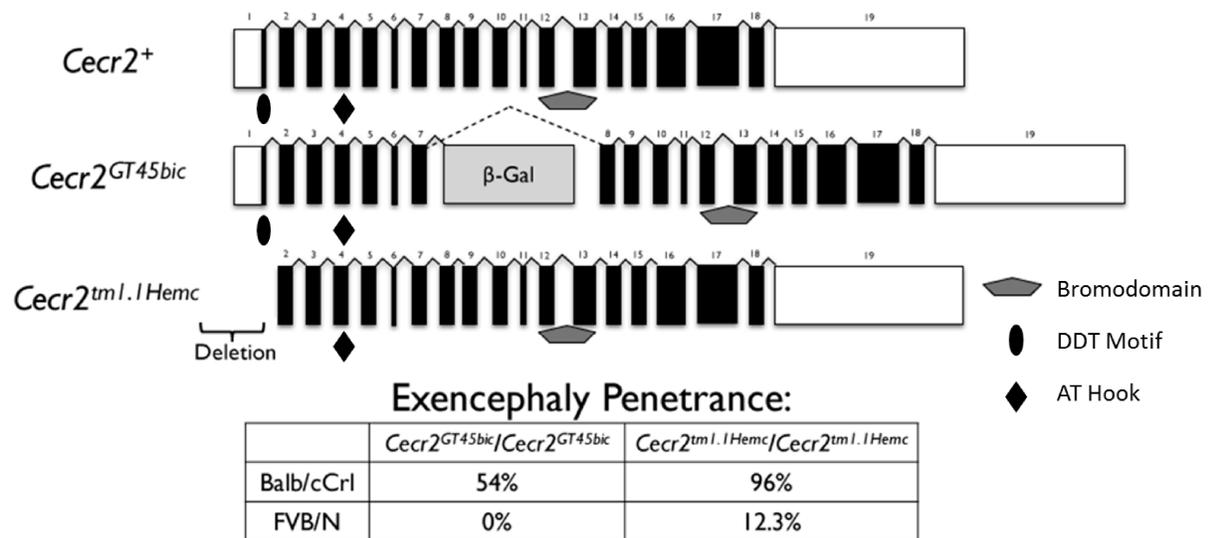
Figure 1.5.1: Full length *Cecr2* and *Cecr2* mouse mutant genes. *Cecr2* is composed of 19 coding exons (black bars, note: coding portion of exon 1 and 19 is small, white bars represent non-coding regions). Key domains include a bromodomain, DDT motif, and AT hook. The *Cecr2*^GT45bic mutation contains a β-galactosidase (β-Gal) inserted in between exon 7 and 8. This gene-trap mutation is hypomorphic due to the low level splicing around the β-gal insert (dotted lines). The *Cecr2*^tm1.1Hemc is more severe as it is thought to lead to a null allele. Exon 1 and about 1000 base pairs upstream has been deleted. The penetrance of exencephaly for each mutation on the BALB/cCrl and FVB/N lines is shown. Figure modified from Leduc et al. (*55*).

1.6: Modifier Loci Affecting Cecr2

The exencephaly phenotype due to the *Cecr2* gene-trap mutation occurs in a strain-dependent manner. It was first reported that BALB/cCrl mice with a homozygous *Cecr2* mutation developed exencephaly about 74% of the time (n=47) (*47*). A more recent penetrance analysis with BALB/cCrl mice indicated that the penetrance of exencephaly is 54% (n=135) with the *Cecr2^Gt45Bic/Gt45Bic* mutation (*55*). Moving the mutation onto the FVB/N genetic background reduces the penetrance of exencephaly in homozygous *Cecr2* mutants to zero (n=45) (*47*). The *Cecr2^tm1.1Hemc* homozygous null mutation leads to an exencephaly penetrance of 96% (n=50) in BALB/c, while the null mutation on the congenic FVB/N genetic background leads to a penetrance of 12% (n=73) (*55*). The difference in penetrance between the BALB/cCrl and FVB/N strains indicated the presence of a single or multiple variant modifier gene(s). One variant of a modifier gene could confer exencephaly resistance in FVB/N mice, while another variant confers exencephaly susceptibility in BALB/cCrl mice.

Analysis of cranial neural folds in FVB/N and BALB/cCrl indicated that embryos from both strains with *Cecr2^Gt45Bic/Gt45Bic* have abnormally wide neural folds in the cranial region relative to wildtype embryos (*84, 90*). The neural fold width does not differ significantly between wildtype BALB/cCrl and FVB/N mice. Heterozygous mutant embryos of both strains do not have significantly wider neural folds but do have a delay in the closure of those neural folds (*90*). The hypothesis is that FVB/N mice with the homozygous *Cecr2* mutation are able to overcome the wide neural folds and close, whereas the BALB/cCrl mice are not able to close the neural tube resulting in exencephaly. Inner ear defects and open eyelids are also observed in mutant BALB/cCrl embryos but not in FVB/N mutants (*47, 84*).

To further characterize the penetrance difference of exencephaly due to the *Cecr2* mutation and identify possible modifier loci, a whole genome linkage analysis was conducted (*91*). Unaffected *Cecr2*^{Gt45Bic/Gt45Bic} congenic FVB/N mice were crossed with *Cecr2*^{Gt45Bic/+} congenic BALB/cCrl mice. The resulting mutant embryos (F1) show a 2.9% (1/35) penetrance of exencephaly, indicating dominance of some or all of the FVB/N resistant alleles (*91*). The F1 were then backcrossed to *Cecr2*^{Gt45Bic/+} BALB/cCrl mice. The penetrance of exencephaly in the resulting homozygous mutant embryos is 28.1%. It was predicted that a single dominant modifier would have a penetrance of about 37%, thus, it is likely that there are multiple modifiers segregating separately (*91*). A significant (P≤0.001, LOD 4.35) linkage peak was mapped on chromosome 19 located at roughly 40.1MB, ranging 30MB. This provides evidence for at least one significant modifier gene in the chromosome 19 modifier region.

The generation of two sub-interval congenic mouse lines, in which a portion of the modifier region from the resistant FVB/N strain was crossed onto susceptible BALB/cCrl mice with *Cecr2*^{Gt45Bic/Gt45Bic}, indicate that more than one modifier gene is found in the modifier region (*90*). A simplified explanation of the experiment is described here. Two lines were produced that each contained non-overlapping fragments (MOD4 and MOD31). A significant drop in exencephaly penetrance is measured in each line (about 30% penetrance in each) (*55*, *90*). This indicates that there is at least one modifier in each of the non-overlapping segments of the modifier region.

Furthermore, mice were produced in which the majority of chromosome 19 modifier loci was either homozygous for the FVB/N alleles or heterozygous for FVB/N and BALB/cCrl alleles while the rest of the background was BALB/cCrl (*55*). Since the FVB/N alleles were shown to be dominant, an additive modifier gene effect would lead to a significant drop in

exencephaly penetrance relative to the MOD4 or MOD31 line. However, the penetrance of exencephaly is similar to MOD4 and MOD31. This indicated that the FVB/N resistant modifier genes in chromosome 19 are not additive. Additionally, since penetrance does not decrease to 2.9% similar to when BALB/cCrl and FVB/N mice were crossed, there are almost certainly other modifiers in other regions of the genome since it was homozygous for the BALB/cCrl alleles.

1.7: Candidate Modifier Genes of Cecr2

The identified modifier locus on chromosome 19 of mice contains hundreds of genes. Since it was not feasible to study every single gene in the region, it was necessary to narrow down the number of genes to a subset of candidates that were more likely to be modifiers of *Cecr2*. To begin this process, a microarray analysis was conducted on BALB/cCrl and FVB/N neurulating embryos with and without the *Cecr2$^{Gt45Bic/Gt45Bic}$* mutation to discover differentially expressed genes during neurulation (*90*). Using two different statistical methods, genes were found that have expression differences between the strains and were validated with qRT-PCR. After validation, a total of nine genes were shown to have significant expression differences between the strains of mice within the modifier region.

Whole exome sequencing was also conducted on a BALB/cCrl and FVB/N female to identify single nucleotide variants (SNVs) or insertions or deletions (INDELs) differing between the two strains. About 2450 SNVs in 371 genes and 430 INDELs in 147 genes were found (*92*). This list was narrowed down to 19 genes by only including genes found in the modifier region. Synonymous and non-coding SNVs and INDELs were filtered out of the list. To prioritize the list of genes and variants, *in silico* analyses were conducted to predict damaging variants. Three tests in particular were utilized: SIFT (Sorting Intolerant from Tolerant), GERP (Genomic

Evolutionary Rate Profiling) and PP2 (PolyPhen-2). SIFT predicts the damaging effect of an amino acid substitution based on amino acid conservation (*93*), GERP analyzes nucleotide conservation (*94*), and PP2 predicts damaging effects of variants based on physical/property changes and conservation of amino acids (*95*). Based on the microarray and whole exome sequence analyses followed by *in silico* predictive analyses, a list of 24 prioritized candidate genes from mouse was produced (*92*).

Since the overall goal of this research is to discover NTD susceptibility genes in humans, the homologues of the 24 mouse genes were then sequenced in a Duke University cohort of 156 unrelated human probands with cranial NTD (*92*). Variants leading to protein coding changes in the 24 genes were identified. Missense, nonsense, and frameshift mutations in the coding regions of the genes were compiled. Since rare variants are more likely to be deleterious, minor allele frequencies (MAFs) were considered (ESP6500 and 1KG databases). Variants with a MAF score of less than 0.03 were considered to be more likely deleterious, thus they were included in the candidate list. In total, 43 variants were found fitting the mentioned criteria, 10 of those variants being novel. Additionally, 10 variants were included for further analysis that had MAF scores higher than 0.03; these variants were included because they may contribute to compound heterozygosity in some probands. Compound heterozygotes are characterized with having two different potentially deleterious mutations within the same gene. After analysis of MAF, GERP, SIFT and PP2 scores of each variant and consideration of gene functions, it was determined that *DNMBP* was the top candidate gene to further analyse, followed by *ZO-2* and *MMS19*. One mouse variant of *Dnmbp* from the BALB/cCrl strain was predicted to be damaging by SIFT and PP2. Three human *DNMBP* variants were predicted to be damaging and two possibly damaging

by PP2. GERP predicted six human *DNMBP* variants to be damaging. Variants of *DNMBP* are further discussed in section 1.13.

1.8: Overview of Dynamin Binding Protein (DNMBP)

  DNMBP is a large gene found on chromosome 10q24.2 in humans (chromosome 19 in mice) encoding a 1577 amino acid scaffold protein with a molecular mass of about 177kDa. The DNMBP protein, also known as Tuba, interacts directly with dynamin as well as with various actin regulatory proteins (*96*). Dynamin is a GTPase that is associated with the fission of endocytic buds, such as clathrin coated pits, from the plasma membrane (*97*). While the precise mechanism of vesicle release from the membrane by dynamin is not fully understood, there is evidence supporting the involvement of the actin cytoskeleton (*98, 99*). For instance, disruption to dynamin leads to disruptions in actin dynamics, while disruptions to actin leads to abnormal endocytosis (*97*). Since DNMBP interacts with both dynamin and actin-associated proteins, it may be one of the links that connects the two groups.

  There are multiple domains that make up the DNMBP protein (Figure 1.8.1). Four Src homology-3 (SH3-1 to -4) domains are found on the N-terminus end, which bind dynamin. There are also two SH3 domains (SH3-5 and -6) found on the C-terminal end of DNMBP. The very last SH3 domain (SH3-6) pulls down a variety of actin regulatory proteins as well as actin itself (*96*). It is not known if all of the proteins pulled-down by the SH3-6 domain bind directly to DNMBP or if an intermediate is involved. There are two other major domains found in the center of DNMBP – the RhoGEF and BAR domains. The RhoGEF (Rho guanine nucleotide exchange factor) domain, sometimes referred to as the Dbl homology (DH) domain, is found in proteins that positively regulate GTPases from the Rho family (*100*). The Rho family of GTPases, which

includes CDC42, Rac1, and RhoA, act as molecular switches in cells regulating cytoskeletal dynamics, vesicle movement, cytokinesis, cell polarity, and gene expression (*101*). GTPases are active when bound to GTP and inactive when bound to GDP. RhoGEF domains activate GTPases by dissociating bound GDP, allowing GTP to bind (Figure 1.8.2). GAPs (GTPase-activating proteins) inactivate GTPases by accelerating the hydrolysis of GTP to GDP (*102*). GDIs (guanine nucleotide dissociation inhibitors) can also bind to Rho GTPases and prevent the activation of the GTPase (*101*). The RhoGEF domain of DNMBP is specific for the CDC42 (Cell Division Cycle 42) GTPase (*96*). The BAR (Bin/amphiphysin/Rvs) domain is also associated with the GEF activity of DNMBP. Almost all RhoGEFs contain PH (pleckstrin homology) domains that are utilized for localization of RhoGEFs to plasma membranes and modulation of GEF activity (*103*). DNMBP is an exception as it contains a BAR domain rather than a PH domain. BAR domains are found in several other dynamin interacting proteins and are used for binding as well as bending plasma membranes (*97*, *104*). BAR domains form "banana" shaped structures that sense and bind to curved membranes through interaction of the positively charged residues at the tips of the structure and negatively charged phospholipids (*97*). Thus, it is thought that the BAR domain fulfils the role of the PH domain when it comes to anchoring the GEF to plasma membranes. The BAR domains of some proteins, such as Arfaptin-2, have also been shown to bind directly to GTPases (*104*). The poor solubility of the DNMBP BAR domain has prevented the analysis of proteins that the domain may bind (*97*, *105*, *106*).

There are two known protein homologues of DNMBP named Tuba2 and Tuba3, both of which are comprised of the C-terminal half of DNMBP with a molecular mass of about 75kDa each (*96*). Tuba2 and Tuba3 are 60% and 41% similar to DNMBP, respectively. The gene encoding Tuba2, located on chromosome 4 of humans (chromosome 3 of mice), results in a

protein that begins in the middle of the RhoGEF domain (Figure 1.8.1). Tuba3, located on

chromosome 5 of humans (chromosome 8 of mice), begins a bit before the RhoGEF domain.

Thus, the RhoGEF domain is likely not functional in Tuba2, but is in Tuba3. Since dynamin is

one of the known interactors of the N-terminal half of DNMBP, it is possible that Tuba2 and

Tuba3 are utilized only for actin regulation.

Expression of DNMBP is strongest in testis, followed by the brain, heart, liver, spleen,

and lung (*96*). Splice variants are also detectable at sizes of about 105kDa, 94kDa and 75kDa

and expressed differently, though the 75kDa protein could also be Tuba2 or Tuba3. In mice, it

has been confirmed through immunofluorescence staining that DNMBP is expressed in

neuroepithelium during neurulation (*92*).
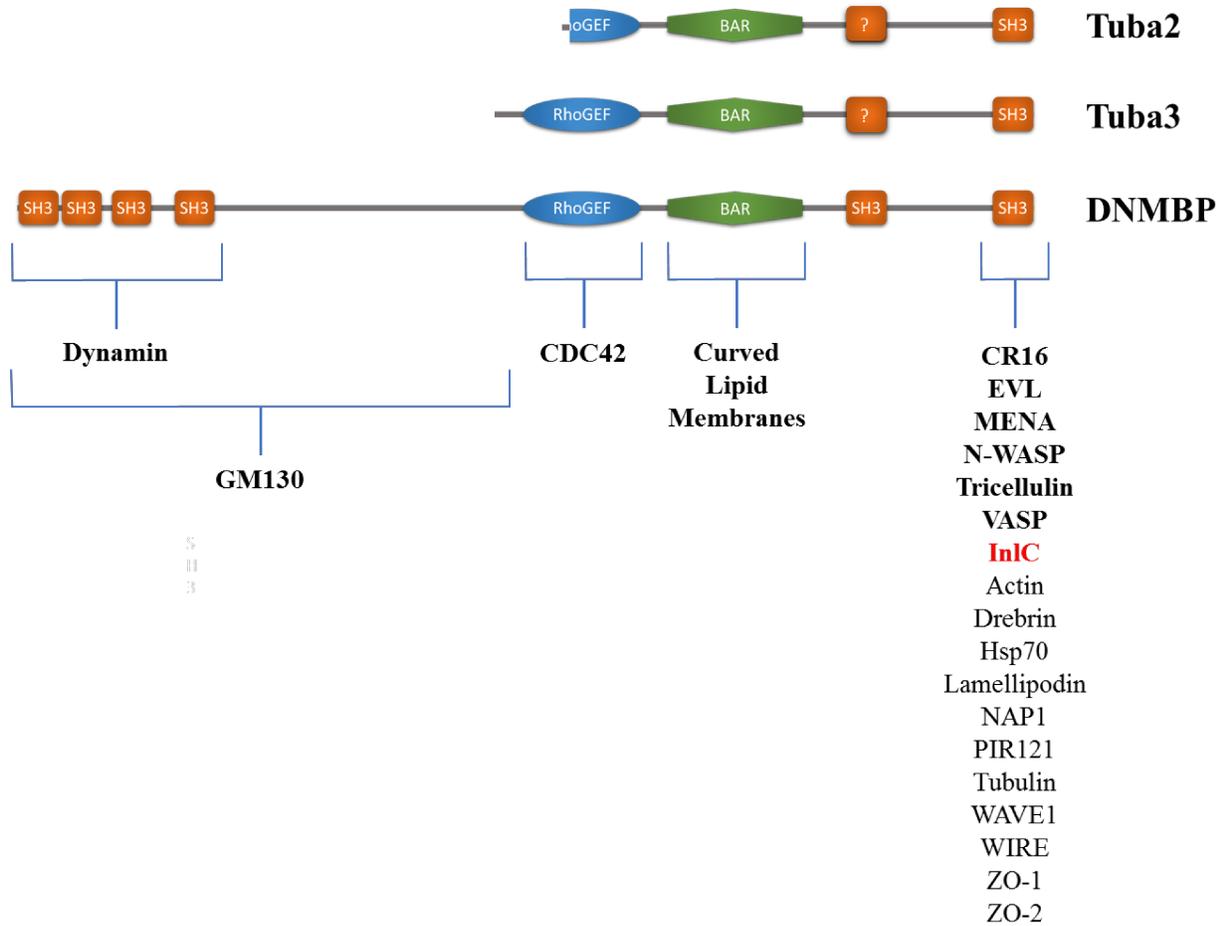
Figure 1.8.1: Protein Domains of DNMBP and homologues of DNMBP. DNMBP is composed of six Src homology 3 (SH3) domains (orange), a RhoGEF domain (blue), and a Bin/amphiphysin/Rvs (BAR) domain (green). There are two homologues of DNMBP, Tuba2 and Tuba3. The presence of the fifth SH3 domain in the homologues has not been confirmed. Predicted interactors of DNMBP have been listed with the predicted site of binding. Bold text indicates confirmed direct binding to specified region of DNMBP. InlC (red) is a bacterial virulence factor that inhibits some DNMBP function.
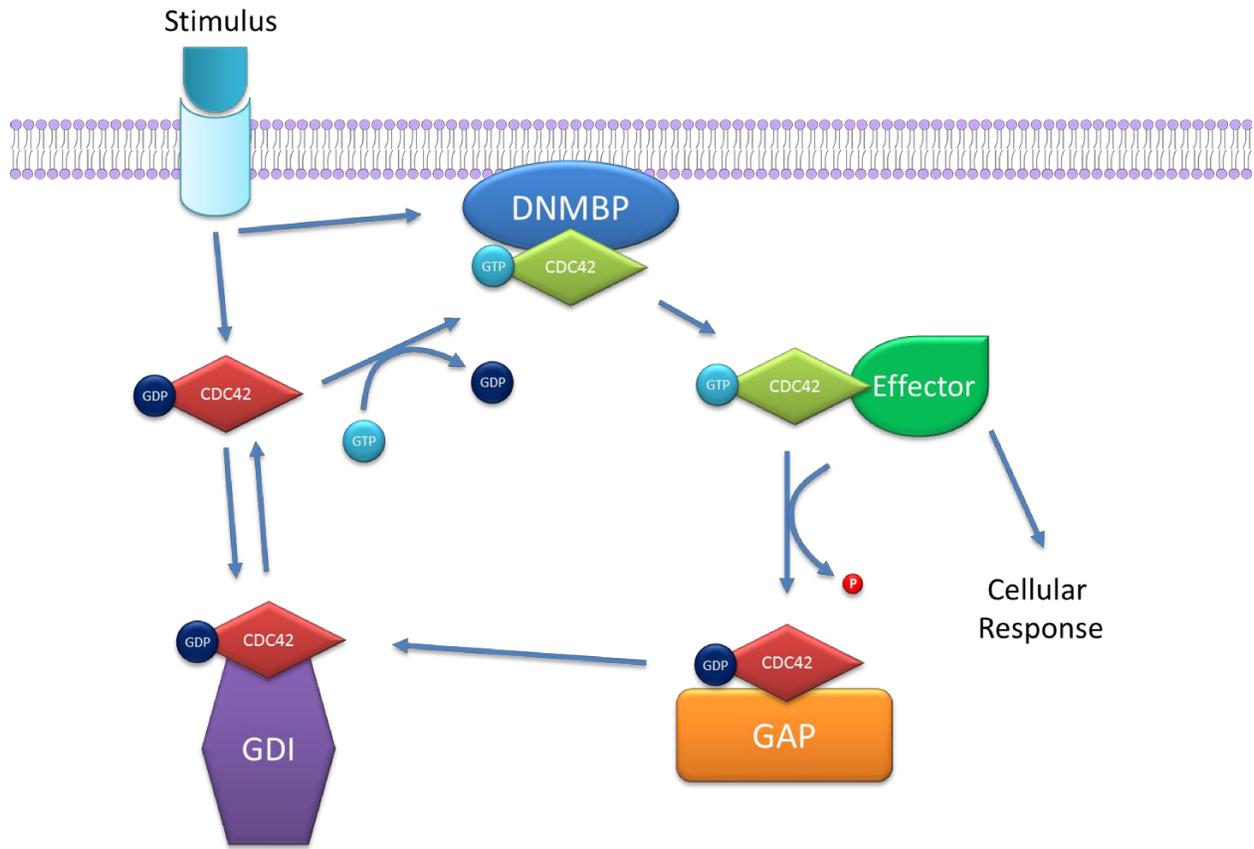
Figure 1.8.2: Activation of Rho GTPase CDC42. A stimulus effects DNMBP and causes CDC42 to become active by binding with DNMBP, leading to an exchange of GDP for GTP. Active GTP-bound CDC42 will then activate effectors leading to cellular changes. If CDC42 binds to a GTP activating protein (GAP), the dissociation of GTP to GDP will be accelerated, leading to GDP-bound CDC42. To prevent further activation, GDP-bound CDC42 can bind to guanosine nucleotide dissociation inhibitors (GDI).

1.9: DNMBP Interactions

DNMBP has multiple domains that bind a variety of proteins. Incubation of a GST fusion protein containing the first four SH3 (SH3-1 to -4) domains of DNMBP with rat brain lysate results in the pull-down of dynamin (*96*). The interaction between DNMBP and dynamin was confirmed to be direct using far-Western blotting. When dynamin is immunoprecipitated from rat brain lysate, endogenous DNMBP is co-precipitated. Individual analysis of each SH3 domain indicated that SH3-4 binds dynamin the strongest, followed by SH3-1, SH3-3, and SH3-2 (*96*). It has been confirmed with other SH3-containing proteins that the proline/arginine-rich domain (PRD) of dynamin binds to SH3 domains (*107*). Thus, it is likely that the PRD of dynamin binds to DNMBP SH3 domains, though this has not been confirmed. *In vivo* interaction of DNMBP and dynamin was further demonstrated in Chinese hamster ovary cells. An HA-tagged protein containing the first four DNMBP SH3 domains was overexpressed to examine the effects of transferrin uptake. Transferrin is an iron-binding glycoprotein that enters a cell through receptor mediated endocytosis, a process involving dynamin (*108, 109*). If dynamin is unavailable due to its PRD being occupied by an isolated SH3 domain from amphiphysin protein (without the rest of the protein), it does not function normally, resulting in an inhibition of receptor-mediated endocytosis and a decrease in the uptake of transferrin (*110*). This decrease in transferrin uptake was observed when the first four DNMBP SH3 domains were overexpressed indicating that they are in fact binding to dynamin (*96*). As expected, the C-terminal SH3 domains of DNMBP, which have not been shown to bind dynamin, do not lead to a decrease in transferrin uptake when expressed in cells.

The C-terminal SH3 domain (SH3-6) interacts with actin associated proteins rather than dynamin. SH3-5, another domain close to the C-terminal, does not pull down any proteins (*96*).

Incubation of a GST fusion protein consisting of DNMBP SH3-6 with rat brain extract results in the pull-down of Lamellipodin, Drebrin, PIR121, NAP1, WAVE1, MENA, VASP, Hsp70, N-WASP, CR16, WIRE, Tubulin, and Actin (*96*). Since this was a simple pull-down experiment with just the SH3-6 domain of DNMBP, it is not known if these proteins all bind directly to DNMBP. It is also possible that the specificity of SH3-6 is reduced when the rest of DNMBP is not present. DNMBP has not been identified as an interactor of actin, however, some proteins that were pulled-down, including Drebrin, WAVE1, VASP, MENA, and N-WASP are all actin binding proteins (*111–113*). Thus, actin was likely pulled-down with those proteins. Pull-down of some proteins was further characterized (*96*). The immunoprecipitation of DNMBP from embryonic day 15 mouse lysate confirmed the co-immunoprecipitation of MENA and N-WASP (*96*).

MENA (Mammalian Enabled, ENAH) is a part of the Ena/VASP family of proteins, which also includes VASP (Vasodialtor-stimulated phosphoprotein) and EVL (ENAH/VASP-like) (*114*). Ena/VASP proteins directly regulate actin assembly, and consequently play a role in cell movement and morphology. To characterize DNMBP interaction with Ena/VASP proteins, ENA/VASP-deficient cells were utilized (*96*). The cells were produced from rat embryo fibroblasts in which Ena/VASP genes are knocked out (*115*). These cells were infected with either MENA, VASP, or EVL along with the SH3-6 domain of DNMBP. The SH3-6 domain was found to bind to all three proteins. Furthermore, the common proline-rich region found in MENA, VASP, and EVL was shown to be necessary for the interaction between MENA and DNMBP. Using a yeast two-hybrid system, the direct interaction between the SH3-6 domain of DNMBP and Ena/VASP proteins was confirmed. Additionally, purification of proteins and

detection of binding was also conducted and confirmed the direct binding of MENA, N-WASP and CR16 to the isolated SH3-6 domain of DNMBP (*96*).

Tricellulin is another actin-associated direct interactor of DNMBP at the SH3-6 domain (*116*). Tricellulin is a protein that localizes to apical tricellular contacts of epithelial cells and regulates F-actin organization through DNMBP and CDC42. The cell-cell borders of epithelial cells, which are usually straight and rigid, become irregular and curved with either a Tricellulin or DNMBP deficiency, indicating abnormalities with actin organization. Specifically, F-actin density at cell-cell junctions is reduced when DNMBP is knocked down or when CDC42 activity is suppressed (*117*). The co-localizing partner ZO-1 (TJP1) of DNMBP has also been shown to interact with DNMBP. The SH3-6 domain of DNMBP is necessary, but is not sufficient, for the interaction with ZO-1 (*117*). The site of interaction on ZO-1 has not been investigated, although it does contain a proline-rich domain. Thus, like other proteins that bind to SH3-6 of DNMBP, the proline-rich domain of ZO-1 might be the binding site. Interestingly, ZO-2 (Zonula occludens-2, TJP2, Tight junction protein-1), a homologue of ZO-1 and one our top candidate modifier genes, also interacts with DNMBP. These tight junction proteins are necessary for F-actin organization (*118*). It is not known if ZO-1 and -2 bind directly to DNMBP. ZO-1 and ZO-2 bind to each other (*119*), therefore it is possible that only one binds to DNMBP, and the other is being pulled-down as well.

The Golgi apparatus-attached protein GM130 is a DNMBP interacting protein that is associated with dynamin rather than actin dynamics. GM130 plays a role in vesicle fusion to the Golgi membrane, glycosylation, cell polarization, and maintenance of the Golgi during mitosis (*120*). Depletion of GM130 results in abnormal centrosomes during interphase, and multipolar spindle during mitosis in different human cell lines; cells do not survive past metaphase (*121*).

Evidence supports that the abnormal centrosome phenotype observed with loss of GM130 is due to a decreased activation of CDC42 by DNMBP (*122*). DNMBP or CDC42 deficiencies result in similar abnormal mitotic phenotypes as GM130 depletion, which can be rescued with the expression of constitutively active CDC42. Pull-down of GM130 leads to the co-precipitation of DNMBP in HeLa cells. It is thought that GM130 binds somewhere in the N-terminal region of DNMBP since GM130 did not pull-down the smaller homologues of DNMBP (Tuba2 and Tuba3) that only contain the C-terminal half of DNMBP. However, if GM130 does bind to the C-terminal, it is possible that a difference in conformation between full length DNMBP and Tuba2 or Tuba3 prevents binding. *In vitro* binding assays indicated that recombinant HA-DNMBP directly interacts with the N-terminal region of GM130 (*122*). When GM130 was knocked down through RNAi in U-2 OS cells, the amount of DNMBP protein remains the same in cells but there is a 50% reduction in active CDC42, indicating that GM130 influences the RhoGEF activity of DNMBP. Figure 1.8.1 summarizes DNMBP interactions.

1.10: The Activation of CDC42 by DNMBP

A third major region of DNMBP interaction, the RhoGEF domain, is found in the center of the protein. This domain functions as a specific activator of the GTPase CDC42. The RhoGEF domain contains three conserved regions (CR1-CR3) at the core of the domain (*103*). Generally, CR1 and CR3 of CDC42 GEFs interact with CDC42 Switch regions – regions of CDC42 that change in conformation based on triphosphate/diphosphate-bound states. There are two Switch regions: Switch 1 is composed of residues 25-39 of CDC42 and Switch 2 is composed of residues 57-75. This interaction has not been confirmed for DNMBP–CDC42, though it is well conserved among other RhoGEFs and GTPases. Binding of the RhoGEF domain to CDC42

results in significant conformational changes of the nucleotide-binding pocket (*103*). The

nucleotide-binding pocket conformations are highly conserved among GTPases. The

conformation change causes the release of bound GDP, allowing GTP to bind to the pocket,

thereby activating the GTPase. While there are multiple GEFs that can activate CDC42, DNMBP

is a major activator. This has been shown in U-2 OS cells that are deficient in DNMBP; when

DNMBP expression is reduced to ~5% by RNAi, the active GTP-bound form of CDC4 is

reduced by 80% (*96*, *122*).

Once active, CDC42 can activate more than 20 downstream effector and adaptor proteins

regulating a variety of cellular processes such as cell proliferation, filopodia formation, motility,

polarity, actin cytoskeleton reorganization, cytokinesis, growth, and cell adhesion (*101*, *123–*

126). Downstream effectors of CDC42 include protein and lipid kinases, phospholipases,

scaffold proteins, cytoskeletal proteins, and actin cytoskeleton structures. CDC42 has been

studied extensively in many different cell and tissue types and model organisms. Abnormal

CDC42 expression/activation leads to detrimental phenotypes and diseases, including ciliary

abnormalities in mouse kidneys leading to renal failure (*127*), cartilage abnormalities in mice

(*128*), cancer (*124*), atherosclerosis (*129*), and craniofacial and cardiovascular defects (*130*),

demonstrating CDC42's importance. To narrow down the scope, here I will focus on what is

known about the effects downstream of CDC42 in the presence of abnormal DNMBP activity.

Out of 70 GEFs tested in MDCK (Madin-Darby canine kidney) epithelial cells, DNMBP

is one of six required for the formation of a cyst/lumen (lumenogensis) (*131*). Under the right

conditions epithelial cell lines such as MDCK and Caco-2 can be grown in culture to form 3D

cysts. Knockdown of DNMBP through shRNA results in reduced and abnormal lumen formation

(multiple lumen forming rather than one) (*131*, *132*). This is likely due to the inactivation of

CDC42 since CDC42 knockdown also results in similar lumenogensis defects and the RhoGEF domain of DNMBP is necessary for lumenogensis (*131*, *133*). The abnormal phenotype can be rescued by nucleofected DNMBP, but not by other CDC42 GEFs, indicating that lumen formation is specifically regulated by DNMBP though CDC42. A constitutively active form of CDC42 also results in rescue of lumenogensis.

The defect in lumenogensis due to the inactivation of CDC42 is thought to be the cause of improper spindle orientation. During lumen formation, the apical surface of epithelial cells are orientated by the mitotic spindles towards the center of the cyst as cells divide (*133*, *134*). Deficiencies in CDC42 result in abnormal spindle positioning. In control Caco-2 epithelial cells, most spindles are oriented perpendicular to the center of the cyst (*133*), whereas the CDC42 deficient cells that managed to produce a cyst structure have spindles randomly oriented resulting in cells dividing within the lumen. Deficiency in DNMBP result in similar spindle orientation abnormalities (*131*). Active CDC42 binds to the polarity protein Par6, which then allows for the activation of atypical protein kinase C (aPKC), which is involved in spindle orientation (*135*). Measuring phosphorylation of aPKC, which indicates activity, in CDC42 or DNMBP deficient cells indicated a decrease in aPKC activity and downstream target phosphorylation (*131*).

Defects upstream of DNMBP also results in lumenogensis abnormalities. When the oncogenic version of *Shp2* (Shp2-E76G) was expressed in MDCK epithelial cells, CDC42 is suppressed and lumen formation is abnormal (*136*). Further investigation indicated that Shp2 is dephosphorylating DNMBP (residues Y430 and Y515), leading to a reduced GEF activity and CDC42 activation.

Along with CDC42 and aPKC, Par3 and Par6 form the Par polarity complex that localizes to primary cilia and is used in ciliogensis (*137, 138*). In MDCK cells, ciliogensis is inhibited when CDC42 is deficient (*139*). Kidney-specific deficiencies of CDC42 in mouse and zebrafish also display signs of abnormal ciliogensis (*127*). Upstream of CDC42, knockdown of DNMBP results in the absence of cilia in MDCK cells that were differentiated to grow cilia (*137*). Cell polarity is not affected by DNMBP knockdown in these differentiated cells, thus, this was not an indirect effect. In zebrafish, DNMBP localizes to ciliated organs. DNMBP was knocked down in zebrafish embryos leading to downward-curled tails, ciliary defects, small eyes, pericardial edema, and hydrocephalus (*137*). When low doses of DNMBP and CDC24 antisense morpholino oligonucleotides are injected individually, a small percentage (less than 8%) develop defects. However, co-injection of both morpholinos at the same concentrations as the low doses results in 68% of embryos developing abnormal phenotypes. Thus, this additive effect shows that DNMBP and CDC42 are involved in ciliogensis in the same pathway. Morpholinos against two other GEFs did not result in these phenotypes.

In epithelial cells, DNMBP and CDC42 also have a role in cell-cell junctional configuration (*117*). DNMBP localizes to the apical cell-cell junctions of epithelial cells. Under normal circumstances, epithelial cells produce a honeycomb-like pattern with relatively straight cell-cell boundaries. Depletion of DNMBP results in slackened/curved cell-cell boundaries. Expression of DNMBP without the RhoGEF domain also leads to abnormal cell-cell boundaries, indicating that CDC42 is involved with junctional configuration. This was confirmed through the expression of a dominant-negative form of CDC42 that led to slackened membrane boundaries. Further analyses indicated that F-actin and E-cadherin organization is irregular in DNMBP deficient cells. Since both DNMBP and CDC42 interact with proteins involved in actin

dynamics, disruptions to F-actin at membrane boundaries is likely the cause of abnormal cell-cell

boundaries. Depletion of N-WASP, an actin regulatory protein that binds directly to the SH3-6

domain of DNMBP and is also an effector of CDC42, results in a similar phenotype as DNMBP

or CDC42 depletion. Binding of CDC42 to N-WASP alleviates the auto-inhibition state of N-

WASP, which then allows for the activation the Arp2/3 complex involved in actin nucleation

(*96*, *140*). Since DNMBP binds both active CDC42 and N-WASP, DNMBP likely functions to

bring these two proteins together for activation of N-WASP and subsequent actin nucleation

(*141*).

DNMBP has been associated with lipid vesicle movement due to its interaction with

dynamin and some actin regulatory proteins (*96*, *142*). More recently, DNMBP has also been

linked to lipid vesicles through its RhoGEF domain. Sato et al. (*143*) have shown that CDC42

regulates the production of dense-core vesicles, as well as the fusion of these vesicles to the

membrane (*143*). Dense-core vesicles are found only in neural cell types and some cells with

neural origins, budding from the Golgi complex and carrying neuropeptides or hormones (*143*,

144). The components of the membranes of dense-core vesicles differ from regular synaptic

vesicles (*144*). In PC12 (rat adrenal medulla pheochromocytoma) cells, expression of

constitutively active CDC42 leads to increased exocytosis of dense-core vesicles and an

increased number of these vesicles in the cytoplasm (*143*). Constitutively active CDC42 contains

a G61L mutation that prevents dissociation of GTP. Increasing the expression of wildtype

CDC42 also leads to an increase in dense-core vesicle numbers and exocytosis, but less

significantly than constitutively active CDC42. As expected, overexpression of DNMBP through

transfection leads to similar results due to the increased activation of CDC42. Inhibition of

CDC42 through expression of an inhibitory domain of neuronal N-WASP decreases the number

of vesicles relative to control cells. The mechanism of biogenesis of dense-core vesicles by CDC42 is not known, though it could be related to F-actin regulation as F-actin has been implicated with exocytosis (*143*, *145*).

1.11: Cellular Localization of DNMBP

DNMBP is expressed in almost all tissue types. The cellular localization of DNMBP is quite diverse, though it is usually related with the localization of CDC42 or actin/dynamin. In the brain, DNMBP is concentrated in areas where clathrin-mediated endocytosis occurs (*96*). DNMBP also localizes to the Golgi complex and areas with concentrated presynaptic actin in the brain. The localization of DNMBP at the Golgi has been confirmed in U-2 OS cells, in which DNMBP also localizes to the plasma membrane, cytosol, ER exit sites, and pericentriolar regions (*122*). In B16 melanoma cells, DNMBP has been observed to localize to the head and actin tail of lipid vesicles (*142*). In zebrafish, DNMBP localizes to ciliated organs (*137*). Within ciliated cells, DNMBP localizes to ciliary areas to activate CDC42.

Localization of DNMBP in epithelial cells has been well studied by Otani et al. (*117*). As described in the previous section, epithelial cell lines such as Caco-2 form honeycomb-like patterns with very straight and taut cell-cell boundaries. DNMBP is highly concentrated in these cell-cell boundaries, specifically on the apical surface, as observed in Caco-2, DLD-1 and MTD-1A epithelial cell lines. DNMBP co-localizes with the tight junction protein ZO-1, which interacts with the SH3-6 domain. Perinuclear localization of DNMBP also occurs. RNAi depletion of ZO-1 resultes in DNMBP not localizing to cell-cell boundaries. DNMBP mutants missing the RhoGEF domain were able to localize correctly. DNMBP mutants without the N-terminal half of the protein show mainly cytoplasmic localization with some membrane

localization. Localization to the membrane does not occur if DNMBP is truncated before SH3-5. Thus, most of the protein and interaction with ZO-1 is necessary for the correct localization of DNMBP to apical cell-cell boundaries in epithelial cells. Though it has not been confirmed, the BAR domain likely anchors DNMBP to the apical membranes.

1.12: Dnmbp as an NTD Susceptibility Gene

One of the major reasons why *Dnmbp* was chosen as the top candidate modifier gene of *Cecr2* over the other candidates was because of its functional association with NTDs. While there has been no direct link between *Dnmbp* and any type of NTD, some of the interacting partners of DNMBP have been associated with NTDs. A possible mechanism through which DNMBP may be linked to NTDs is actin irregularities. Normal actin dynamics is very important to the formation of the neural tube, as there are multiple examples of actin deficiencies leading to exencephaly. For instance, mutations to the following actin associated genes results in NTDs: *Mena* and *Pfn1* double mutant (*146*), *Mena* and *Vasp* double mutant (*147*), *Abl* and *Arg* double mutant (*148*), RhoGAP *Gr1f1* (*149*), *Marcks* (*150*), *Palld* (*151*), *Shroom3* (*152*), *Abi1* (*153*), *Cfl1* (*154*), RhoGAP *Dlc1* (*155*), and *Vcl* (*156*). Interestingly, the only NTD that occurs when the majority of these genes are mutated is exencephaly. Furthermore, inhibition of F-actin polymerization in rat and mouse embryos when exposed to cytochalasin D results in exencephaly and no other NTDs (*157*, *158*). Thus, closure of the cranial neural tube likely relies on an actin-based mechanism more so than spinal closure. Intriguingly, *Cecr2* mutations in mice only causes cranial neural tube defects, thus, it is possible that actin-related abnormalities are involved.

Similar to epithelial cells, neuroepithelial cells that make up the developing neural tube contain F-actin at apical cell-cell boundaries (*159*). Evidence suggests that the F-actin acts as

"purse strings," causing contraction of the apical ends of the cells with the help of muscle myosin II, which is thought to be necessary for bending of the neural plate and formation of hinge points (reviewed in *159*). From the list of genes mentioned above, *Marcks*, *Shroom3*, *Mena*, and *Vasp* all play a role in apical constriction. *Shroom3* has been extensively studied, and it had been revealed that deficiencies in Shroom protein prevent hinge point formation in frogs (*160*).

MENA and VASP are of interest as both interact with the SH3-6 domain of DNMBP and have been associated with neural tube defects. MENA and VASP bind F-actin and prevent capping, allowing for the growth of long unbranched actin filaments that affect cell motility and morphology (*114*, *147*, *161*). Mutation to either *Mena*, *Vasp* or *Pfn1* (Profilin I) alone in mice does not cause NTDs, possibly due to redundancy of the genes (*146*, *147*). However, double null mutations of both MENA and VASP or a null mutation to MENA and a heterozygous mutation to Profilin I leads to exencephaly (*146*, *147*). Profilin I binds directly to MENA and actin, and regulates actin polymerization. *Vasp* and *Pfn1* double mutants do not show signs of NTDs. Thus, if MENA and VASP are in fact redundant, disruptions in neurulation due to *Mena* and *Pfn1* double mutants must occur through a different pathway then from *Mena* and *Vasp*.

An alternative mechanism in which MENA and VASP may be involved in is the fusion of the neural tube once neural folds appose. MENA and VASP have been associated with the growth of filopodia – cellular protrusions comprised of F-actin (*161*, *162*). During fusion of the neural tube, evidence supports the formation of filopodia at the tips of neural folds (*19*). Another interactor of DNMBP, CDC42, also plays a role in filopodia formation. Once active, CDC42 binds directly to IRSp53 (insulin receptor tyrosine kinase substrate protein of 53kDa), which causes filopodia formation through the recruitment and binding of actin related proteins, such as

WAVE1, N-WASP, and MENA (all actin regulators and interactors of DNMBP), to the SH3 domain of IRSP53 (*163–166*). Interestingly dynamin, a major interactor of DNMBP, also binds to IRSp53 (*163*). Knockdown of dynamin or mutation to the region in dynamin that binds actin (K44A) in neuroblastomas prevents the formation of filopodia. Furthermore, it was shown that dynamin and MENA only localize to filopodia during initiation and growth. It has not been shown that the activation of CDC42 by DNMBP results in filopodia formation, thus, another GEF may be involved. Nevertheless, DNMBP does bind many of the proteins needed for the formation of filopodia, so it possible that DNMBP functions to bring these proteins together to IRSp53.

CDC42 has been a difficult protein to study during neurulation since homozygous null mutations of the gene in mice is embryonic lethal before neurulation begins (*167*). In embryonic stem cells, a CDC42 deficiency prevents actin polymerization. Conditional CDC42 knockouts in mice have allowed the investigation of CDC42 activity during different stages of development (reviewed in *168*). Unfortunately, conditional knockout of *Cdc42* in neuroepithelium during neurulation has yet to be extensively studied. Recently, *Cdc42* conditional knockout experiments were conducted to examine filopodia formation in mouse embryos. *Cdc42* was conditionally knocked-out in the surface ectoderm close to the neural fold tips (part of the surface ectoderm is involved in fusion of the neural tube) (*169*). The conditional knockout was lethal before the neural tube could be completely closed, but allowed for observation of initial closure. Unlike control embryos, filopodia do not form in *Cdc42* conditional knockout mice, suggesting CDC42 is necessary for early stages of neural tube fusion. Experiments conducted in mosaic *Xenopus laevis*, in which some cells expressed a dominant-negative CDC42, showed that neuroepithelial cells have abnormal mitotic spindle orientation during neural tube closure, while epidermal cells

are normal (*170*). This leads to abnormal orientation of dividing cells. It was not mentioned if

neural tubes completely closed, although if all cells contained the CDC42 mutation, closure

would likely not occur since orientation of cells is necessary for proper neural tube formation.

Aside from filopodia formation, effectors of CDC42 have been linked to other actin

regulatory processes (*171*, *172*). For example, active CDC42 binds to Par6, which then allows

for the activation of aPKC and stabilization of microtubules (*173*). Another effector of CDC42,

IQGAP1, is involved in reorientation of microtubules in the cell (*174*). As already mentioned,

binding of CDC42 to N-WASP alleviates the auto-inhibition state of N-WASP, which then

allows for the activation the Arp2/3 complex involved in actin nucleation (*96, 140*). DNMBP's

role in actin regulation through interaction with proteins, namely CDC42, and those that bind to

the SH3-6 domain have been described in the section 1.9. Additionally, it has been shown that

the SH3-6 domain is able to recruit F-actin (*96*). A fusion protein was produced with a

mitochondrial anchoring sequence attached to the SH3-6 domain of DNMBP. Expression of this

protein leads to increased F-actin localization to the mitochondria, showing DNMBP's ability to

increase F-actin recruitment to areas where it is localized. Thus, DNMBP's involvement in actin

regulation is highly evident, and may very well be the pathway it affects if it is modifying *Cecr2*.

The actin regulatory tight junction scaffold protein ZO-1, which also interacts with

DNMBP, has somewhat been associated with neural tube defects. In chick embryos, the

expression of ZO-1 increases in the neural plate as tube formation progresses (*175*). Knockout of

this gene in mice is embryonic lethal (*176*). While embryos do not survive to the stage of neural

tube closure, likely due to defects in other developmental processes, it was observed that there is

disorganization in the notochord and neural tube area due to abnormal apoptosis. It would be

interesting to conditionally knockout ZO-1 in the neural plate to determine if neurulation is

disrupted. ZO-2 is a related protein with a similar function to ZO-1 in tight junction formation, with some functional redundancy though not enough to completely compensate if one is deficient (*118*, *177*). ZO-2 deficiency in mice is lethal before the start of neurulation (*178*). However, an indirect link between ZO-2 and NTDs has been shown through the interacting partner SCRIB (*179*). *Scrib* mutant mice develop the most severe NTD, craniorachischisis (*180*). Along with *Dnmbp*, *Zo-2* is one of the three top candidate genes thought to be modifiers of *Cecr2*. Interestingly, *Zo-2* and *Dnmbp* are found in the different non-overlapping segments of chromosome 19 modifier regions. These two distinct regions lead to a decrease in exencephaly penetrance (see section 1.6). Since the modifier genes of *Cecr2* are expected not to have additive effects, modifiers are likely in the same or related pathways. Since there is evidence of an interaction between DNMBP and ZO-2, there is support for these proteins being in the same pathway, and thus, providing more support for *Dnmbp* being an NTD susceptibility gene and modifier of *Cecr2*.

1.13: Variants of DNMBP

The number and predicted deleteriousness of *Dnmbp* variants found by Dr. Renee Leduc was another major reason for choosing to focus on *Dnmbp* over other candidate modifier genes. Through sequencing of *Dnmbp* in the FVB/n and BALB/cCrl mouse strains and in 156 human cranial NTD probands, nine variants were identified that were predicted to be deleterious (*92*). Variants are listed in Table 1.13.1 in the order they appear in the protein sequence of DNMBP. One variant was found in the susceptible BALB/cCrl strain that differed from the resistant FVB/N strain (variant 4) and the other eight variants were sequenced in human NTD samples. The mouse variant was homozyogous in BALB/cCrl while the human variants were

heterozygous. Some human variants (1, 3, and 8) were found in multiple probands. As discussed in section 1.7, minor allele frequencies and GERP, SIFT and PP2 scores were utilized to predict deleteriousness of each variant. Of the human variants, two were novel and three had MAF scores below 0.03. Some variants were maintained on the list even with insignificant scores because they were thought to be involved with compound heterozygosity in some probands.

Figure 1.13.1 shows the location of each mutation in the gene. Variant 1 (N373K) is found in the N-terminal half of DNMBP. It is not within any major domains and MAF, SIFT, PP2 and GERP scores were not significant, making this variant the least likely to be deleterious on its own. Variants 2 to 4 (P820L, M831T, and P945L) are found within the RhoGEF domain with variant 4 being from the susceptible BALB/cCrl strain. The RhoGEF domain is specific for the activation of CDC42. The BAR domain contains the two novel variants (5 and 6). Variant 5 (R1024X) is predicted to be the most deleterious of all as it results in a truncation. The BAR domain allows for DNMBP to bind to membranes and the C-terminus is required for localization, thus, disruptions here could result in abnormally localized DNMBP. The last three variants (R1376H, C1413W, and T1462M) are found between the SH3-5 and SH3-6 domains. They do not fall within the very important SH-6 domain, though variant 7 and 9 are both rare and predicted to be damaging by PP2 and GERP analyses. Truncating the protein right before SH3-5 results in localization abnormalities (*117*). Thus, if the mutations in variants 7 to 9 are in regions of *Dnmbp* that play a role in localization, the protein may not localize correctly. Furthermore, the proband from which variant 9 was sequenced also contained a *Cecr2* variant predicted to be damaging (*92*). Variant 8 is another less likely variant to be deleterious out of all variants due to its position and scores. Based on the location of mutations alone, variants 2 to 6 are likely to be the most damaging since they are found within key domains. The most damaging is likely variant

5 as it contains an early stop codon, resulting in almost half of the protein being lost. Processes related to CDC42 activation and/or processes that need correct localization of DNMBP are most likely to be affected if variants are deleterious.

Table 1.13.1: DNMBP protein-coding sequence variants of interest. Variants identified by next generation sequencing in cranial NTD probands were Sanger sequence validated in probands. One mouse variant from BALB/cCrl is included. Variants are listed in order of amino acid residue. Table modified from Leduc (*92*).

Variant	Probands Containing Variant	AA Change	MAF	SIFT	PP2	GERP
1	8	N373K	0.04	0.89	0.049	-1.33
2	1	P820L	0.001744	0.62	0.015	**3.96**
3	2	M831T	0.03186	0.45	*0.673*	**5.99**
4	Mouse	P945L	-	**0.01**	**0.989**	-0.345
5	1	R1024X	-	-	-	**3.85**
6	1	R1154Q	-	0.59	**1**	**5.82**
7	1	R1376H	0.000116	0.07	**1**	**5.42**
8	8	C1413W	0.380349	0.19	*0.6*	0.604
9	1	T1462M	0.002209	0.09	**0.986**	**3.49**

Abbreviations – AA: amino acid, MAF: minor allele frequency, SIFT: sorting intolerant from tolerant, PP2: Polyphen2, GERP: genomic evolutionary rate profiling
MAF Score < 0.03 – rare variant, Novel variants are not in the database, thus a MAF score is not assigned
Underlined MAF Scores – Variant may contribute to compound heterozygosity in some probands
Bold numbers – Predicted to be damaging (SIFT <0.05, PP2 >0.956) or evolutionarily conserved (GERP>2)
Italicized numbers – Predicted to be possibly damaging (PP2 between 0.453 and 0.956)
Red text – nonsense mutation
Note: scores are from exome sequencing project (ESP6500), which may differ from other sources such as Ensembl.
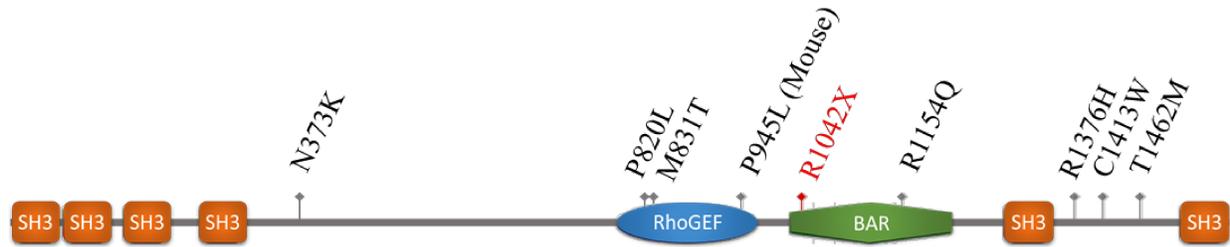
Figure 1.13.1: Location of variants on DNMBP protein. Eight human variants and one mouse variant (P945L) were predicted to be deleterious. One human variant was found that leads to an early stop codon (red text).

1.14: Aim, Objectives, and Significance

The overall aim of this research was to test the hypothesis that the candidate gene *Dnmbp* is an NTD susceptibility gene in humans and a modifier of *Cecr2* in mice. To test this hypothesis, variants of DNMBP predicted to be damaging that were found in human NTD probands were studied in cell lines to verify their functionality. One variant that was found in the susceptible BALB/cCrl mouse strain was also studied. Each variant was expressed in cultured cell lines. If the protein product of a particular variant was not functioning normally in an assay, it would provide support for the hypothesis that *Dnmbp* is necessary for neurulation. Three main objectives were utilized to test each variant, which evaluated protein-protein interaction, CDC42 activation and cellular localization:

Objective 1: Are *Dnmbp* variants able to interact with protein partners?

DNMBP has numerous protein partners, many of which are regulators of dynamin and actin. Since it was not feasible to examine the interaction of every protein partner, a select few proteins were examined. These were chosen based on their function related to neurulation. Two

important interactors are MENA and VASP, proteins that are involved in the regulation of actin polymerization (*96*). Using co-immunoprecipitation on cells expressing DNMBP variants, I attempted to show the interaction between DNMBP variants and MENA or VASP.

Objective 2: Are *Dnmbp* variants able to correctly activate CDC42?

Since there are six variants with mutations in either the RhoGEF or BAR domain, I attempted to determine if CDC42 activation was affected by variant DNMBP using an ELISA based CDC42 activation assay kit as well as analyzing dense-core vesicles in cultured neural cells. Additionally, to determine if decreased CDC42 activation is associated with exencephaly penetrance in *Cecr2* mutants, CDC42 heterozygous deletion mutants were produced on a dominant FVB/N resistant background with a homozygous *Cecr2* mutation. A deficiency in CDC42 would mimic a decrease in the active form of CDC42 similar to what would occur if DNMBP was not fully activating CDC42. Thus, if *Dnmbp* is modifying *Cecr2* through *Cdc42*, a deficiency in *Cdc42* should result in an increase in exencephaly penetrance on the resistant FVB/N background.

Objective 3: Are *Dnmbp* variant-containing proteins able to localize correctly in the cell?

If a protein does not localize to the correct cellular region, overall function could be affected. This is because the protein may not be close enough to other proteins it must interact with or targets it must act on. For this reason, I determined if DNMBP variants maintain normal localization in cultured epithelial cells. DNMBP is expected to localize at apical cell-cell junctions of epithelial cells in order to regulate cell junction configuration (*117*). Thus, variants

were expressed in the Caco-2 cell line and localization was assessed through immunofluorescence.

Significance: I tested *Dnmbp* as a putative modifier of *Cecr2*, a gene implicated in neural tube defects. Characterizing the function, localization, and interactions of *Dnmbp* variants could shed light on a possible NTD susceptibility gene. This study will also further elucidate the complexity of neural system development and possibly identify new components.

Chapter 2:

Materials and Methods

2.1: Bacterial Transformations

Bacterial transformations were conducted for the purposes of cloning and replicating plasmid vectors. 100ul of *E. coli* DH5α competent cells previously produced in our lab were thawed on ice. Up to 100ng of plasmid DNA (1-5ul) was added to the cells and the mixture was incubated on ice for 30 min. Cells were heat shocked in a 42°C waterbath for 1.5 min and then incubated on ice for 2 min. 250ul of pre-warmed Super Optimal Broth (SOB) media (2% bacto-tryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10mM $MgSO_4$, 10mM glucose) was added to the cells and incubated at 37°C for 1 hour. Cells were then plated on lysogeny broth (LB) agar plates (1% bacto-tryptone, 0.5% yeast extract, 1% NaCl, 1.5% agar, pH 7.5) containing either 50ug/ml of kanamycin or 100ug/ml of ampicillin for plasmid selection. Plates were incubated overnight at 37°C. Colonies were randomly chosen and used to inoculate 5ml of LB medium (1% tryptone, 0.5% yeast extract, and 1% NaCl, pH of 7.5) containing the respective selectable marker (50ug/ml of kanamycin or 100ug/ml of ampicillin). The cells were grown overnight at 37°C on a shaker (200RPM). Plasmid DNA was extracted from bacterial cells using the QIAquick® Spin Miniprep Kit (cat. no. 27104) following the manufacturer's protocol. For endotoxin-free plasmid DNA to be used for cell transfections, the QIAGEN Plasmid Maxi Kit (Cat. no. 12163) was utilized, following the manufacturer's protocol.

2.2: DNMBP Variant Construction

The Gateway® cloning system was utilized for the construction and expression of *Dnmbp*. The system is derived from the bacteriophage Lambda's site-specific recombination mechanism. During the lysogenic cycle, the bacteriophage will insert its genome into the genomes of *E. coli* (*181*). This occurs through an integrative recombination reaction between the

attP (attachment P) site in the bacteriophage genome and the attB site in the *E. coli* genome; this reaction requires integrase and integration host factor proteins (*182*). The recombination of these sites produces hybrid attR and attL recombination sites flanking the inserted DNA. During the lytic cycle, the bacteriophage genome can be excised through recombination of the attL and attR sites, which is mediated by integrase, excisionase, and integration host factor proteins.

This system has been engineered by Life Technologies to efficiently shuttle DNA sequences between different plasmids (*182*). Gateway plasmids contain open reading frames in which sequences can be inserted and flanked by att sites.

A human reference (wildtype) *Dnmbp* gene (Appendix A) was produced using synthetic oligonucleotides and inserted into a plasmid with flanking attB sites (attB1 and attB2) (produced by Life Technologies). These attB sites then underwent recombination with attP1 and attP2 sites, respectively, of a pDONR^{TM}221 Gateway® donor vector, creating a *Dnmbp* pENTR^{TM}221 Gateway® entry clone with flanking attL sites as well as a kanamycin selectable marker (prepared by Life Technologies). The entry vector allows for easy moving of *Dnmbp* into a variety of attR-containing Gateway® expression vectors through a simple recombination reaction. The *Dnmbp* gene in the entry vector was Sanger sequence verified using primers (Table 2.2.1) designed with the program Primer3 (http://bioinfo.ut.ee/primer3/) (*183*) and synthesized by Integrative DNA Technologies (IDT, Inc.).

Table 2.2.1: List of primers used for sequence verification of DNMBP

containing vectors.

Name	Sequence
DNMBP-pENTR^{TM}221-1-Forward	ACAACGTTCAAATCCGCTCC
DNMBP-pENTR^{TM}221-2-Forward	AAGAACTGCCGCTCTTTGTG
DNMBP-pENTR^{TM}221-2-Reverse	CTCATCCACAGTCCTCAGGG
DNMBP-pENTR^{TM}221-3-Forward	TCTGCTCAGCTGGATGAAGA
DNMBP-pENTR^{TM}221-3-Reverse	GCAGAGCCATGGTTTCCTC
DNMBP-pENTR^{TM}221-4-Forward	TTCTGGCGACCTTGGAAGAT
DNMBP-pENTR^{TM}221-4-Reverse	GCTCCCTCCCACTGTAGAAT
DNMBP-pENTR^{TM}221-5-Forward	GACTCTCCCACATCTGACCC
DNMBP-pENTR^{TM}221-5-Reverse	GTTGTGTCAGCTTCGAGTCC
DNMBP-pENTR^{TM}221-6-Forward	AGCAGCAACACATTCACAGG
DNMBP-pENTR^{TM}221-6-Reverse	TCCCGCTCCATTTCCTCAAT
DNMBP-pENTR^{TM}221-7-Forward	ATTGAGGAAATGGAGCGGGA
DNMBP-pENTR^{TM}221-7-Reverse	CGGTGACCAAGAAACACAGG
DNMBP-pENTR^{TM}221-8-Forward	CTGTGTTTCTTGGTCACCGG
DNMBP-pENTR^{TM}221-8-Reverse	AGGAGCAAAGCCAGTGAGAT
DNMBP-pENTR^{TM}221-9-Forward	GCGGTCAAGGAAATCAACGT
DNMBP-pENTR^{TM}221-9-Reverse	GGCCCTGTAAACATGCTCAG
DNMBP-pENTR^{TM}221-10-Forward	CAGTTTGAGAGGGTGCATCG
DNMBP-pENTR^{TM}221-10-Reverse	ACTGGCGGTCAATGGTTTTC
DNMBP-pENTR^{TM}221-11-Forward	CGAAGAGCACAGCAGAGTTC
DNMBP-pENTR^{TM}221-11-Reverse	CATGCTGCTGGGGTTGAAG
DNMBP-pENTR^{TM}221-11.5-Forward	AGGCTTCGTGTACAGCTCTT
DNMBP-pENTR^{TM}221-11.5-Reverse	TTGGGTTTCGTGCCTTGAAG
DNMBP-pENTR^{TM}221-12-Forward	GTGCAAGAACAGCCCAGG
DNMBP-pENTR^{TM}221-12-Reverse	CTTGTGCAATGTAACATCAGAGA
DNMBP-pENTR^{TM}221-13-Forward	TTGTTGCAACGAACAGGTCA
DNMBP-pENTR^{TM}221-13-Reverse	GGCATAAATTCCGTCAGCCA

2.3: Site-Directed Mutagenesis

After sequence validation of the entry vector containing wildtype *Dnmbp*, the nine

variants were produced through site-directed mutagenesis. This method allowed for the precise

mutation of single base pairs. The protocol, adapted from Niederriter et al. (*184*) and Zheng et al.

(*185*), involved the designing of primers using the web-based software PrimerX

(http://www.bioinformatics.org/primerx). A pair of complementary primers were produced for

each variant with the desired mutation in the center of the primer (Table 2.3.1). To optimize mutagenesis, six different PCR reactions were carried out with varying concentrations of primers and DMSO (Table 2.3.2). Mutagenesis reactions were carried out on Peltier Thermal Cycler-200 (MJ Research) with the following program: 1) 95°C for 5 min, 2) 95°C for 50 seconds, 3) 55°C for 1 min, 4) 68°C for 9 min, 5) repeat steps 2-4 18 times, 6) 68°C for 7 min, 7) 4°C hold. The primers anneal to the template, creating a replicate of the vector that incorporates the mutation. After the amplification process was complete, 4μl of the PCR products of the 12.5μl reaction were run on a 0.8% agarose gel to determine which of the six mutagenesis reactions resulted in the best amplification. Once determined, the remaining amount of the best reaction (based on band intensity) was treated with 0.5μl of *DpnI* for two hours at 37°C, which cleaves at DAM methylated sites. Since only the template plasmids were methylated due to their production in bacterial cells, the templates were destroyed and the mutant plasmids remained. Sequences of the Dnmbp-pENTR^{TM}221 constructs can be found in Appendix A.

The mutagenesis reaction was then transformed into competent *E. coli* cells. Cells were grown at 37°C overnight on agar plates containing kanamycin. Four colonies were then picked at random for each variant and grown overnight at 37°C on a shaker at 200 RPM in liquid LB media containing kanamycin. Plasmid DNA was extracted using a QIAquick® Spin Miniprep Kit (cat. no. 27104), which was then Sanger sequence verified for the presence of the mutation and integrity of the rest of the gene and attL sites.

Table 2.3.1: List of primers used for site-directed mutagenesis.

Site of mutation indicated in bold.

Name	Sequence
DNMBP - N373K Forward	GTATGACACAGACAGAAA**G**TCTTATCAGGACGAGGAC
DNMBP - N373K Reverse	GTCCTCGTCCTGATAAGA**C**TTTCTGTCTGTGTCATAC
DNMBP - P820L Forward	GCAGCAGGCACAGGTAC**T**AAACATTGATTTTGAGG
DNMBP - P820L Reverse	CCTCAAAATCAATGTTT**A**GTACCTGTGCCTGCTGC
DNMBP - M831T Forward	GAGGGACTTTTTGGAAATA**C**GCAGATGGTGATTAAGGTC
DNMBP - M831T Reverse	GACCTTAATCACCATCTGCG**T**ATTTCCAAAAAGTCCCTC
DNMBP - P945L Forward	CCCCAGAATCCCACCC**T**GATAAAGTGCCTTTAAC
DNMBP - P945L Reverse	GTTAAAGGCACTTTA**C**CGGGGTGGGATTCTGGGG
DNMBP - R1024X Forward	GATTGATTAAGTCTTTTATC**T**GAGACCTGTCTCTACCTCC
DNMBP - R1024X Reverse	GGAGGTAGAGAGACAGGTCTC**A**GATAAAAGACTTAATCAATC
DNMBP - R1154Q Forward	GCTGCAGTCGGCCC**A**GAACAACTATGAGG
DNMBP - R1154Q Reverse	CCTCATAGTTGTTC**T**GGGCCGACTGCAGC
DNMBP - R1376H Forward	CCAGGTTCCCACAC**A**CCAGAACAGCGGC
DNMBP - R1376H Reverse	GCCGCTGTTCTGGT**G**TGTGGGAACCTGG
DNMBP - C1413W Forward	CTCCGCCAAAAGAATG**G**GACCAAGGAACTCTC
DNMBP - C1413W Reverse	GAGAGTTCCTTGGTC**C**CATTCTTTTGGCGGAG
DNMBP - T1462M Forward	CAACCCACTGCCA**T**GCCGAGGAGCTAC
DNMBP - T1462M Reverse	GTAGCTCCTCGGC**A**TGGCAGTGGGTTG

Table 2.3.2: Site-directed mutagenesis conditions used for optimization.

	Reaction 1	Reaction 2	Reaction 3	Reaction 4	Reaction 5	Reaction 6
Forward Primer (2.5uM)	1.5μl	1.5μl	1μl	1μl	0.5μl	0.5μl
Reverse Primer (2.5uM)	1.5μl	1.5μl	1μl	1μl	0.5μl	0.5μl
dNTPs (40uM)	0.5μl	0.5μl	0.5μl	0.5μl	0.5μl	0.5μl
Phusion Buffer (5X High Fidelity)*	2.5μl	2.5μl	2.5μl	2.5μl	2.5μl	2.5μl
***DNMBP*-pDONR221 DNA template (2ng/uL)**	1μl	1μl	1μl	1μl	1μl	1μl
milliQ H$_2$0	4.8μl	4.2μl	5.8μl	5.2μl	6.8μl	6.2μl
Phusion High Fidelity DNA Polymerase*	0.2μl	0.2μl	0.2μl	0.2μl	0.2μl	0.2μl
Dimethyl Sulfoxide (DMSO)*	-	0.6μl	-	0.6μl	-	0.6μl
Total	12.5μl	12.5μl	12.5μl	12.5μl	12.5μl	12.5μl

*Reagents from New England Biolabs

2.4: Subcloning of DNMBP Variants into HA Expression Vectors

Once the entry clones containing *Dnmbp* variants flanked by attL sites were sequence verified, wildtype *Dnmbp* and variants were subcloned into a pCIHA vector supplied by Dr. Fred Berry, which is a modified pCI vector (Promega Corp.). It had been modified to contain an inframe HA-tag (between the Nhel and EcoRI sites within the multiple cloning site) and an inframe gateway cassette containing flanking attR sites. The pCIHA vector contains an ampicillin selectable marker and an N-terminal HA tag in the open reading frame, resulting in *DNMBP* carrying an HA tag when expressed from this vector. The open reading frame contains attR sites that recombine with attL sites, leading to a transfer of *Dnmbp* into the expression vector. In the subcloning reaction, 75ng of the entry clone (100ng/μl stock) and 75ng of the pCIHA vector (150ng/μl stock) were combined with 2.5μl of Tris-EDTA buffer (pH 8,

Integrated DNA Technologies) and 1μl of Gateway® LR Clonase II Enzyme Mix (Life

Technologies), which is a blend of integrase, excisionase, and integration host factor proteins,

resulting in the recombination of the sites. After a one hour incubation at 25°C, Proteinase K was

added to the reactions and incubated at 37°C to degrade the LR Clonase II Enzymes. 1μl of the

reaction was transformed into DH5α *E. coli* cells and grown overnight at 37°C on agar plates

containing ampicillin. Two colonies were picked at random for each variant and were grown in

LB medium containing ampicillin overnight on a shaker at 37°C The plasmid DNA was Sanger

sequenced to confirm the presence of the HA tag and also to ensure that that start codon of the

HA tag was in-frame with the start codon of *Dnmbp*. Once confirmed, the QIAGEN Plasmid

Maxi Kit (cat. no. 12163) was used to extract the plasmid DNA for transfection. The

concentration and quality of the DNA was measured using Nanodrop 1000 (Thermo

Scientific™) and brought to 1000ng/μl.

2.5: Cell Culture Conditions

HEK293 cells, originally supplied by Dr. Rosealine Godbout, U-2 OS cells supplied by

Dr. Fred Berry, SK-N-BE(2) supplied by Dr. David Eisenstat, and Caco-2 were utilized for

variant DNMBP protein analyses. Depending on the experiment, cells were grown in Corning®

T25 Vented Flasks, Sarstedt® TC 60mm Dishes, or Themo Scientific™ BioLite 12-well and 24-

well plates. Cells were incubated in a 37°C incubator (Sanyo Electric Co.,) with 5% CO_2. Cell

culture work was conducted in a sterile biological safety cabinet. Materials being used were

sprayed with 70% ethanol to prevent contaminations.

2.6: Cell culture Media and other Culture Reagents

For HEK293, Caco-2, and U-2 OS cell lines, complete growth media consisted of Dulbecco's Modified Eagles Medium (with 4500 mg/L glucose, L-glutamine, sodium pyruvate, and sodium bicarbonate, Sigma-Aldrich®), 10% Heat Inactivated Fetal Bovine Serum (Gibco), 1% L-Glutamine, and 1% Penicillin/Streptomycin (Gibco). For SK-N-BE(2) cells, complete growth media consisted of 1:1 mixture of Minimum Essential Medium Eagle (with L-glutamine and sodium bicarbonate, Sigma-Aldrich®) and Nutrient Mixture F-12 Ham (with L-glutamine and sodium bicarbonate, Sigma-Aldrich®), 10% Heat Inactivated Fetal Bovine Serum (Gibco), 1% L-Glutamine, and 1% Penicillin/Streptomycin (Gibco). The media was mixed together and filtered through a Nalgene™ Rapid-Flow Vacuum Sterile Bottle Top Filter with a 0.20μm pore size (Thermo Scientific™). Media was always pre-warmed in a 37°C water bath prior to use with cells. For cell dissociation during passages or harvesting, 0.25% 1X Tryspin-EDTA (Gibco) was used.

2.7: Thawing Cells

To revive cryopreserved cells, frozen vials containing cells were quickly transferred from liquid nitrogen storage to a 37°C waterbath for about one min. The contents of the vial were then slowly transferred to a 15ml falcon tube containing of 9ml of warm media. Cells were pelleted using a centrifuge at 75xg for 5 min. Media above the pellet was aspirated off, and the cells were resuspended in 4ml of fresh media and plated.

2.8: Passaging, Harvesting, and Cryopreserving Cells

Cells were passaged or harvested when approximately 70-90% confluent. Media was aspirated off dishes and 1X trypsin-EDTA was added. Cells were incubated for 5 min at 37°C or until cells were completely dissociated. The cells were then transferred to a 15ml falcon tube containing three times the amount of medium to trypsin, which deactivates the trypsin-EDTA. Cells were pelleted using a centrifuge at 75xg for 5 min followed by aspiration of the supernatant.

If cells were being passaged, they were resuspended in warm media and counted using a hemocytometer. After calculating the density, cells were diluted and plated at the desired concentration.

If cells were being harvested, the pellet of cells was washed with ice-cold 1X PBS and transferred to a pre-chilled microcentrifuge on ice. Cells were again pelleted in a microcentrifuge at 2000xg for 4 min. PBS was aspirated off and cells were then ready to be used for experiments.

If cells were being cryopreserved, the pellet was resuspended in 1ml of media containing 10% DMSO and transferred to a 2ml cryogenic vial. The vial was then placed in an insulated container, which was stored at -80°C overnight for a slow-freeze. The vial was then transferred to liquid nitrogen.

2.9 Transient Transfection of Expression Plasmids into Cells

Cells were transiently transfected with vectors using TranIT-LT1 Transfection Reagent (Mirus Bio LLC) or Lipofectamine 2000 (Thermo Fisher Scientific). Following the manufacturer's protocol, cells were plated at a density that would result in ~80% confluence the next day (Table 2.9.1). Room temperature TransIT-LT1 reagent or Lipofectamine, Opti-MEM I

Reduced-Serum Medium (Gibco), and plasmid DNA (1μg/μl stock) were mixed in a microcentrifuge tube and incubated at room temperature (Table 2.9.2). The mixture was then applied dropwise to cells and mixed into the media by rocking the dish. After 48 hours, cells were harvested as described above for experiments. For non-transfected control cells, plasmid was omitted.

Transfection efficiency was measured for wildtype transfected HEK293, U-2 OS, and SK-N-BE(2) cells and for each vector transfected into Caco-2 cells by conducting immunofluorescence assays based on expression of the desired protein (Section 2.17) 48 hours after transfection. At least six different areas of cells were observed at random. The percent of cells transfected was calculated for each area and averaged to arrive at the transfection efficiency.

Table 2.9.1: Cell culture volumes and cell densities.

	Growth Media (ml)	Trypsin-EDTA Applied for Dissociation (ml)	Density of Cells Plated (cells/ml)		
			HEK293	Caco-2	U-2 OS
T25	4.0	2	$3.0x10^5$	-	-
60mm	3.0	1.5	$3.0x10^5$	$3.0x10^5$	$2.0x10^5$
12-well	1.0	0.5	-	$0.5x10^5$	$1.0x10^5$
24-well	0.5	-	$1.0x10^5$	$1.0x10^5$	$0.5x10^5$

Table 2.9.2: Volumes of media and reagents used for transfections.

	Reagent	Growth Media (ml)	Reduced-Serum Medium (µl)	DNA to Reagent Ratio (µg:µl)		
				HEK293	Caco-2	U-2 OS
60mm	TransIT®-LT1	4.0	400	4.0:12	-	4.0:12
60mm	Lipofectamine 2000	4.0	400	0.8:20	0.8:20	0.8:20
12-well	TransIT®-LT1	1.0	100	1.0:3.0	1.0:2.0	1.0:2.0
12-well	Lipofectamine 2000	1.0	100	1.6:4.0	0.8:20	0.8:20
24-well	TransIT®-LT1	0.5	50	0.5:1.5	0.5:1.0	0.5:1.0
24-well	Lipofectamine 2000	0.5	50	0.8:2.0	0.8:20	0.8:20

2.10: Protein Lysate Extraction

For protein extractions, non-denaturing lysis buffer (20mM Tris-HCl pH 8.0, 420mM NaCl, 10% glycerol, 1% IGEPAL CA-630, 2mM EDTA, 1% 100x proteinase inhibitor cocktail) was added to the pellets of cells obtained after harvesting. Cells were rocked for 30 min at 4°C. Protein was clarified by centrifuging for 20 min at 9500xg in a 4°C cold room. The supernatant was transferred to a pre-chilled microcentrifuge tube and the concentration was quantified using the DC™ Protein Assay (Bio-Rad). Protein was diluted 1:10 with lysis buffer and reagents were added following the manufacturer's protocol. After 15 min, absorbance was measured using a spectrophotometer set at 750nm and concentration was calculated using a pre-determined standard curve. If necessary, concentrations were equalized by diluting with lysis buffer. Protein was either used right away or was flash-frozen in liquid nitrogen and stored at -80°C.

2.11: Western Blotting

      To separate and detect for proteins in a sample, Western blots were conducted. Protein samples were combined with 1X NuPAGE™ loading buffer (Invitrogen) and 100mM of 2-mercapoethanol and denatured by incubating for 5 min in boiling water. Up to 20μg of protein sample was loaded onto a SDS-PAGE gel with a molecular weight marker (PageRuler Prestained Protein Ladder, Thermo Scientific™). The gel consisted of a 5% acrylamide stacking gel (12.5% of 40% 37.5:1 acrylamide/bisacrylamide stock, 50mM Tris-HCl ph 6.8, 0.1% sodium dodecyl sulfate, 0.1% ammonium persulfate, 0.1% tetramethylethylenediamine) and a 7.5% resolving gel (18.75% of 40% 37.5:1 acrylamide/bisacrylamide stock, 50mM Tris-HCl pH 8.8, 0.1% sodium dodecyl sulfate, 0.1% ammonium persulfate, 0.1% tetramethylethylenediamine). The Mini-PROTEAN® Electrophoresis System (Bio-Rad) was utilized to run the samples in the gel. After the addition of SDS-PAGE running buffer (25mM Tris, 0.19M glycine, 1% SDS), samples were loaded and run for approximately 10 min at 120V. Once the samples were in the resolving gel, the voltage was increased to 175V. The voltage was stopped once the dye front reached the bottom of the gel. The separated protein samples in the gel were then transferred to a 0.45μm polyvinylidene fluoride membrane (Immobilon-P Membrane, Millipore Sigma) using the Mini-PROTEAN (Bio-Rad) transfer apparatus. Methanol was applied to the membrane to activate it, followed by transfer buffer (25mM Tris, 190mM glycine, 10% methanol) for 5 min. The membrane and gel were set into the apparatus, which was set at 350mA for 30 min. The membrane, which would now have the molecular weight ladder and separated proteins on it, was washed in Tris-buffered saline with 0.05% Tween-20 (TBST) for 5 min followed by blocking in 5% skim milk in TBST for 45 min on a shaker. The blocking solution was removed, and primary antibodies diluted in 5% skim milk were applied and incubated overnight at 4°C. The next day,

the primary antibody was removed and the membrane was washed three times with TBST for 5 min on a shaker. A horseradish peroxidase (HRP) conjugated secondary antibody diluted in skim milk was then applied for 1 hour, incubating at room temperature on a shaker. A list of antibodies used can be found in Table 2.11.1. The membrane was then washed in TBST 3 times for 10 min. The HRP was detected by using the SuperSignal® West Pico Stable Peroxide Solution and Luminol Enhancer Solution (Thermo Scientific™) which was mixed in a 1:1 ratio, then applied to the membrane for 5 min. Lastly, the membrane was exposed to X-ray film (FUJIFILM) for 20 seconds, 1 min, 5 min, and 2 hours, and developed using a Kodak X-OMAT 2000 automated developer.

If a membrane needed to be re-probed for a different set of proteins, the antibodies on the membrane were stripped. This was done by incubating the membrane in stripping buffer (2% SDS, 12.5% 0.5M Tris pH 6.8, 0.8% 2-mercapoethanol) at 50°C for 40 min. The membrane was then rinsed under running water for one hour followed by a 5 min wash in TBST. After this wash, the protocol described above beginning at the blocking stage was followed.

**Table 2.11.1: List of antibodies used for Western blots (WB) and
immunofluorescence (IF) and their optimal concentration**

	WB Dilution	IF Dilution
Primary Antibodies		
anti-HA (Rabbit) (Proteintech Group, 51064-2-AP)	1:5000	1:400
anti-HA (Mouse) (Proteintech Group, 66006-1-Ig)	-	1:400
anti-DNMBP (Rabbit) (Proteintech Group, 17191-1-AP)	1:5000	1:400
anti-VASP (Rabbit) (Proteintech Group, 13472-1-AP	1:10000	-
anti-MENA (Rabbit) (Abcam, ab176820)	1:2000	-
anti-ZO-1 (Mouse) (Thermo Fisher Scientific, 33-9100)	-	1:300
anti-Tubulin (Mouse) (Sigma-Aldrich, T6199)	1:5000	-
anti-CgA (Rabbit) (ThermoFisher Scientific, SP12)	-	1:100
Secondary Antibodies		
anti-Mouse IgG-Peroxidase (Sigma-Aldrich, A0168)	1:10000	-
Goat Anti-Rabbit-HRP (Bio-Rad, 1705046)	1:5000	-
Goat anti-Rabbit IgG Alexa Fluor® 488 (Thermo Fisher Scientific, A11008)	-	1:2000
Cy3 anti-Mouse (Jackson ImmunoResearch Laboratories, Inc., 115-166-006)	-	1:2000

2.12: Co-immunoprecipitations

Co-immunoprecipitations were conducted to observe DNMBP protein-protein

interactions in HEK293 and U-2 OS cells. Protein was extracted from cells as described in

section 2.10. The general co-immunoprecipitation protocol involved using magnetic

Dynabeads® (Novex), which were added to a microcentrifuge tube and placed on a magnetic

rack. The beads were pulled to the wall of the tube, allowing for the removal of the supernatant.

Beads were resuspended in 200μl of 1X PBS-T (0.02% Tween-20). 1μg of an antibody specific

for the protein being pulled-down was added to the beads and rocked for 10 min at room

temperature. The tube was then placed on the magnetic rack and the solution was removed.

Beads were washed with PBS-T and then removed. Protein lysate (up to 500µg) was added to the beads along with other reaction buffers (Appendix B). The mixture of beads and protein was rocked for at least 2 hours at 4°C. The tube was then placed on the magnetic rack to remove the solution. Beads were washed at 4°C with 500µl of buffer. After the last wash, the beads were transferred to a new tube and all of the supernatant was completely removed. To elute complexes attached to the beads, 20µl of DTT and 20µl of 4X NuPAGE™ loading buffer (Invitrogen) were added to the beads and boiled in water for 5 min. The supernatant that contained the immunoprecipitated protein and any proteins that were bound to it was transferred to a new tube, allowing for analysis through Western blotting. Normal Rabbit IgG was used as a non-specific control. This general protocol led to inconsistent results, thus, attempts at optimizing the reactions were made. Optimization attempts included protein cross-linking using formaldehyde (24), transfection of DNMBP and protein partners to increase the frequency of interaction in cells, varying the concentrations of certain components of our lysis and reaction buffers, such as the salt, pH, and detergent, as well as incubation times for reactions and washes. A detailed summary of co-immunoprecipitation attempts can be found in Appendix B.

2.13: CDC42 Activation Assay

Activation of Cdc42 was assessed using the CDC42 G-LISA Activation Assay (Cytoskeleton, Inc.). The assay allows for the pull-down of GTP-bound CDC42 from cell protein lysate. HEK293 or U-2 OS cells were plated on Sarstedt® TC 60mm Dishes (Table 2.9.1). Cells were either serum starved for 24 or 48 hours, transfected with DNMBP (wildtype or variants), or grown under normal conditions. After 48 hours from transfections, cells were harvested and assayed following the manufacturer's protocol. The supplied lysis buffer in the kit was very

harsh leading to lysis of the nucleus and the addition of DNA to the protein lysate. To fix this issue, it was suggested by the manufacturer to dilute the lysis buffer with water in a 1:1 ratio. Briefly, protein lysate was incubated with an effector protein that binds specifically to the active form of CDC42. Inactive CDC42 and other proteins where washed away. A primary CDC42 antibody is applied followed by a horseradish peroxidase conjugate (HRP) secondary antibody. HRP substrate is added and a microplate spectrophotometer was utilized to measure absorbance, which corresponds to the amount of active CDC42.

2.14: Dense-core Vesicle Analysis

To determine if DNMBP variants were able to increase the amount of dense-core vesicles in neural cells similar to wildtype DNMBP as shown by Sato et al. (*143*), variants were transfected into SK-N-BE(2) cells. Cells were plated at a density of 5×10^4 cells/ml in 0.5ml in a 24-well plate with uncoated coverslips. Cells were transfected with Mirus Trans TranIT-LT1. Two days after transfection, immunofluorescence was carried out. Medium was aspirated from wells and cells were fixed by incubating in 4% formaldehyde for 20 min. This was followed by three gentle washes in 1X PBS for 5 min each. Cells were permeabilized by applying 0.1% Triton X-100 (Sigma) in 1X PBS (PBS-TX) for 5 min at room temperature. After removing the PBS-TX, cells were blocked in 3% bovine serum albumin (BSA) in 1X PBS with 0.1% Triton X-100 for 45 min. The blocking solution was removed and primary antibodies (anti-HA Mouse, anti-CgA Rabbit) diluted in fresh blocking solution were added to wells for one hour at room temperature. Following three washes in 1X PBS, secondary fluorescent antibodies were applied for 30 min. Antibody concentrations are listed in Table 2.11.1. Cells were washed three times with 1X PBS, and then incubated in DAPI for 5 min, diluted to 1:5000 with 1X PBS. After 3

washes for 5 min with 1X PBS, coverslips were carefully removed and mounted onto glass slides using Fluoromount G (SouthernBiotech). Coverslips were sealed with nail polish. Cells were viewed under oil immersion at 60X magnification on a Nikon Eclipse 80i Confocal Microscope. Controls included the absence of primary antibody to test for nonspecific binding of the secondary antibody and the absence of secondary antibody to determine if autofluorescence was occurring. At least 25 cells were imaged for each analyzed variants and non-transfected cells. The experiment was repeated three times. The program CellProfiler (*186*) was utilized to quantify various phenotypes of each cell including size and fluorescence intensity of CgA and HA-DNMBP staining.

2.15: Deletion of Cdc42 in Cecr2 Homozygous Mutants

Experiments involving mice were approved by the Animal Care and Use Committee of the University of Alberta (University of Alberta AUP 00000094). Mice were housed in a facility with a 14 hour light/10 hour dark cycle at room temperature (22 ± 2°C). Mice were fed either LabDiet Laboratory Rodent Diet 5001 (non-breeding) or LabDiet Mouse Diet 9F 5020 (breeding). After weaning (~2 weeks after birth), ear notches were used for the identification and genotyping of mice.

Cdc42$^{loxP/loxP}$ mice were purchased from Jackson Laboratory (Stock No: 027576) that contain floxed sites around the second exon of *Cdc42* on a predominately C57BL/6J background. The introduction of Cre leads to a deletion of exon 2 and expression of CDC42 is not detected (*125*). Figure 2.15.1 summarizes crosses utilized for the production of *Cd42* heterozygous-*Cecr2* homozygous mutants. To produce heterozygous *Cdc42$^{+/Del}$* mice, Cdc42$^{loxP/loxP}$ mice were crossed to EIIa-cre FVB/N-C57 mice (The Jackson Laboratory, Stock No: 003314). The Cre

transgene is downstream of the adenovirus EIIa promoter, which expresses the Cre recombinase throughout the mouse embryo including the germ line (*187*). This produced mice that were heterozygous for the *Cdc42* deletion on the mixed FVB/N-C57BL/6J background. Mice were genotyped to confirm deletion (described in section 2.16 below). The *Cecr2*^{GTbic45} mutation or the more severe *Cecr2*$^{tm1.1Hemc}$ mutation were then introduced into *Cdc42*$^{Del/+}$ mice by crossing them with *Cecr2*$^{GTbic45/GTbic45}$ or *Cecr2*$^{tm1.1Hemc/+}$ mice on a congenic FVB/N background. Resulting heterozygous mice for both mutations (*Cdc42*$^{Del/+}$; *Cecr2*$^{GT/+}$ or *Cdc42*$^{Del/+}$; *Cecr2* $^{tm1.1Hemc/+}$) were on a mixed FVB/N-C57BL/6J background. Since the modifying loci of FVB/N are dominant, the mixed background should have the equivalent to pure FVB resistance to exencephaly due to the *Cecr2* mutations. Thus, it was not necessary to produce a pure FVB/N line with the compound mutations. To test if a partial deficiency in CDC42 in combination with the absence of CECR2 resulted in exencephaly in resistant FVB/N mice, *Cdc42*$^{Del/+}$; *Cecr2*$^{GT/+}$ or *Cdc42*$^{Del/+}$; *Cecr2* $^{tm1.1Hemc/+}$ mutants were crossed to *Cdc42*$^{+/+}$; *Cecr2*$^{GT/GT}$ mice on a congenic FVB/N background.

The resulting embryos (expected genotypes listed in Figure 2.15.1) were dissected out of euthanized pregnant dams at 14.5-18.5 days post fertilization. Dissections were conducted in 1X PBS and the presence of any abnormalities including exencephaly or spina bifida was determined. Tail samples were taken from embryos for genotyping.
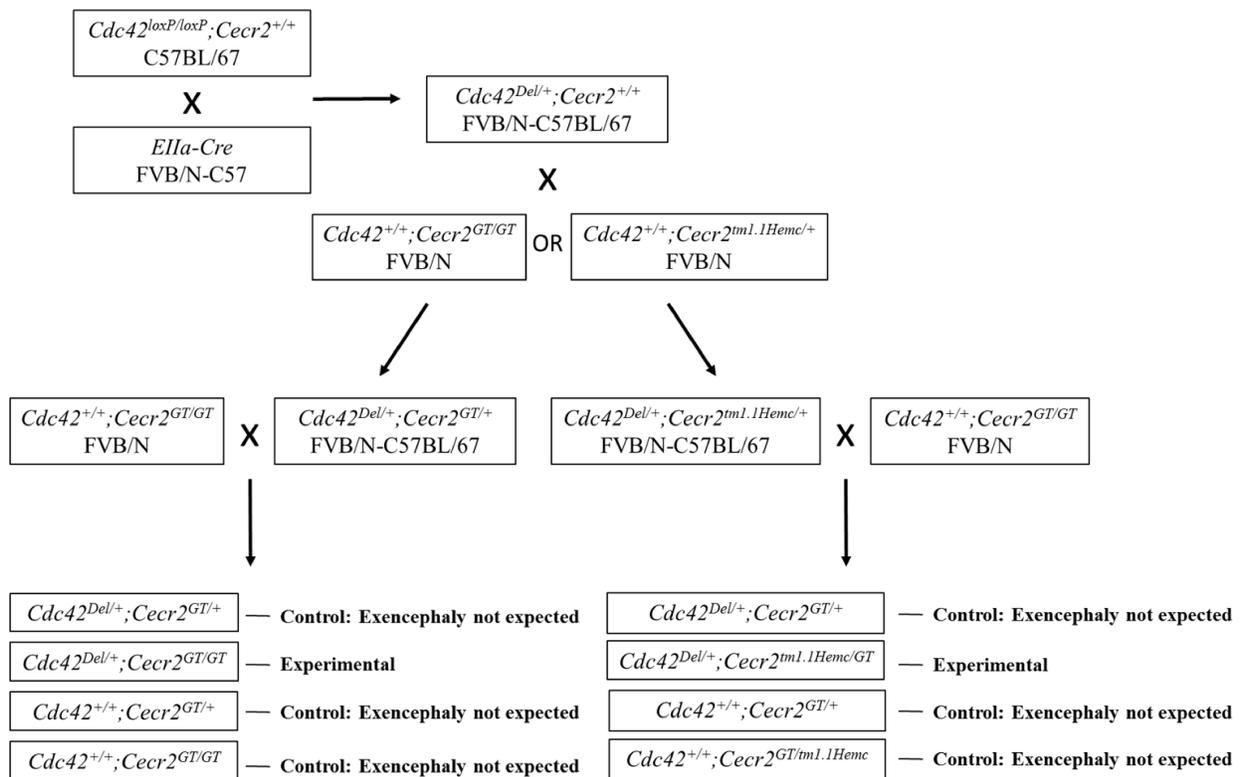
$Cdc42^{loxP/loxP};Cecr2^{+/+}$
C57BL/67

X

$Cdc42^{Del/+};Cecr2^{+/+}$
FVB/N-C57BL/67

EIIa-Cre
FVB/N-C57

X

$Cdc42^{+/+};Cecr2^{GT/GT}$
FVB/N   OR $Cdc42^{+/+};Cecr2^{tm1.1Hemc/+}$
FVB/N

$Cdc42^{+/+};Cecr2^{GT/GT}$
FVB/N   **X** $Cdc42^{Del/+};Cecr2^{GT/+}$
FVB/N-C57BL/67

$Cdc42^{Del/+};Cecr2^{tm1.1Hemc/+}$
FVB/N-C57BL/67   **X** $Cdc42^{+/+};Cecr2^{GT/GT}$
FVB/N

$Cdc42^{Del/+};Cecr2^{GT/+}$ — **Control: Exencephaly not expected**

$Cdc42^{Del/+};Cecr2^{GT/GT}$ — **Experimental**

$Cdc42^{+/+};Cecr2^{GT/+}$ — **Control: Exencephaly not expected**

$Cdc42^{+/+};Cecr2^{GT/GT}$ — **Control: Exencephaly not expected**

$Cdc42^{Del/+};Cecr2^{GT/+}$ — **Control: Exencephaly not expected**

$Cdc42^{Del/+};Cecr2^{tm1.1Hemc/GT}$ — **Experimental**

$Cdc42^{+/+};Cecr2^{GT/+}$ — **Control: Exencephaly not expected**

$Cdc42^{+/+};Cecr2^{GT/tm1.1Hemc}$ — **Control: Exencephaly not expected**

Figure 2.15.2: Mouse cross for the production *Cdc42* heterozygous and *Cecr2* homozygous compound deletion mutants. Relevant genotypes and progeny are shown after each cross. Final crosses show expected genotypes of each embryo and the purpose of each embryo in the experiment.

2.16: Mouse DNA Extraction and Genotyping

Extraction of genomic DNA from ear clippings of mice or tail clips of embryos began with an incubation at 95°C in 75µl of 50mM NaOH for one hour followed by vortexing. The reaction could then be used for amplifying DNA through PCR.

For genotyping *Cdc42* deletion mutants, a PCR reaction was set up containing 1X DreamTaq PCR buffer (ThermoFisher Scientific), 2.5μM MgCl$_2$, 0.2μM dNTPs, 0.6μM of each primer (Table 2.16.1), 1U of DreamTaq DNA polymerase (ThermoFisher Scientific), and 1μl of DNA. PCR reactions were run on a T100 Thermocycler (Bio-Rad) with the following program: 1) 94°C for 10 min, 2) 94°C for 1 min, 3) 55°C for 1 min, 4) 72°C for 1 min, 5) repeat steps 2-4 35 times, 6) 72°C for 10 min, 7) 4°C hold. Once complete, 15μl of PCR products were mixed with 3μl of Orange G loading buffer and separated on a 1.5% agarose gel at 100V with a no-template control. A 600bp band indicated wildtype *Cdc42*, a 700bp band indicated floxed *Cdc42*, and a 200bp band indicated *Cdc42* deletion.

For genotyping *Cecr2*^{GTbic45} mutants, a PCR reaction was set up containing 1X DreamTaq PCR buffer (ThermoFisher Scientific), 0.25μM dNTPs, 1μM of each primer (Table 2.16.1), 1U of DreamTaq DNA polymerase (ThermoFisher Scientific), and 1μl of DNA. PCR reactions were run on a T100 Thermocycler (Bio-Rad) with the following program: 1) 94°C for 1.5 min, 2) 94°C for 15 seconds, 3) 60°C for 20 seconds, 4) 68°C for 40 seconds, 5) repeat steps 2-4 36 times, 6) 68°C for 5 min, 7) 4°C hold. Once complete, PCR products were mixed with 3μl of Orange G loading buffer and separated on a 2.0% agarose gel at 130V with a no-template control. A 376bp band indicated wildtype *Cecr2*, while a 537bp band represented *Cecr2*^{GTbic45}.

For genotyping *Cecr2*$^{tm1.1Hemc}$ mutants, a PCR reaction was set up containing 1X DreamTaq PCR buffer (ThermoFisher Scientific), 0.25μM dNTPs, 1μM of each primer (Table 2.16.1), 1U of DreamTaq DNA polymerase (ThermoFisher Scientific), and 1μl of DNA. PCR reactions were run on a T100 Thermocycler (Bio-Rad) with the following program: 1) 94°C for 1.5 min, 2) 94°C for 15 seconds, 3) 60°C for 20 seconds, 4) 68°C for 40 seconds, 5) repeat steps 2-4 36 times, 6) 68°C for 5 min, 7) 4°C hold. Once complete, PCR products was mixed with 3μl

of Orange G loading buffer and separated on a 2.0% agarose gel at 130V with a no-template

control. A 200bp band indicated wildtype *Cecr2*, while a 450bp band represented *Cecr2*^{tm1.1Hemc}.

Table 2.16.1: List of primers used for genotyping

Genotyping	Name	Sequence
Cdc42^{Del}	SA2f	AGACAAAACAACAAGGTCCAGAAAC
	LA1r	CTGCCAACCATGACAACCTAAGTTC
Cecr2^{GT45bic}	Intron7 F4	CCCCATTTATTTGCTTGAGCTG
	Cecr2 Intron7 R4	CACGAACAATGGAAGGAATGA
	pGT1R4	ACGCCATACAGTCCTCTTCACATC
Cecr2^{tm1.1Hemc}	Ingenious SDL2	GTAGCGCCTATTTGTAATGGTCA
	Lox-CECR2_DEL3R	AATGGTGGCGAAATCAACTC
	IngeniousLox 1	TTAGAATAGGTGAGGGAGGAG

2.17: Caco-2 Cell Immunofluorescence

 Immunofluorescence of Caco-2 cells was conducted in ThemoScientific™ BioLite 12-

well and 24-well plates. Circular coverslips (1.5 thickness) were placed in the wells and coated

with Poly-D-Lysine (Sigma) following the manufacturer's protocol. Caco-2 cells were plated in

each well at specific densities (Table 2.9.1). After 24 hours, cells were transfected with wildtype

or variant DNMBP plasmids following the protocol described in section 2.9. Control cells

underwent the transfection protocol without the inclusion of plasmid DNA. Two days after

transfection, media was aspirated from wells and cells were fixed by incubating in ice cold

methanol at -20°C for 20 min. This was followed by three gentle washes in 1X PBS for 5 min

each. Cells were permeabilized by applying 0.1% Triton X-100 (Sigma) in 1X PBS (PBS-TX)

for 15 min at room temperature. After removing the PBS-TX, cells were blocked in 5% skim

milk in 1X Tris-buffered saline with 0.1% Triton X-100 for 30 min. The blocking solution was

removed and primary antibodies (anti-DNMBP, anti-HA, or anti-ZO-1) diluted in fresh blocking

solution were added to wells for one hour at room temperature. Following three washes in PBS-TX, secondary fluorescent antibodies were applied for 30 min. Antibody concentrations are listed in Table 2.11.1. Cells were washed three times with 1X PBS, and then incubated in DAPI for 5 min, diluted to 1:5000 with 1X PBS. After 3 washes for 5 min with 1X PBS, coverslips were carefully removed and mounted onto glass slides using Fluoromount G (SouthernBiotech). Coverslips were sealed with nail polish. Cells were viewed under oil immersion at 60X magnification on a Nikon Eclipse 80i Confocal Microscope. Controls included the absence of primary antibody to test for nonspecific binding of the secondary antibody and the absence of secondary antibody to determine if autofluorescence was occurring.

Chapter 3:

Results

3.1: DNMBP Variant Expression and Protein-Protein Interactions

  To test the function of human and mouse DNMBP variants, each variant gene was subcloned into an expression vector containing an N-terminal HA-tag. Each vector was then individually transfected into a specific cell line. The expression of transfected wildtype DNMBP and its variants was confirmed through Western blot analysis in U-2 OS cells (Figure 3.1.1). Using an antibody against the HA-tag, all constructs were expressed at close to equal levels with the exception of variant 5 (R1024X), which showed higher expression levels possibly due the size of the gene being transcribed (discussed further in section 4.1). This pattern was also seen when HEK293 cells were transfected with variants. The loading control (Tubulin) indicated that protein concentrations were well normalized (Figure 3.1.1)**.** Interestingly, cells transfected with variant 5 showed the lowest concentration of DNMBP when a DNMBP antibody was used for probing (Figure 3.1.1). Since the HA-tag is on the N-terminal of variants, the truncated protein was successfully probed with HA antibody. However, the DNMBP antibody binds to a site after the truncation site of variant 5, thus, only endogenous protein was probed with the DNMBP antibody in cells transfected with variant 5. I was expecting the endogenous DNMBP in variant 5 transfected cells to be at similar levels as non-transfected cells, but it was considerably lower based on the Western blot. The loading control indicated that the protein concentration was slightly lower in variant 5 extract, which may explain this discrepancy.

  Using immunofluorescence, transfection efficiency was determined for the wildtype DNMBP expression vector using HEK293 and U-2 OS cells. Transfection efficiency was ~40% in HEK293 cells and ~60% in U-2 OS cells. In Caco-2 cells, efficiency was about 10% for cells transfected with wildtype DNMBP. Each variant also transfected with an efficiency close to 10%, with the standard deviation being 1.83% in Caco-2 cells.

The transfected constructs were used to test whether the variants affected the strength of endogenous protein interactions with DNMBP. Using co-immunoprecipitations, I focused on MENA and VASP due to evidence of their direct interaction with DNMBP and their association to NTDs (*96*, *147*). Co-immunoprecipitations were first conducted in HEK293 cells and later attempted in U-2 OS cells due to the higher transfection efficiency. However, co-immunoprecipitation of DNMBP and protein partners was inconsistent. Interaction between DNMBP and VASP was verified in non-transfected and WT transfected cells through the pull-down of VASP (Figure 3.1.2), though I was not always able to repeat this after many optimization attempts. Appendix B summarizes optimization attempts and successful reactions. Of 19 reactions, only reactions 1, 3, and 4 showed the interaction between DNMBP and VASP. Reciprocal immunoprecipitation of VASP showed a similar inconsistency. The interaction between DNMBP and MENA was never observed in this study using co-immunoprecipitations.

Fixing cells in formaldehyde, a technique utilized to capture transient protein interactions, led to a very low protein yield after extraction with RIPA lysis buffer. Increasing formaldehyde concentration gradually from 0.4% to 2.0% led to an even lower protein yield. Since a higher concentration of formaldehyde is more likely to cross-link protein complexes than a lower concentration, I used a 0.8% formaldehyde concentration because it led to a protein yield suitable for co-immunoprecipitation. However, it did not lead to co-immunoprecipitation of DNMBP and VASP.

VASP was then co-transfected with DNMBP in hopes that the excess protein would lead to a higher frequency of interactions than with the endogenous protein alone. Overexpression of DNMBP and VASP in cells did not change the inconsistency of capturing interactions with co-

immunoprecipitation. Co-immunoprecipitation in U-2 OS cells also did not result in consistent results. Using more protein, antibody, and/or beads did not lead to more reliable results either.

However, when VASP was immunoprecipitated, MENA was pulled-down even if DNMBP was not pulled-down. This was seen for at least three different co-immunoprecipitation conditions, indicating that the co-immunoprecipitations were successful in showing some interactions, but not for the interactions with DNMBP (Figure 3.1.3).
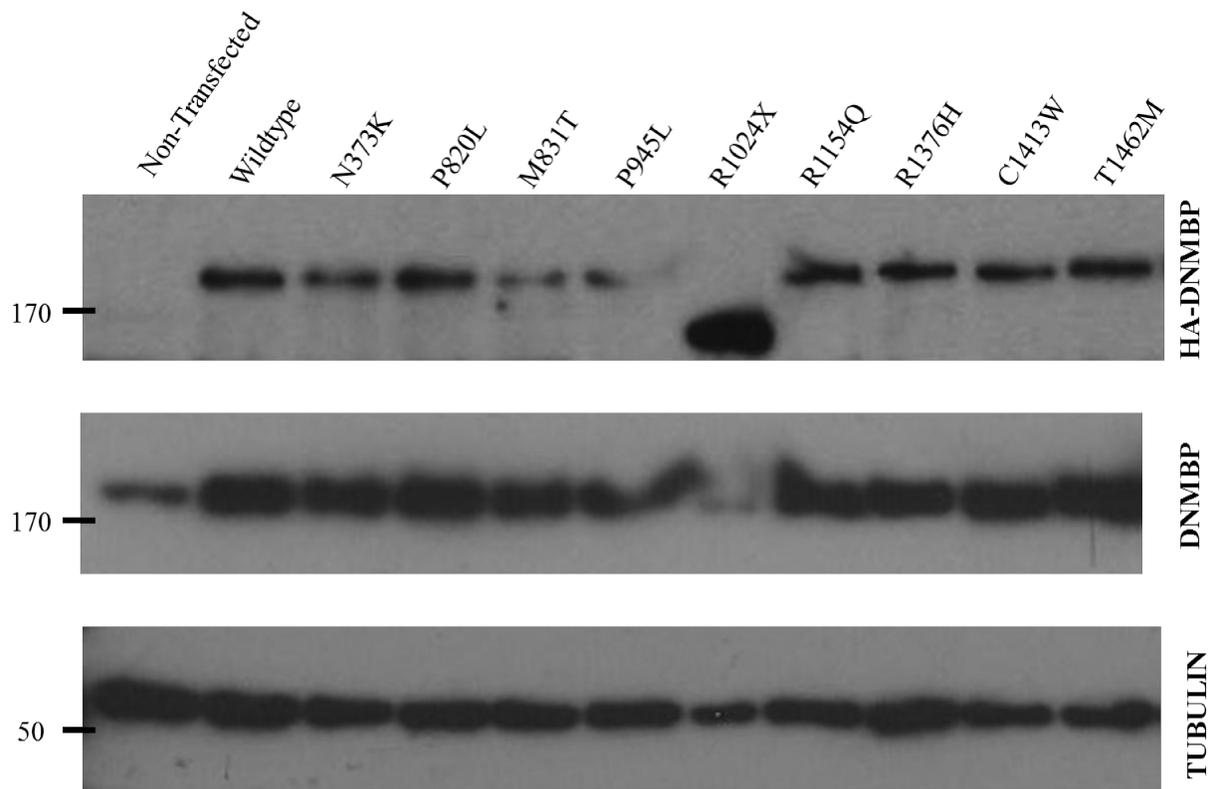
Figure 3.1.1: Expression of transfected DNMBP in U-2 OS cells. U-2 OS cells were transfected with wildtype or variant DNMBP. Protein lysate was extracted and quantified, concentration was equalized, and separated by Western blotting. (Top) Anti-HA was utilized to show expression of transfected protein. (Middle) The same Western blot was stripped and re-probed with anti-DNMBP to show expression of endogenous and transfected DNMBP. (Bottom) Tubulin was probed for as a loading control. This expression analysis was completed twice in U-2 OS cells and once in HEK293 cells.
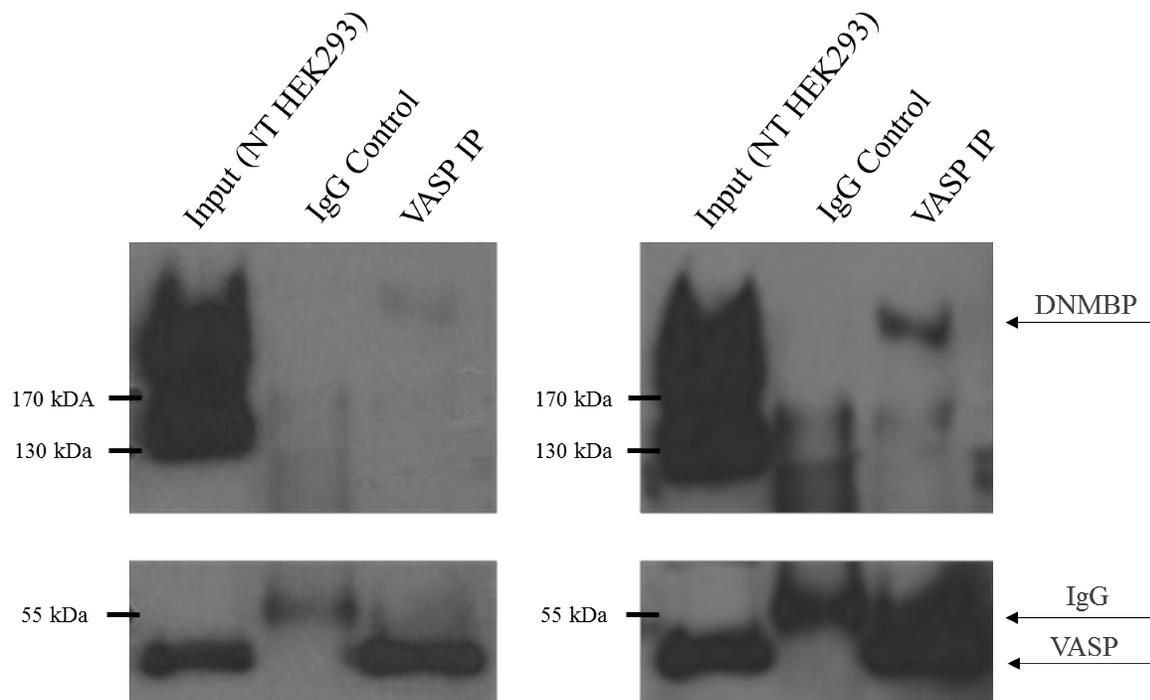
Figure 3.1.2: Immunoprecipitation of endogenous VASP in HEK293 cells leading to co-immunoprecipitation of endogenous DNMBP. VASP was pulled-down leading to a co-precipitation of DNMBP in non-transfected cells. The left panel shows a one minute exposure of the Western blot while the right panel shows a two minute exposure. The input lane shows multiple bands, which are likely isoforms of DNMBP. The band in the IgG control lane, which can also be slightly seen in VASP IP lane, represents antibody lighter chains that dissociate from beads after co-immunoprecipitation. Antibodies were crosslinked to beads, leading to less IgG signal.
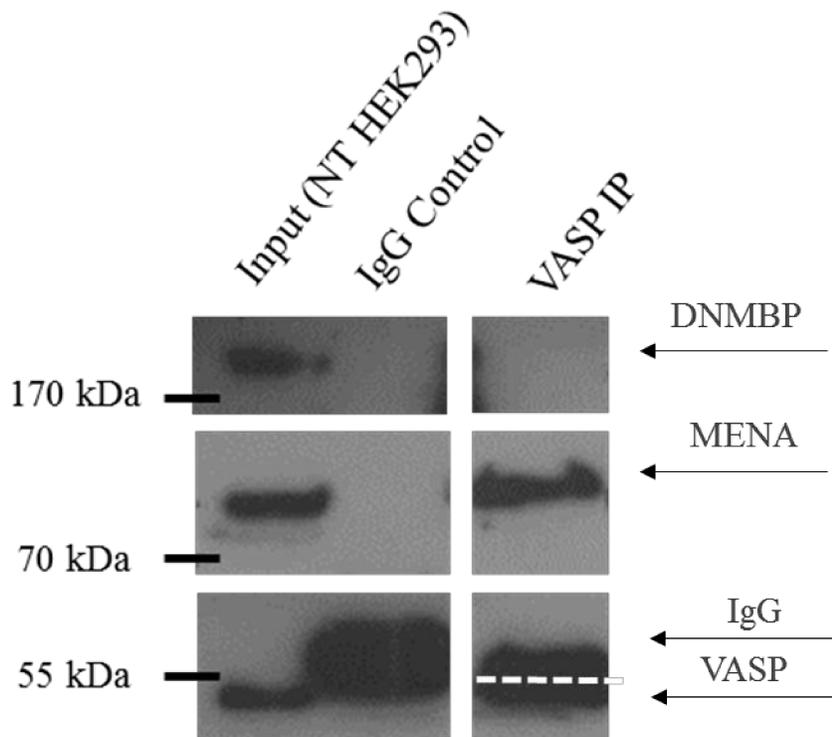
Figure 3.1.3: Immunoprecipitation of endogenous VASP in HEK293 cells. VASP antibody was used to pull-down endogenous VASP in non-transfected HEK293 cells. DNMBP was not pulled-down. However, MENA, an interacting partner of both VASP and DNMBP, was co-immunoprecipitated with VASP, indicating a functioning co-immunoprecipitation experiment. The band in the IgG control lane, which can also be seen in VASP IP lane, represents antibody lighter chains that dissociate from beads after co-immunoprecipitation. Note: The amount of input protein was reduced relative to Figure 3.1.2 to reduce signal of endogenous proteins.

3.2: CDC42 Activation Assay

DNMBP is a major activator of the Rho GTPase CDC42. If a variant of DNMBP is not fully functional, there would be less active CDC42 in the cells, leading to downstream abnormalities. The CDC42 G-LISA Activation Assay (Cytoskeleton, Inc.) allows for the pull-down of active GTP-bound CDC42 from cell protein lysate. The amount of active CDC42 is quantified through measurement of absorbance after the addition of HRP substrate. In the first assay attempted, I looked at the effect of serum starvation on HEK293 cells, which was suggested by the protocol to bring down the level of endogenously active CDC42 (Figure 3.2.1). Cells were grown in serum free media (48 or 24 hours prior to assay) or normal complete media. The control was purified constitutively active CDC42 protein supplied with the kit to ensure that the values given by the assay were within the expected range. Replicates were performed and the assay showed the opposite of what was expected – serum starved cells had an increase in active CDC42 relative to cells in normal media. The positive control was under the range given by the manufacturer. Due to the expense of the assay and the need for optimization, replicates were not conducted after the first attempt. The protocol suggested increasing the HRP substrate incubation time from 15 to 25 min. This was done for the next assay, which included non-transfected cells and cells transfected with either wildtype DNMBP, variant N373K (randomly chosen), or variant R1024X (likely to lead to decreased CDC42 activation) (Figure 3.2.2A). The increased incubation time in HRP substrate led to increased absorbance values. Cells transfected with wildtype DNMBP had a lower amount of active CDC42 than non-transfected cells. Variant N373K showed a similar absorbance value to non-transfected cells while variant R1024X had a slightly higher absorbance. Since the absorbance values were outside of the linear range of 1.0, the assay was repeated with non-transfected and wildtype transfected HEK293 cells with an HRP

substrate incubation time of 17 min (Figure 3.2.2B). Still, wildtype transfected cells showed a slight decrease in active CDC42 than non-transfected cells. The assay was then conducted with U-2 OS cells to determine if changing the cell line would have an effect on CDC42 activation in response to overexpression of DNMBP. U-2 OS cells were used due to their higher transfection efficiency. Relative to non-transfected U-2 OS cells, cells transfected with wildtype DNMBP or variant R1024X showed somewhat decreased levels of active CDC42 (Figure 3.2.2C). I was expecting there to be a decrease in activation with R1024X compared to wildtype DNMBP, which was seen, though wildtype DNMBP should have had higher activation levels relative to non-transfected cells and R1024X should have had similar activation levels to non-transfected cells. Since replicates were not conducted, we cannot say that there was a definite change, especially between values very close to each other. Based on the optimization attempts, cost of the kit, and the availability of other techniques to test CDC42 activation, it was not the best option to continue with the CDC42 G-LISA assay to determine DNMBP variant function.

Figure 3.2.1: Serum starvation of HEK293 cells and analysis of CDC42 activation. HEK293 cells were grown in normal complete media or serum starved (serum-free media, SF). The G-LISA CDC42 Activation Assay was carried out following the manufacturer's protocol (Cytoskeleton Inc.). Absorbance was measured at 490nm. Control represents constitutively active CDC42. Replicates were conducted for (A), while all other assays did not contain replicates, thus, error bars are not shown. (A) HEK293 cells grown in normal media and serum-free media (48 hours). (B) HEK293 cells grown in normal and serum-free media (24 hours). HRP substrate incubation was increased to increase absorbance values.
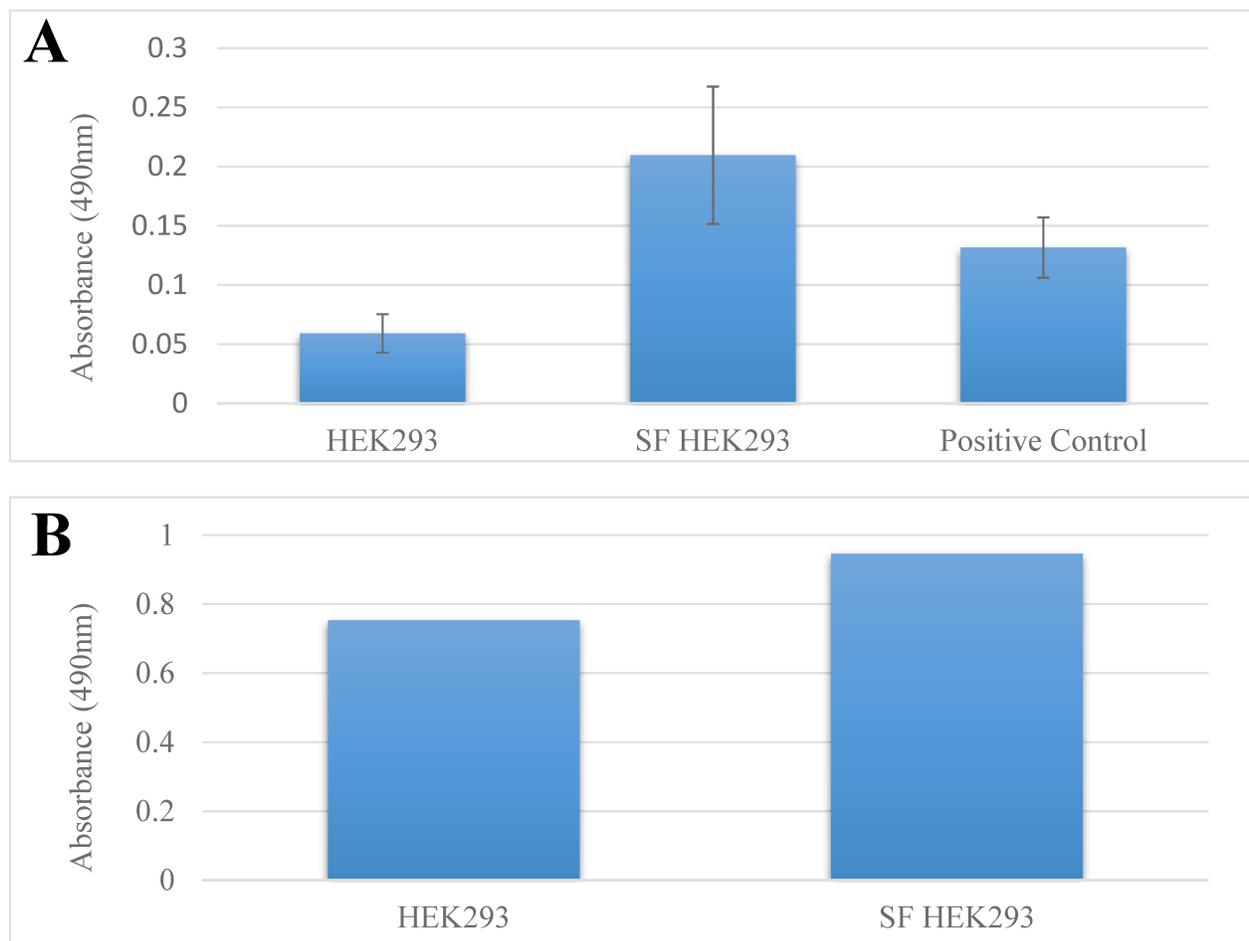
Figure 3.2.2: Transfection of HEK293 and U-2 OS cells and analysis of CDC42 activation.
HEK293 and U-2 OS cells were grown in normal complete media and transfected with DNMBP. The G-LISA CDC42 Activation Assay was carried out following the manufacturer's protocol (Cytoskeleton Inc.). Absorbance was measured at 490nm. Control represents constitutively active CDC42. Replicates were not conducted, thus, error bars are not shown. (A) Non-transfected, and transfected HEK293 cells. Note: HRP substrate incubation was increased leading to an absorbance past the linear range of 1.0. (B) Non-transfected and wildtype transfected HEK293 cells. HRP substrate incubation time was decreased, leading to decrease in absorbance measurement. (C) Non-transfected and transfected U-2 OS cells.
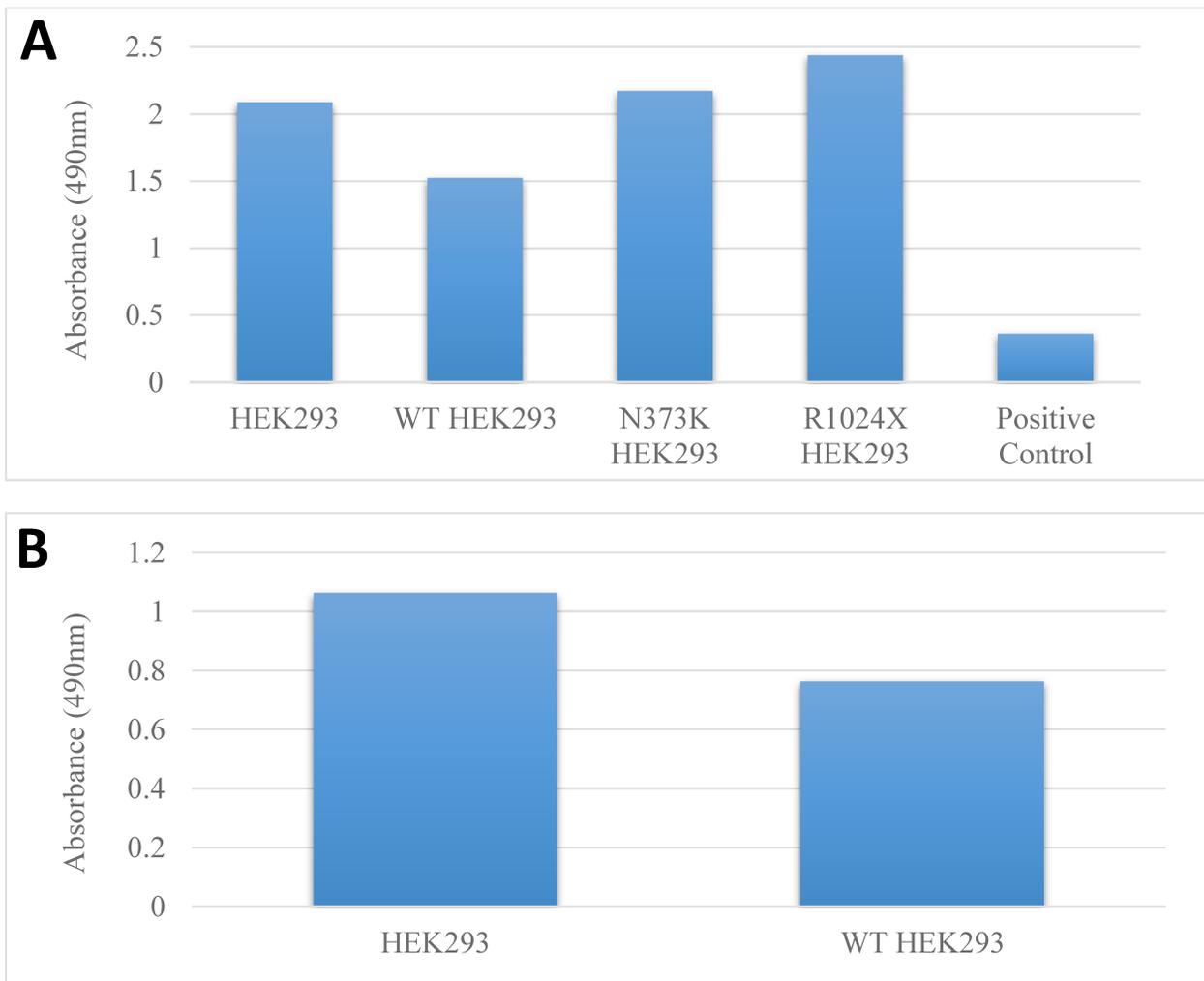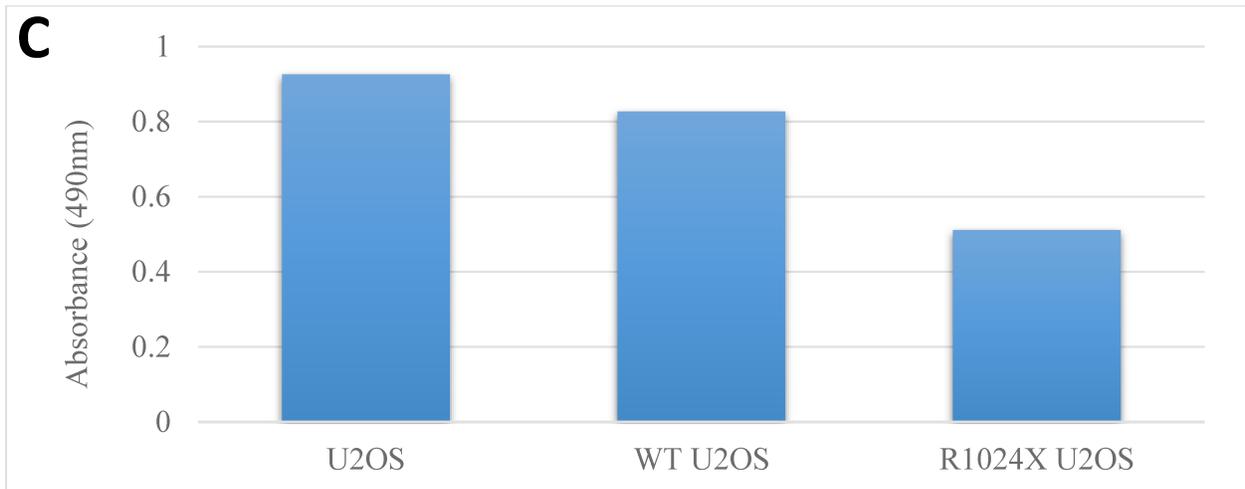
Figure 3.2.2 continued

3.3: Dense-core Vesicle Analysis

Increased activation of CDC42 through increased expression of DNMBP has been shown to lead to an increase in the number of neural cell-specific dense-core vesicles in PC12 cells (*143*). Here, I attempted to show that the expression of wildtype HA-DNMBP in PC12 cells also leads to an increased activation of CDC42 by quantifying the number of dense-core vesicles. This would then be compared to the number of dense-core vesicles in cells transfected with each variant. If a variant was activating CDC42 normally, the number of dense-core vesicles would be comparable to wildtype transfected cells. A variant with a mutation affecting CDC42 activation would show a significant change in the number of dense-core vesicles relative to cells transfected with wildtype DNMBP.

This experiment was originally to be conducted in PC12 cells similar to the original experiment by Sato et al. (*143*), however, our PC12 cells were growing abnormally. Instead of a flat morphology, the cells were spherical, with the nuclei occupying most of the cell space.

Staining these cells did not allow for the observation of dense-core vesicles. After attempting to optimize the conditions by growing the PC12 cells in various conditions (i.e. cell media, plate coating) and trying two different PC12 lines obtained from other labs, I was not able to grow them in a manner that would allow for observation of dense-core vesicles. Thus, I opted for the use of another neural cell-line, SK-N-BE(2), which was shown to express dense-core vesicles (Figure 3.3.1).

When anti-Chromogranin A (CgA) was utilized to stain dense-core vesicles, the vesicles were only individually distinguishable in some cells. Those cells that did not display individual vesicles were still stained by the CgA antibody. It is possible that the dense-core vesicles were being stained, but the resolution of the microscope was not high enough to resolve the vesicles from each other. Transfection efficiency was less than 20% with SK-N-BE(2) cells, thus cells that were transfected successfully and that showed clearly distinguishable vesicles were rare. For this reason, I measured the mean fluorescence intensity for each cell, as was done by Hao et al. in a similar experiment using PC12 cells (*135*), instead of counting individual vesicles. The mean intensities were then averaged for each variant. Theoretically, if there are more dense-core vesicles, there should be more staining by the CgA antibody and a higher mean intensity. The analysis was first conducted with non-transfected cells and cells transfected with wildtype DNMBP and two variants likely to affect CDC42 activation (P820L and R1024X) (Figure 3.3.2). Figure 3.3.3 shows wildtype transfected cells along with non-transfected cells. Non-transfected cells showed the highest average mean intensity (Figure 3.3.4). However, variation was very high for each measurement set, as shown by the standard deviation bars in Figure 3.3.4. Thus, I was not able to make any conclusions based on these results. The HA-DNMBP staining was also utilized for normalizing CgA fluorescence intensity data from cells that were transfected, though

this did not reduce the amount of variation. Due to the high variation in fluorescence intensity, other variants were not analyzed using this method.

Interestingly, in almost all cells observed, wildtype HA-DNMBP and variant DNMBP seemed to be forming aggregations of protein within the cytoplasm (Figure 3.3.2). However, this was not ever seen in cells transfected with variant 5 (R1024X), indicating that it could be related to localization abnormalities. These aggregations of protein were not observed with endogenous DNMBP (Figure 3.3.1). The aggregates did not overlay CgA staining, thus, they are likely not indicative of dense-core vesicles (see variant P820L in Figure 3.3.2).
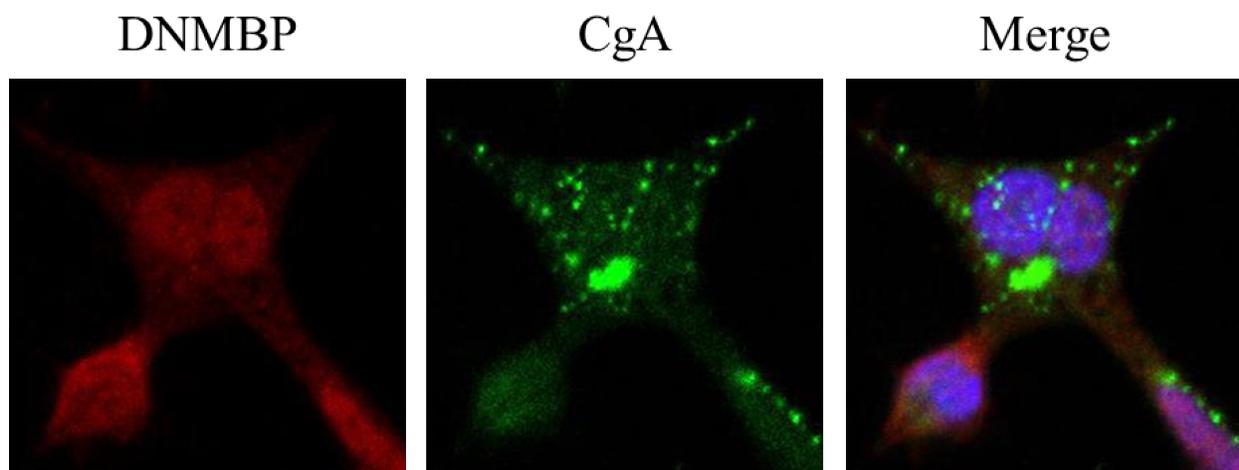
Figure 3.3.1: Immunofluorescence staining of DNMBP and dense-core vesicles in non-transfected SK-N-BE(2) cells. Cells were grown on coverslips and immunofluorescence was conducted. Chromogranin A (CgA), a protein marker of dense core-vesicles, was stained to show the presence of dense-core vesicles (Green) in non-transfected cells. A DNMBP antibody was used to show endogenous DNMBP (Red). DAPI was utilized to stain the nucleus (Blue). Images were taken with a confocal microscope at a magnification of 60X and processed through Photoshop.

Figure 3.3.2: Visualization of dense-core vesicles in transfected SK-N-BE(2) cells. Cells were grown on cover slips and transfected with either wildtype DNMBP or variant P820L or R1024X. Immunofluorescence was conducted, staining for HA-DNMBP (Red) and Chromogranin A (Green). DAPI was utilized to stain the nuclei (Blue). The merge images show the overlay of HA-DNMBP, CgA, and DAPI. Images were taken with a confocal microscope at a magnification of 60X and processed through Photoshop.
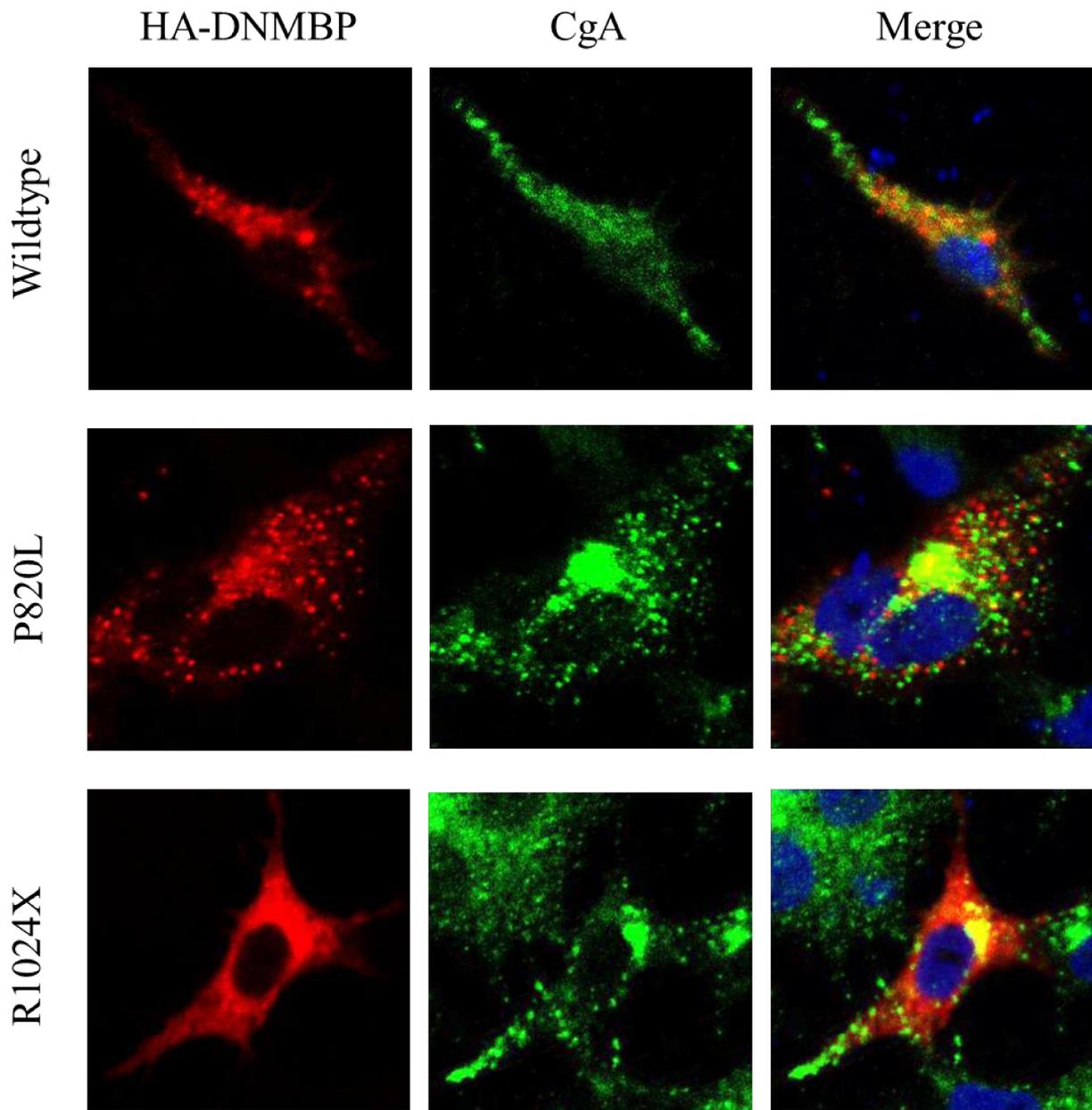
Figure 3.3.3: Visualization of dense-core vesicles in non-transfected and transfected SK-N-BE(2) cells. Cells were grown on cover slips and transfected with wildtype DNMBP. Non-transfected and successfully transfected cells are shown side-by-side for comparison of dense-core vesicle staining. Immunofluorescence was conducted, staining for HA-DNMBP (Red) and Chromogranin A (Green). DAPI was utilized to stain the nuclei (Blue). The merge images show the overlay of HA-DNMBP, CgA, and DAPI. Images were taken with a confocal microscope at a magnification of 60X and processed through Photoshop.
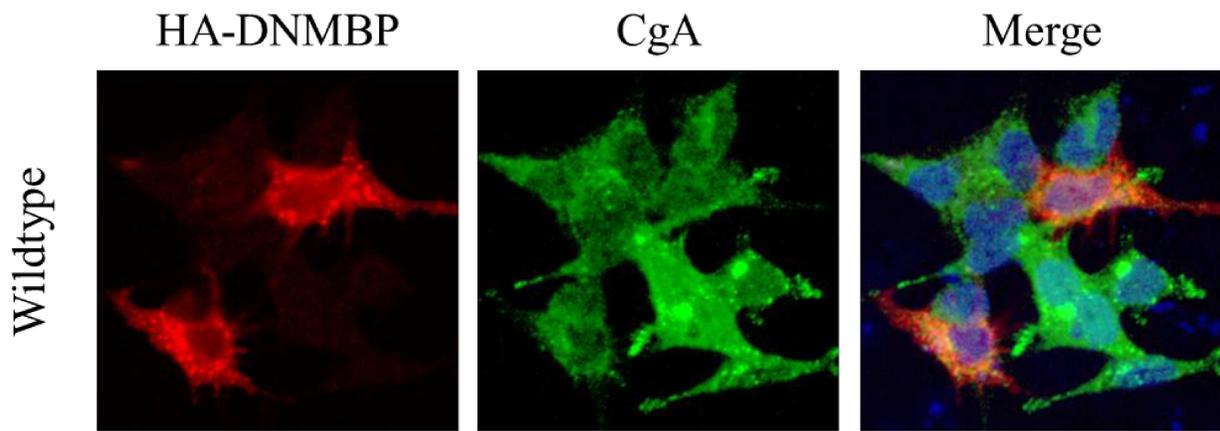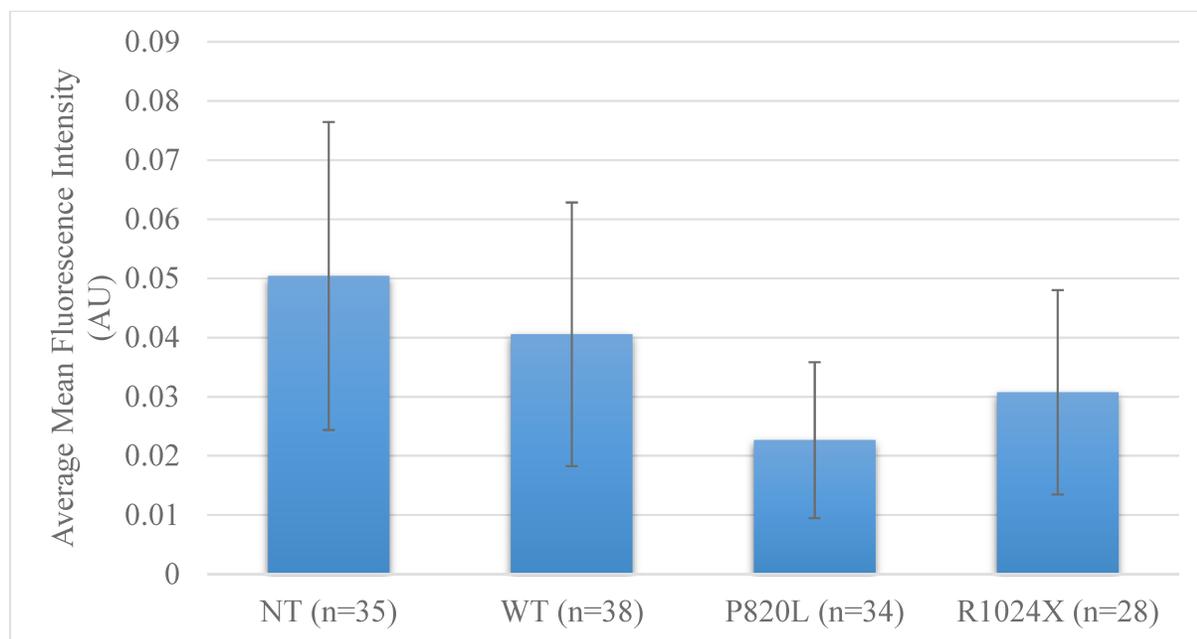
Figure 3.3.4: Average mean fluorescence intensities of Chromogranin A in SK-N-BE(2) cells. Non-transfected (NT) cells and cells transfected with wildtype (WT) DNMBP or variants were stained with CgA antibody through immunofluorescence. Using a confocal microscope, images were obtained of cells, and the mean fluorescence intensity was measured for each cell and averaged. The number of cells observed for each group is indicated by n. Error bars represent standard deviation.

3.4: Analysis of Cdc42 and Cecr2 Genetic Interaction

The interaction of genes was also studied *in vivo*. To determine if *Dnmbp* was modifying *Cecr2* through *Cdc42*, a mouse model was produced in which there was a heterozygous *Cdc42* deletion and homozygous *Cecr2* mutation on a dominant FVB/N background. A homozygous *Cecr2* gene-trap (*Cecr2*[GTbic45]) mutation alone does not result in exencephaly in the resistant FVB/N strain, while a homozygous deletion of *Cecr2* (*Cecr2*[tm1.1Hemc]) results in a penetrance of only ~12%, both significantly less than what is observed on the BALB/C background (*55*). The

human variants of *DNMBP* being studied were only found in one allele of an NTD fetus. Therefore, at most a deleterious DNMBP variant affecting CDC42 activity would reduce activation of CDC42 by 50%. Consequently in the mouse model, a 50% *Cdc42* deficiency would mimic a decrease in DNMBP function that may be occurring when a DNMBP variant is deleterious. Thus, if lowered CDC42 activity due to a deleterious DNMBP variant is causing exencephaly in BALB/C mice when *Cecr2* is deficient, a 50% reduction in *Cdc42* should also lead to exencephaly in resistant FVB/N mice with a *Cecr2* deficiency.

To test this, FVB/N mice with a heterozygous *Cdc42* deletion mutation ($Cdc42^{Del/+}$) and either a homozygous gene-trap mutation ($Cecr2^{GTbic45/GTbic45}$) or *Cecr2* gene-trap and deletion mutation ($Cecr2^{GTbic45/tm1.1Hemc}$) were produced (Figure 2.15.2). In total, 105 embryos were examined in the group containing just the *Cecr2* gene-trap mutation (Table 3.4.1) while 55 embryos were examined in the group containing both *Cecr2* gene-trap and deletion mutations (Table 3.4.2). Exencephaly was not observed in any heterozygous *Cdc42* deletion mutants. Genotypes of embryos were not significantly different from expected ratios (P>0.05), thus, *Cdc42* heterozygous mutants are likely not being lost early in development. Exencephaly was observed in one embryo with wildtype *Cdc42* and a $Cecr2^{GTbic45/tm1.1Hemc}$ mutation, which would be expected in some embryos with the deletion mutation since a homozygous deletion ($Cecr2^{tm1.1Hemc/tm1.1Hemc}$) results in exencephaly in ~12% embryos with a FVB background.

Thus, this data indicates that reducing *Cdc42* in combination with *Cecr2* deficiencies does not result in NTDs on the resistant FVB/N background.

Table 3.4.1: Analysis of exencephaly penetrance of embryos produced through crossing a heterozygous *Cdc42* deletion mutant with a homozygous *Cecr2^{GTbic45}* mutant (FVB/N background).

Genotype	Exencephaly	Normal	Total
Cdc42$^{Del/+}$; Cecr2$^{GTbic45/+}$	0	25	25
Cdc42$^{Del/+}$; Cecr2$^{GTbic45/GTbic45}$	0	26	26
Cdc42$^{+/+}$; Cecr2$^{GTbic45/+}$	0	28	28
Cdc42$^{+/+}$; Cecr2$^{GTbic45/GTbic45}$	0	26	26
Total	0	105	105

Table 3.4.2: Analysis of exencephaly penetrance of embryos produced through crossing a compound heterozygous *Cdc42* and *Cecr2* deletion mutant with a homozygous *Cecr2^{GTbic45}* mutant (FVB/N background).

Genotype	Exencephaly	Normal	Total	% Exencephaly
Cdc42$^{Del/+}$; Cecr2$^{GTbic45/+}$	0	10	10	0.00%
Cdc42$^{Del/+}$; Cecr2$^{GTbic45/tm1.1Hemc}$	0	15	15	0.00%
Cdc42$^{+/+}$; Cecr2$^{GTbic45/+}$	0	10	10	0.00%
Cdc42$^{+/+}$; Cecr2$^{GTbic45/tm1.1Hemc}$	1	19	20	5.26%
Total	1	54	55	1.85%

3.5: Localization of DNMBP Variants

Mislocalization of a protein can affect its function. Mislocalization would therefore provide evidence for functional abnormalities in DNMBP variants. The localization of each DNMBP variant was tested to determine if it concentrated to similar areas of the cell as wildtype DNMBP. DNMBP is known to localize to apical cell-cell junctions of epithelial cells with some perinuclear localization (*117*). This was confirmed using the epithelial cell line Caco-2. Non-transfected Caco-2 cells stained by immunofluorescence with a DNMBP antibody showed strong endogenous expression at cell membranes (Figure 3.5.1). Cytoplasmic expression of DNMBP was strongest close to the nucleus. The ZO-1 antibody was used as a control for membrane binding, and as expected, ZO-1 expression was observed only at the membrane and overlapped

with DNMBP. An HA-tag antibody was utilized to stain only transfected DNMBP, which have an N-terminus HA-tag. Non-transfected control cells stained with HA-tag antibody did not show any signal indicating the antibody was not binding non-specifically (data not shown). Primary and secondary antibody controls also did not show signal with transfected cells indicating that the secondary was specific and cells were not autofluorescing (data not shown).

Transfection of wildtype (WT) HA-tagged DNMBP into Caco-2 cells resulted in similar localization patterns as endogenous DNMBP detected with anti-DNMBP in non-transfected cells. Transfected WT DNMBP stained with an HA-tag antibody was concentrated at the membrane with some perinuclear localization (Figure 3.5.2). Expression of transfected DNMBP between cells was variable. Some cells showed higher concentration of transfected DNMBP in the cytoplasm and membrane while other cells showed lower concentrations (Figure 3.5.3). As well, in some images the entire membrane does not show a consistent staining of DNMBP. This is because the Caco-2 cells are not completely flat on the slides. Since confocal microscopes focus on a single section or plane at a time, a portion of the apical surface may not be in focus leading to only a part of the membrane stained with DNMBP being observed. This prevented the analysis of subtle difference using the fluorescence intensity.

For each of the nine variants, this experiment was repeated twice, observing at least 50 cells in total. With the exception of variant R1024X, all variants of DNMBP transfected into Caco-2 cells showed the expected localization pattern in each cell observed (Figure 3.5.2). Variant 5 (R1024X) did not show membrane localization in any of the 50 transfected Caco-2 cells observed. Instead, the transfected R1024X protein was spread throughout the cell with a uniform consistency, indicating improper localization.

Spots of intense fluorescence were observed in the cytoplasm of about 10-20% of the Caco-2 cells imaged that were transfected with either wildtype DNMBP or a variant. They were seen at a similar rate in HEK293 cells. These aggregates were not observed in cells transfected with variant 5 (R1024X) or with endogenous protein in non-transfected cells. Similar aggregations of transfected DNMBP and other transfected proteins was seen by Otani et al. (2006) who stated it was excess protein. I also saw aggregations of other transfected proteins in the cytoplasm, such as *Cecr2* in an unrelated experiment, but not endogenous protein.

Overall, this analysis indicated that variant 5 (R1024X) had abnormalities in localizing to the apical membrane of Caco-2 cells and confirmed that this is a deleterious DNMBP allele from a patient with anencephaly.

Figure 3.5.1: Localization of endogenous DNMBP in Caco-2 cells. Through immunofluorescence, localization of the endogenous DNMBP (Green) in Caco-2 cells was determined. A ZO-1 antibody was used to stain the membrane in Red to show where DNMBP is expected to localize. The merge images show the overlay of ZO-1, HA-DNMBP, and DAPI (Blue). Images were taken with a confocal microscope at a magnification of 60X and processed through Photoshop.

Figure 3.5.2: Localization of transfected DNMBP in Caco-2 cells. Wildtype DNMBP and variants were transfected individually into Caco-2 cells. Through immunofluorescence, localization of the transfected DNMBP was determined. A ZO-1 antibody was used to stain the membrane in red to show where DNMBP is expected to localize. An HA-tag antibody stained only the transfected DNMBP (Green). The merge images show the overlay of ZO-1, HA-DNMBP, and DAPI (Blue). Images were taken with a confocal microscope at a magnification of 60X and processed through Photoshop.
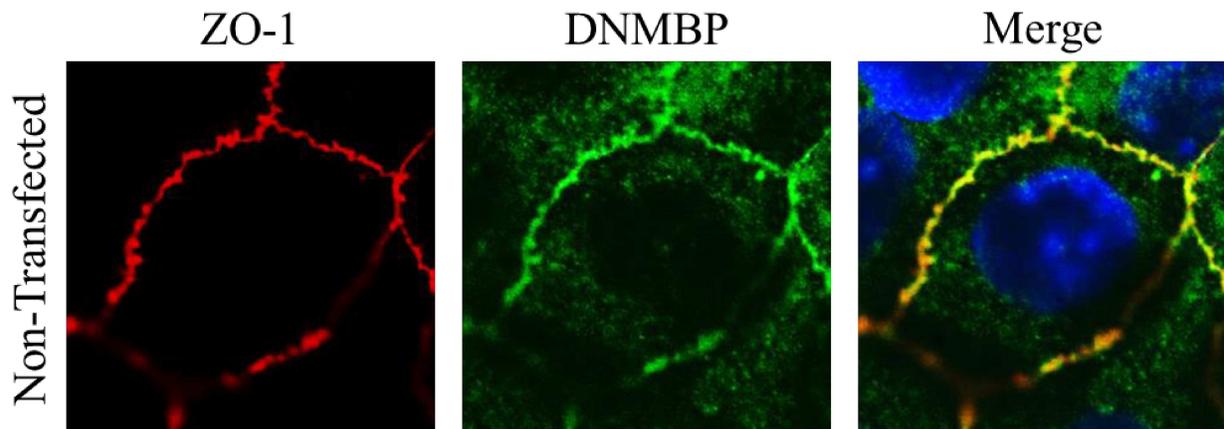
Figure 3.5.2 continued

ZO-1　　　HA-DNMBP　　　Merge

M831T

P945L

R1024X

R1154Q

Figure 3.5.2 continued

	ZO-1	HA-DNMBP	Merge
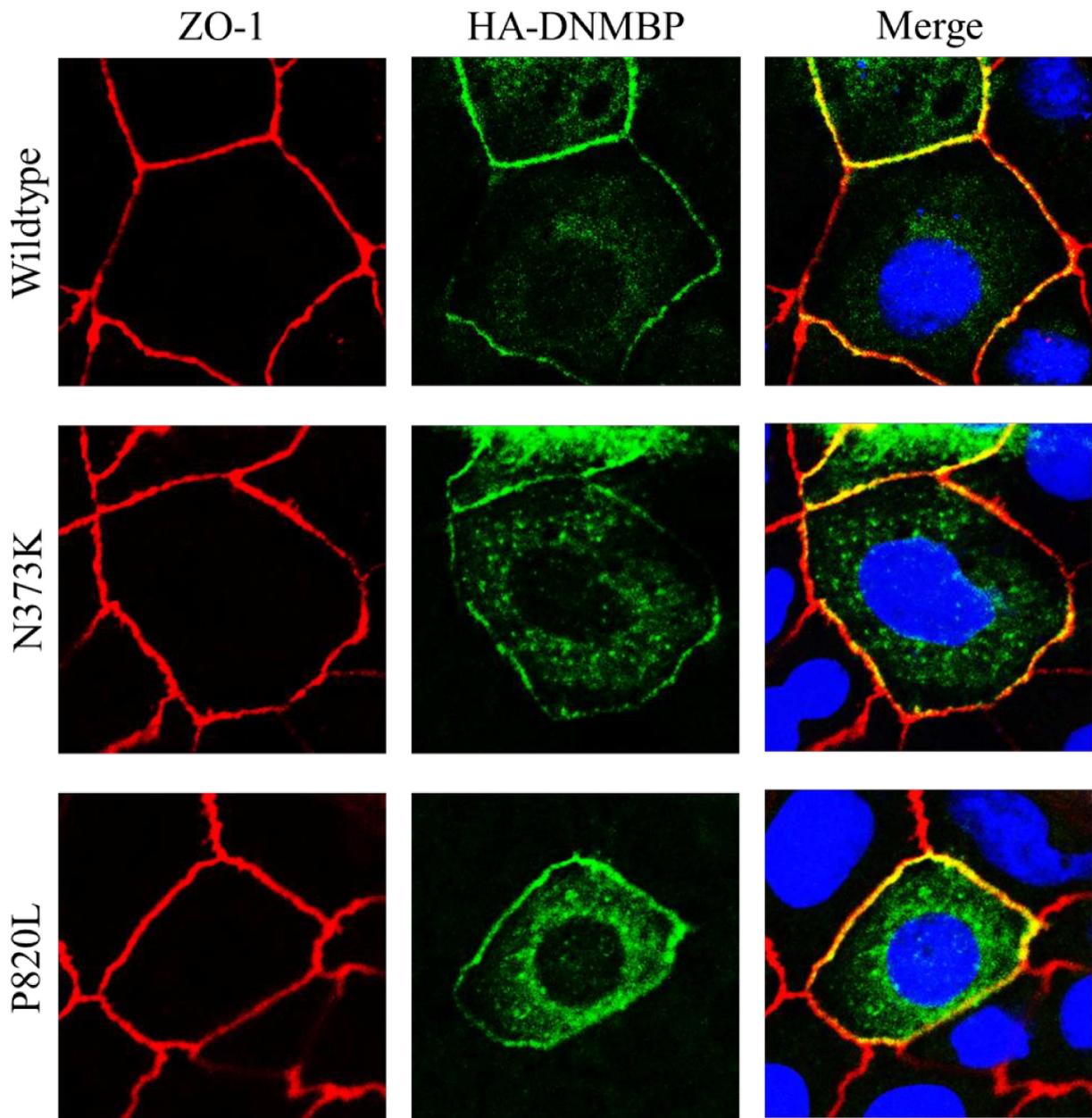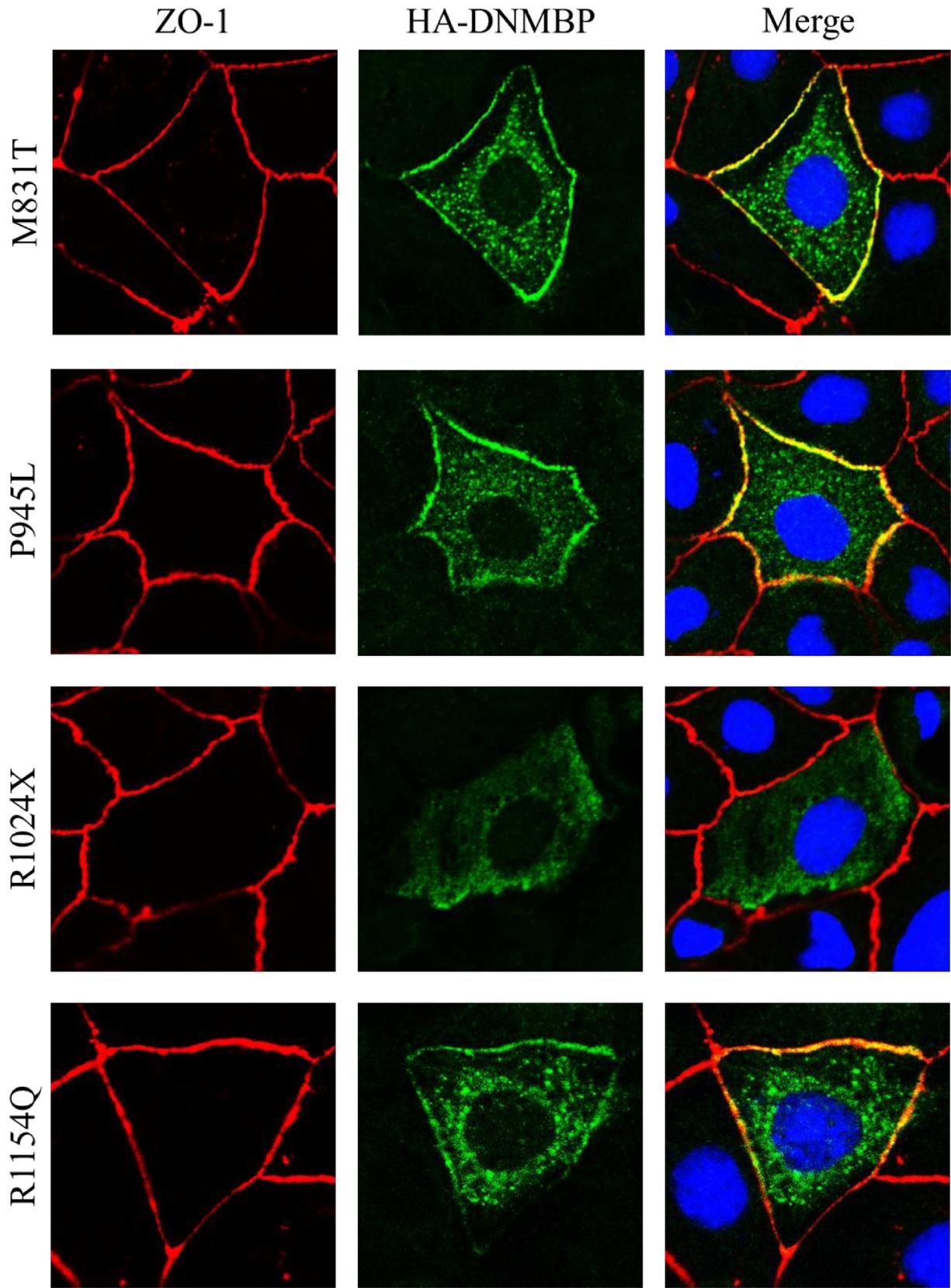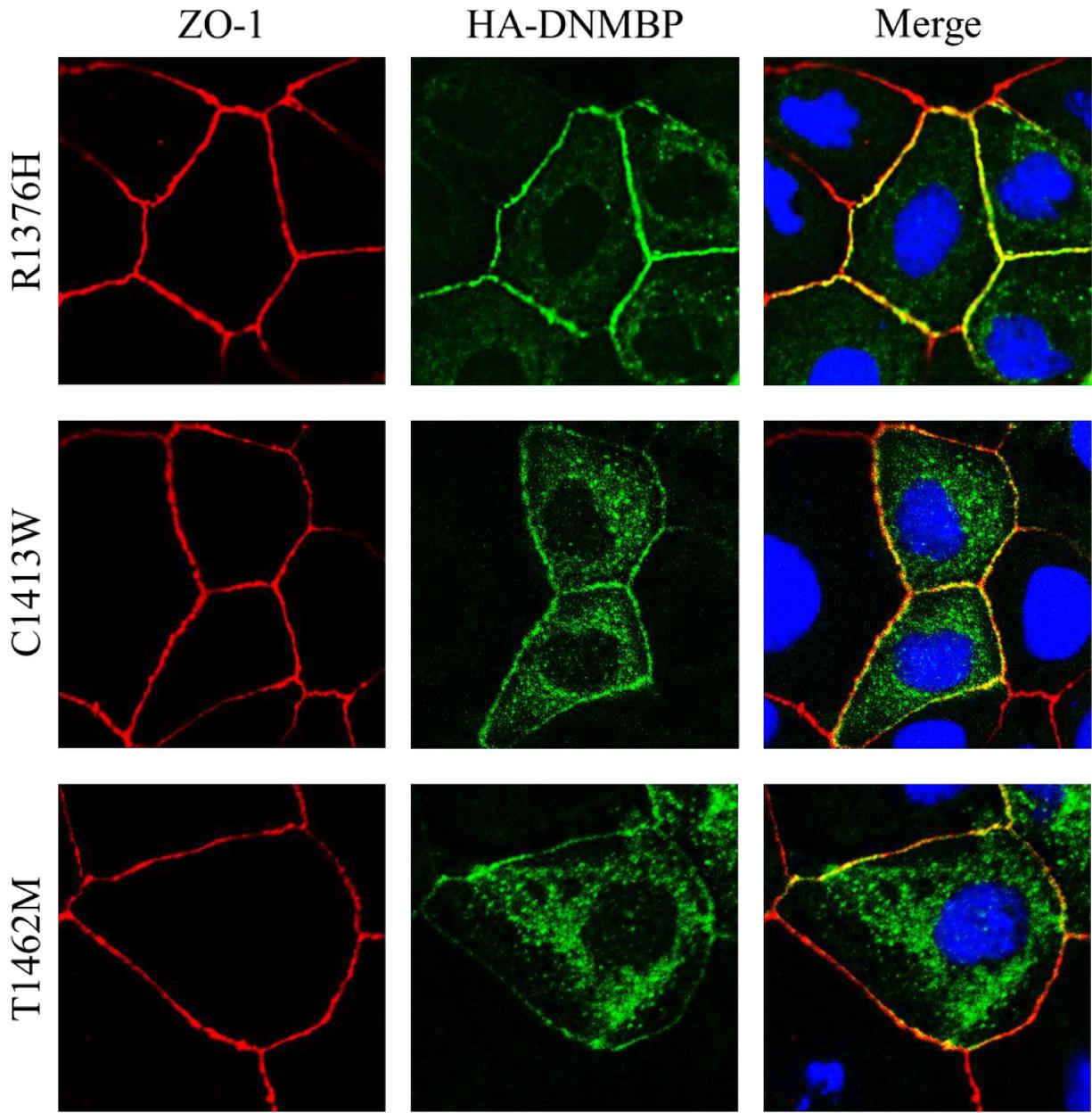
Figure 3.5.3: Variation in expression of transfected DNMBP in Caco-2 cells. Wildtype DNMBP was transfected into Caco-2 cells. Through immunofluorescence, expression of the transfected DNMBP can be observed. A ZO-1 antibody was used to stain the membrane in red to show where we expect DNMBP to localize. An HA-tag antibody stained only the transfected DNMBP (Green). The merge images show the overlay of ZO-1, HA-DNMBP, and DAPI (Blue). The relative amount of DNMBP expression can be associated with the intensity of fluorescence. Cells regularly showed noticeable differences in fluoresce intensity, indicating that they are likely not expressing the same amount of transfected DNMBP. Images were taken with a confocal microscope at a magnification of 40X and processed through Photoshop.
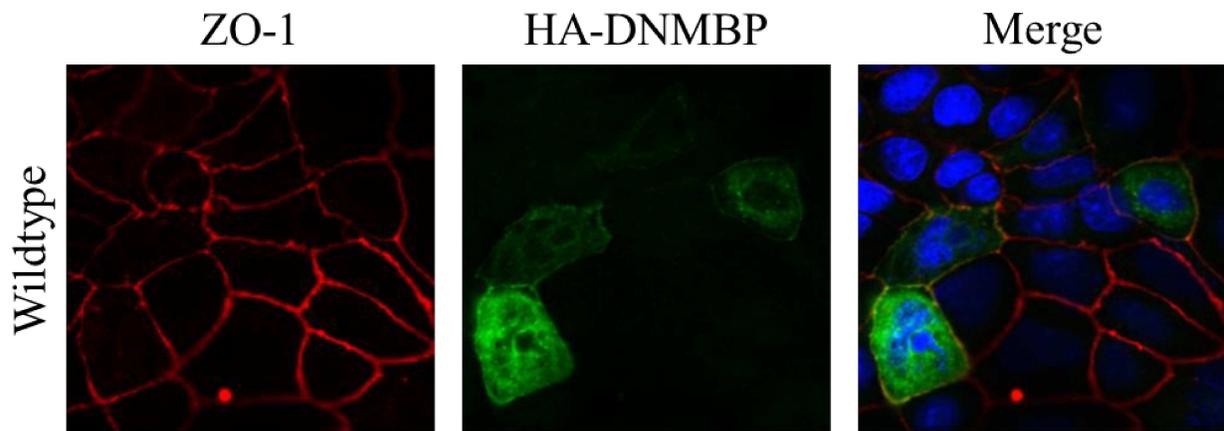
Chapter 4:

Discussion

The process of neurulation is quite complex and can lead to neural tube defects if disrupted. In this research, I focused on the genetic factors leading to NTDs, specifically with the *Cecr2* gene. In mice, mutations to *Cecr2* results in exencephaly in the BALB/cCrl strain but not in the FVB/N strain, indicating the presence of at least one modifier gene effecting *Cecr2*. A linkage analysis indicated that a strong modifier locus is located on chromosome 19 of mice. Genes in this region were sequenced and variants were found that differed between the strains of mice in coding or expression. The top candidate variant genes were then sequenced in human cranial NTD fetuses. Mouse and human candidate genes were ranked based on their predicted deleteriousness. One of the top candidate genes is *Dnmbp*. Eight variants were found in human NTD probands and one mouse variant from the susceptible BALB/cCrl strain, all predicted to lead to a deleterious phenotype. Here, I conducted functional analyses of each variant in order to determine if *Dnmbp* is a modifier of *Cecr2* and an NTD susceptibility gene.

4.1: DNMBP Variant Expression and Protein-Protein Interactions

Through transfection of each DNMBP variant into HEK293 and U-2 OS cell lines, I was able to show that each variant produced was being expressed and that variants could be distinguished from endogenous DNMBP using the HA-tag. Variant 5 (R1024X) showed a higher level of expression based on Western blot band intensity relative to all other variants, but the loading control indicated that the total lysate concentration was a bit less for variant 5 compared to other variants. Thus, if there was to be any difference in band intensity, the band for variant 5 should have been less intense, similar to the loading control rather than more intense compared to other variants. The concentration of the plasmid DNA for each variant was measured and equalized multiple times to ensure that there was little variation in concentrations between

variants. Transfection efficiency between variants was also similar. For these reasons, differences in plasmid DNA concentration is likely not the cause of the higher expression of variant 5. One possibility for this discrepancy in expression could be the size of the gene being transcribed and translated. Since variant 5 leads to an early stop codon close to the start of the BAR domain leading to only ~66% of the gene being transcribed and translated, theoretically the time it would take to process variant 5 is much shorter than the other full-length variants. Since it is not being expressed from the genome, the increase in expression of variant 5 is likely not a feedback mechanism occurring where the cell senses that variant 5 is not functioning normally, leading to an increase in its expression to compensate. If we were to mutate the endogenous *DNMBP* to lead to a similar early stop codon as variant 5, feedback could possibly explain the change in expression. However, since these genes are being expressed from plasmids with unrelated promoters, there is likely no way for the cell to increase the expression of variant 5 through a feedback mechanism. In bacterial cells, the expression of truncated genes has also been shown to lead to an increase in expression relative to full length genes (*188*). If sensitive assays are being conducted that quantify the functionality of variants, such as the G-LISA CDC42 assay, it would be best if the expression of variant 5 was optimized by decreasing the amount of plasmid DNA being transfected, thereby normalizing the amount of protein produced by each variant.

Assessing the ability of variant genes found in human NTD samples to maintain protein-protein interactions has been utilized by some researchers to successfully associate variants to NTD susceptibility. For instance, Yang et al. (*189*) analyzed the PCP gene *Sec24b*, which when mutated in mice leads to craniorachischisis (*190*). Four nonsynonymous *SEC24B* variants were found in a cohort of 163 Chinese fetuses with NTDs and were not found in matched controls or

the Genome 1000 database (*189*). Through transfection of variants and co-immunoprecipitation, it was shown that two of the four variants disrupt the interaction between SEC24B and VANGL2, a core PCP protein also associated with craniorachischisis in mice (*26*). Thus, co-immunoprecipitations have been successfully used to provide evidence of how mutations in *SEC24b* in humans may lead to NTDs due to functional disruptions in the interaction with VANGL2.

Similarly, I tried to determine if human and mouse DNMBP variants were able to maintain protein-protein interactions with known partners, though I was unsuccessful due to the inconsistency of the co-immunoprecipitation experiments. Immunoprecipitation of VASP only co-immunoprecipitated DNMBP a few times out of numerous attempts. Pull-down of DNMBP and MENA together did not ever occur, though the majority of the co-immunoprecipitations I conducted focused on DNMBP and VASP. While some DNMBP co-immunoprecipitation experiments by others have used the endogenous protein (*96*), many have instead only co-immunoprecipitated transfected fusion proteins of DNMBP and partners (*96*, *116*, *117*). It is easier to co-immunoprecipitate these fusion proteins since transfection of the gene leads to higher than normal expression levels. Some researchers used fusion proteins of DNMBP in their assays due to the low solubility of the BAR domain (*96*). Thus, it is possible that low solubility of DNMBP is causing inconsistencies in co-immunoprecipitation experiments. Also, interactions between DNMBP and MENA or VASP may only occur in specific conditions within the cell that are difficult to replicate *in vitro*. While I tried numerous optimization reactions, capturing the interaction may require a very specific balance of conditions (i.e. salt concentration, pH, detergent strength) that were not attempted. The protocols of successful co-immunoprecipitations of DNMBP and partners conducted in other studies (*117*, *122*) were attempted here as well, but

did not ever lead to successful results. The only major difference between those studies and this one was the type of beads used during the immunoprecipitation. For example, Otani et al. (*117*) used Sepharose 4FF beads (GE Healthcare) while I used magnetic Dynabeads (ThermoFisher), which is the newer and superior method for immunoprecipitation due to its specificity. Thus, the beads are less likely to be the cause of the inconsistency that I had with co-immunoprecipitation. Increasing the amount of DNMBP and VASP inside the cell through transfection did not increase the chances of co-immunoprecipitation of both. In HEK293 cells, the transfection efficiency was ~40% while the transfection efficiency in U-2 OS cells was ~60%. Since co-immunoprecipitation of DNMBP and VASP did occur twice using non-transfected cells, the amount of these proteins in the cell is likely not the issue. Interestingly, though I was not able to co-immunoprecipitate DNMBP and VASP or MENA, the interaction between VASP and MENA was usually seen when VASP was pulled-down using some of the attempted protocols, indicating that the co-immunoprecipitations were working. The interaction between DNMBP and VASP or MENA may just have a low affinity or transient to capture through immunoprecipitation.

There are some alternative methods that could be used to analyze the interaction between variant DNMBP and protein partners. For example, a different system of co-immunoprecipitation exists using magnetic MicroBeads (MACS Miltenyl Biotec). Compared to Dynabeads, MicroBeads are significantly smaller, which reduces the chances of the beads interfering with the DNMBP interactions; if the epitope used for immunoprecipitation is somewhat covered by an interacting protein, the MicroBeads are more likely able to "squeeze in" between the protein without disrupting the potentially weak interactions. The specificity of the MicroBeads are said to be much higher, which means washing procedures can be less stringent

allowing for the co-immunoprecipitation of fragile interactions. Due to the sensitivity of MicroBeads, incubation of MicroBeads with protein lysate is much shorter as well (30 min), while incubation periods with Dynabeads can be anywhere from three hours to overnight. The shorter incubation periods are beneficial as it is less likely that protein interactions will be disrupted in that shorter time frame. Preliminary experiments have been conducted with these beads but did not show co-precipitation of DNMBP and partners.

An alternative method to co-immunoprecipitation for the examination of protein interactions is Fluorescence Resonance Energy Transfer (FRET). In FRET, the proteins of interest would be tagged with distinct fluorescent markers referred to as donor and acceptor fluorescent molecules. The donor fluorescent molecule attached to one of the proteins (i.e. DNMBP) would be excited at with a specific wavelength that would not excite the other fluorescent molecule. Once excited, the donor would emit energy that is able to excite the acceptor fluorescent molecule attached to the interacting protein (i.e. VASP). However, in order for the acceptor to be excited by energy emitted from the donor, the two fluorophores must be within 10nm of distance to each other (*191*). This very short distance between proteins is usually only observed when the proteins are interacting. A benefit to this assay is that it can be performed using standard immunofluorescence protocols with coupled fluorescent antibodies and interactions are observed within the cell. A disadvantage is that the fluorophores need to be oriented close to each other – if they are found on opposite sides of a protein complex that is larger than 10nm, the system will not work. As well, if a variant leads to a conformation change resulting in a proximity greater than 10nm between fluorophores, it would indicate that the two proteins are no longer interacting, but in reality this may not be the case, the protein may still be bound even with the conformation change. FRET would be worth attempting due to its relatively

simple and inexpensive design, though it can sometimes require a reasonable amount of optimization.

Another alternative to examine protein-protein interactions is the Two-Hybrid system. The Two-Hybrid system was originally conducted in yeast cells, but has since been modified for mammalian cells (*192*). In the Two-Hybrid system, the protein of interest is fused with a specific DNA-binding domain that usually binds upstream of a reporter gene, while one of the interacting partners is fused with a transcriptional activation domain for that reporter gene. If the two proteins bind, the transcriptional activation domain will activate transcription of the reporter gene, which can be measured. An advantage of the Two-Hybrid system is that transient interactions can be captured and the amount of interaction can be easily quantified. However, the system is disadvantageous because it would require the production of fusion proteins for each variant and each protein partner being analyzed. The interactions would also be occurring in a less 'natural' state.

A newer method for the analysis of protein interactions makes use of a rare protein modification, biotinylation, to label interacting proteins with biotin (*193*, *194*). In this system called BioID, the protein of interest (bait) is fused to a mutant 35kDa biotin ligase (BirA) from *E. coli* cells (*195*). The mutant biotin ligase is able to attach biotin to nearby amine groups. Thus, interacting proteins of the bait protein will be labelled with biotin. Protein lysate can then be extracted and biotin-labelled proteins can be isolated with magnetic beads with high affinity for biotin due to their streptavidin coating (*195*). Biotin-labelled isolated proteins can then be separated and detected through Western blot or mass spectrometry. A drawback of this system is that it may require stable expression of fusion proteins (*193*), which can sometimes be challenging. However, there are many benefits over other systems such as co-

immunoprecipitation and Two-Hybrid. For instance, the interactions only need to occur in their

normal environment – they do not need to be maintained after lysis. This allows for more

stringent lysis of cells and not having to optimize conditions after lysis. Transient and weak

interactions, like those DNMBP may be undergoing, can still be captured with this system. Thus,

BioID may be the most suitable system to use to determine if DNMBP variants are still able to

maintain interactions.

4.2: Activation of CDC42 by DNMBP Variants

One of DNMBP's major functions is to activate the Rho GTPase CDC42, a regulator of

cellular processes such as cytoskeleton organization (*196*). There is some evidence associating

CDC42 deficiencies to NTDs (*169, 170*). For instance, conditional knockouts of *Cdc42* in mice

show abnormalities in filopodia formation during neurulation (*169*). Filopodia are necessary

structures for the fusion of the neural tube. Knockdown of DNMBP through RNAi leads to an

80% reduction in active CDC42 in cells (*122*). Thus, a deficiency in DNMBP would lead to a

major deficiency in active CDC42, resulting in downstream cellular abnormalities such as those

seen with filopodia formation.

Another family member in the Rho GTPase family, RhoA, has been implicated in NTDs

using human variants (*197*). The PCP associated gene *Dact1* is associated to the regulation of

RhoA (*198*). Loss of *Dact1* in mice leads to NTDs and a reduction in the activity of RhoA (*198,

199*). *Dact1* was sequenced in a cohort of 167 Chinese NTD fetuses to look for damaging

variants (*197*). Five missense mutations predicted to be damaging were discovered and were

further examined. The variants do not lead to protein interaction or cellular localization

abnormalities. However, of the five variants, one variant leads to a decrease in the active GTP-

bound form of RhoA while another variant leads to an increase in active RhoA, which could possibly participate in the development of NTDs.

Similar to the DACT1-RhoA analysis, I wanted to determine if DNMBP variants had an effect on the activation of CDC42. There were five variants of particular interest found in the RhoGEF and bar domains of DNMBP (variants 2-6) since those two domains are utilized in CDC42 activation. The function of each DNMBP variant as an activator of CDC42 was originally to be tested using the CDC42 G-LISA Activation Assay (Cytoskeleton, Inc.). The assay allows for the pull-down of active GTP-bound CDC42 from cell protein lysate. DNMBP was transfected into HEK293 or U-2 OS cells and the amount of active CDC42 was measured. Replicates were not conducted with this assay because I was first assessing if the kit was suitable to use for this experiment. The results were inconclusive as I did not see an increase in active CDC42 when wildtype DNMBP was transfected into cells relative to non-transfected cells, indicating that the transfected DNMBP was not having an effect on CDC42.

One possibility is that the starting amount of inactive CDC42 compared to the active form in the cell was low, meaning that an overexpression of DNMBP would not lead to a detectable increase in active CDC42 due to the unavailability of the inactive form. To increase the amount of inactive CDC42 in cells, it was suggested by the G-LISA CDC42 Assay protocol to serum starve cells. Serum starvation is a commonly used technique in cell-based assays to reduce the endogenous activation of certain proteins. However, it is not a well-defined technique and can require a substantial amount of optimization to reduce activation; for some proteins, the activation change in response to serum starvation is quick and transient (*200*). In my analysis, serum starvation led to an increase in active CDC42. Due to the small number of assays in the

kit, there was little room for optimization. Because of this, I decided to conduct the assay without serum starving cells to see if transfection of DNMBP would still lead to an increase.

Transfection of wildtype DNMBP led to an apparent decrease in the active form of CDC42 rather than an increase in both HEK293 cells and U-2 OS cells. It is possible that the transfection of plasmids into cells were having adverse effects on the activation of CDC42. Control non-transfected cells were treated with only transfection reagents; thus, large plasmids were not entering the cell. Still, this would not explain the decrease in activation of CDC42 since the overexpression of DNMBP should lead to a major increase. Transfection efficiency for HEK293 and U-2 OS cells was ~40% and ~60%, respectively. This of course is not the most ideal efficiency. However, the G-LISA kit is reported to be very sensitive, therefore, the change in CDC42 activation should have been measurable at these transfection levels. While the transfection efficiency of each variant was not measured for HEK293 and U-2 OS cells, it is expected that there would be little variation since the vectors are nearly identical except for a single nucleotide. As well, transfection efficiency between variants in Caco-2 cells only varied by a few percentage points. Thus, it is likely that there would be little variation between variants when transfecting into other cell lines.

Since the transfection efficiency before each assay was not checked, it is always possible that transfection may have not occurred for some of the assays. This is very unlikely though since I had transfected using the exact same protocol for other experiments and always had successful transfections when checked. Interestingly, in the last assay conducted a reduction in CDC42 activation was observed in U-2 OS cell transfected with variant 5 (R1024X) relative to cells transfected with wildtype DNMBP. Variant 5 was one of the most likely variants to lead to a reduction in active CDC42 since the BAR domain is absent. Because replicates were not

conducted for that assay, we cannot say for sure if it was a significant decrease. Since the kit was expended, I was not able to repeat this attempt. Based on these preliminary results, we decided it would be best to study CDC42 activation in other ways.

The CDC42 G-LISA activation assay was the most favourable assay since it is highly sensitive and is carried out on CDC42 activated within the cell. There are also less ideal assays that are conducted on purified protein such as the GEF assay. In this assay, DNMBP variants would be purified and incubated with purified CDC42 to assess the ability of each variant to activate CDC42. Purification of proteins is also sometimes complex because the exact condition needed to maintain conformation may not be known. The BAR domain would also cause problems because it has been reported to be insoluble, which has led to some researchers to create fusion proteins lacking the BAR domain of DNMBP for purification purposes (*96*). Since there are variants with mutations in the BAR domain, we might not be able to study them with this assay. Salazar et al. (*96*) also showed the GEF assay is functional with only the RhoGEF domain of DNMBP. In the cell, other domains of DNMBP can lead to changes in the regulation of CDC42 due to interaction with proteins such as GM130 (*122*). A variant with a mutation might reduce CDC42 activation in a cell due to disruptions in interaction with other proteins, but this would not be captured by the GEF assay, making it less robust compared to other assays.

There are also some indirect approaches that can be used to measure CDC42 activation by DNMBP within cells through examination of cellular changes in response to CDC42 activation. DNMBP has been associated with various lipid vesicles found within certain cell types due to the its interaction with dynamin and some actin regulatory proteins (*96, 97, 142*). Sato et al. (*143*) have shown that overexpression of DNMBP leads to an increase of dense-core vesicles through activation of CDC42. Dense-core vesicles are found only in neural cell types

and carry peptide hormones (*143*, *144*). Expression of constitutively active CDC42 leads to increased exocytosis of dense-core vesicles and an increased number of these vesicles in the cytoplasm of PC12 cells (rat adrenal medulla pheochromocytoma) (*143*). As expected, overexpression of DNMBP through transfection leads to similar results due to the increased activation of CDC42. To determine if DNMBP variants were able to activate CDC42 at similar levels as wildtype DNMBP, I looked at the change in the number of dense-core vesicles in the neuroblastoma cell line SK-N-BE(2). I originally started the analysis in PC12 cells similar to Sato et al. (*143*), but I had trouble growing them flat on coverslips even after optimization, which prevented visualization of dense-core vesicles. Wildtype DNMBP and variant DNMBP plasmids were individually transfected into SK-N-BE(2) cells. A Chromogranin A (CgA) antibody was used to stain dense-core vesicles while successfully transfected cells were identified by staining with an HA-tag antibody.

Dense-core vesicles were not clearly distinguishable in every cell that was transfected, likely due to the resolution power of the microscope used. For this reason, instead of counting individual vesicles, mean fluorescence intensity of CgA staining for each cell was measured and averaged for each variant group. However, variation was very high in mean intensity measurements. A possible explanation for this could be variation in the amount of HA-DNMBP plasmid that each cell is up-taking. Varying levels of DNMBP protein could lead to variation in activation of CDC42 and therefore the number of dense-core vesicles. I normalized the CgA mean fluorescence intensities using mean fluorescence intensity measurements for HA-DNMBP, though it did not lead to a decrease in variation (data not shown). Thus, I was not able to show changes in CDC42 activation with the transfection of DNMBP using this assay. Since Sato et al. (*143*) only showed that transfecting DNMBP into rat PC12 cells leads to an increase in dense-

core vesicles, it is possible that it may not be the case for human neuroblastoma cell lines such as SK-N-BE(2). Furthermore, the increase measured by Sato et al. (*143*) was relatively small (33%). Measuring changes of this size sometimes requires a high level of optimization. Since this assay is relatively easy and fast to conduct, it would be worthwhile to investigate alternative cell lines that clearly show individual dense-core vesicles, which would make it easier to quantify an increase when DNMBP is transfected into cells rather than measuring the mean fluorescence intensity.

To further examine the possible role of CDC42 in neural tube defect development, I used a *Cdc42* mutation in conjunction with a *Cecr2* mutation to look for a genetic interaction. Using the Cre-Lox recombination system, a heterozygous *Cdc42* knockout mutation was produced on a background that was at least 50% FVB/N, which is dominant exencephaly-resistant. Since the human variants of DNMBP sequenced were nucleotide changes in only one allele, theoretically at most there would only be a 50% decrease in CDC42 activation by DNMBP if a variant had an effect on function. The *Dnmbp* variant found in the susceptible BALB/cCrl strain (P945L) is homozygous, making it possible that there is more than a 50% decrease in CDC42 activation with this variant. However, this is unlikely since a complete loss of DNMBP function would likely cause widespread abnormalities in the embryo. Thus, a heterozygous *Cdc42* knockout, which leads to a 50% reduction in CDC42 would likely represent the most extreme phenotype that a DNMBP variant could lead to. The hypomorphic *Cecr2*^GTbic45/GTbic45^ or the more severe *Cecr2*^GTbic45/tm1.1Hemc^ mutation was combined with a heterozygous *Cdc42* deletion. In theory, the *Cecr2* mutations alone should not lead to a high penetrance of exencephaly on the FVB/N background since it has been shown that this background is resistant to NTDs in a dominant fashion (*47*). However, if *Dnmbp* is modifying *Cecr2* through a reduction of CDC42 activity, a

deficiency of *Cdc42* in the resistant mice with the *Cecr2* mutation should lead to an increase in the penetrance of exencephaly.

The results indicated that a combination of *Cdc42*$^{Del/+}$ with either *Cecr2*$^{GTbic45/GTbic45}$ or *Cecr2*$^{GTbic45/tm1.1Hemc}$ on the resistant FVB/N strain did not lead to exencephaly or any other abnormal phenotypes at a penetrance observable in the number of animals produced to date. In total, 24 *Cdc42*$^{Del/+}$; *Cecr2*$^{GTbic45/GTbic45}$ and 15 *Cdc42*$^{Del/+}$; *Cecr2*$^{GTbic45/tm1.1Hemc}$ embryos were examined. While a homozygous *Cecr2*^{GTbic45} mutation on the FVB/N line does not lead to exencephaly (*47*, *55*), a homozygous *Cecr2*$^{tm1.1Hemc}$ deletion mutation is more severe and leads to about 12% of embryos with exencephaly on the FVB/N strain (*55*, *76*). Exencephaly has not been observed in heterozygous *Cecr2*$^{tm1.1Hemc}$ deletion FVB/N mice (*92*). *Cecr2*$^{GTbic45/tm1.1Hemc}$ mutations on the FVB/N had previously not been studied, thus it was not known if this combination would lead to exencephaly, though it would be predicted to be around 6%. Here, one *Cecr2*$^{GTbic45/tm1.1Hemc}$ mutant (with wildtype *Cdc42*) had exencephaly, leading to a penetrance of 5.26%. The number of embryos examined is too small for a penetrance analysis, thus more embryos would need to be studied to fully conclude that *Cdc42* is not affecting the penetrance of exencephaly, however, a large change in penetrance is not evident.

If further embryo collection shows no more than the expected exencephaly, the simplest explanation for the lack of exencephaly in the presence of *Cdc42* and *Cecr2* deficiencies in FVB/N mice is that *Dnmbp* is not modifying *Cecr2* through *Cdc42* or is not a modifier of *Cecr2*. There are multiple ways that *Dnmbp* can be modifying *Cecr2* as discussed in section 1.13. Thus, another interacting partner of DNMBP may play a role in the modification of *Cecr2*. It is also possible that mice were able to compensate for the missing *Cdc42* in some way during neurulation. It would be best to conduct a Western blot analysis on embryos to determine if

CDC42 expression decreased relative to wildtype embryos during neurulation. However, even if CDC42 expression is decreased, there would need to be a significant decrease in the active form of CDC42 to determine if *Dnmbp* is a modifier of *Cecr2* through *Cdc42*. This is because disruptions in DNMBP would lead to a decrease in activation of CDC42 rather than lowering the expression of CDC42. It is possible that DNMBP may just activate a higher percentage of the pool of available CDC42 in cells of *Cdc42* deficient mice. For example, if in a normal embryo 30% of the CDC42 found in each cell needs to be activated for correct neurulation, a $Cdc42^{Del/+}$ mutant embryo with only half the amount of CDC42 relative to wildtype mice could compensate through activating 60% of its pool of CDC42. Consequently, both the mutant and wildtype embryos would have a similar amount of active CDC42 leading to normal neurulation. For this reason, it would be necessary to perform the G-LISA CDC42 Activation Assay on a heterozygous CDC42 mutant and a wildtype sibling pairs to determine if activation is significantly decreased during neurulation in mutants. If activation is not significantly reduced, this mouse model cannot be used to test the hypothesis that states *Dnmbp* modifies *Cecr2* through *Cdc42*.

4.3: Localization of DNMBP Variants in Caco-2 Epithelial Cells

　　Localization of a protein is important for normal function of that protein. If a protein does not locate to the correct cellular region, the protein may not be close to interacting partners or targets it must act on. For this reason, I determined if DNMBP variants maintained normal localization. Previously it has been reported that some human variants of mouse NTD susceptibility genes show localization abnormalities. For instance, three of the four SEC24B variants (discussed in section 4.1) led to VANGL2 localization abnormalities (*189*). Robinson et

al. (*201*) have also reported localization abnormalities with human variants of the PCP genes *CELSR1* and *SCRIB*. Mutations to either *Celsr1* or *Scrib* in mice lead to craniorachischisis (*27, 202*). *CELSR1* and *SCRIB* were sequenced in 36 human fetuses with craniorachischisis to look for coding sequence variants (*201*). Normally, both proteins localize to the plasma membrane. Four variants of *CELSR1* show significantly decreased membrane localization and increased cytoplasmic localization, which was also seen with the *Celsr1* mouse mutant protein that leads to craniorachischisis. One *SCRIB* variant shows significantly reduced localization to the membrane, also similarly seen with the *Scrib* mouse mutant gene that leads to craniorachischisis. The interactions of CELSR1 and SCRIB with protein partners were analyzed and were found not to be disrupted. Thus, localization abnormalities of these proteins likely plays a role in the development of craniorachischisis in mice and possibly in humans.

Like CELSR1 and SCRIB, DNMBP also localizes to the cell membrane, specifically to the apical surface membrane in epithelial cells (*117*). Through transfection of each DNMBP variant into Caco-2 epithelial cells followed by immunofluorescence, I was able to visualize localization of DNMBP. Expression of wildtype transfected DNMBP localized to the membrane and cytoplasm, indicating that the overexpressed transfected DNMBP is able to localize to the membrane. All variants showed normal localization patterns except for variant 5 (R1024X). Variant 5 contains a nonsense mutation, resulting in a truncated protein. The majority of the BAR domain is lost in this variant, which is used to bind to curved membranes (22). Furthermore, the C-terminal end of DNMBP is needed for localization because it interacts with the co-localizing partner ZO-1. Without the interaction between ZO-1 and DNMBP, DNMBP does not the localize to the membrane, though ZO-1 is able to retain membrane localization (*117*). There was no clear localization of variant 5 protein to the membrane. Instead, the protein

was spread throughout the cell. If only variant 5 DNMBP protein was expressed, rather than mixed with endogenous DNMBP, I would expect the apical membranes in epithelial cells to be much more curved and slackened as reported by others (*117*, *203*). Apical tension of neuroepithelium is important during neural tube folding (*159*), thus, the presence of variant 5 could contribute to neural tube formation abnormalities.

It was not possible to quantify the fluorescence intensity of variant DNMBP protein at the membrane due to the large amount of variation in fluorescence signal between cells transfected with the same variant. A cause of this variation could be that some cells are taking up more plasmid DNA during transfection, resulting in more variant protein being expressed in the cell. As well, the 3D geometry of cells and the localization of DNMBP only to the apical surface prevented visualization of the localization pattern of the entire membrane in some cells due to the confocal microscope only being able to focus on one plane. It is possible to take images of different planes of the cells and then overlay each image, however, this would not solve the issue with some cells up-taking more plasmid than others and may add the problem of photobleaching. Thus, measuring the intensity in many cells and averaging the measurements to quantify subtle difference would likely not have been effective. Only significant or complete loss of localization to the membrane would be measurable.

While this localization analysis indicated that DNMBP shows pattern of concentration at the apical membrane of epithelial cells and the localization pattern is not observed with variant 5, we cannot infer that localization of other variants is completely normal in other specialized cells. This is because DNMBP is known to localize to other parts of the cell such as the Golgi complex, endoplasmic reticulum exit sites, and synapses in brain cells (*96*, *121*). It is possible for a variant to show normal localization to the cell membrane in epithelial cells but abnormal

localization to another specific region in a different cell type. Thus, it would be beneficial to analyze localization of DNMBP variants to other regions of the cell.

Aggregates of transfected DNMBP protein were seen in the cytoplasm of 10-20% of the Caoc-2 cells transfected with wildtype and variant DNMBP vectors. However, aggregation of the protein was not seen in cells transfected with variant 5 or non-transfected cells. A similar pattern was seen when SK-N-BE(2) cells were transfected. The aggregates were much larger and occurred at a frequency close to 100% in SK-N-BE(2) cells. Aggregates did not form in variant 5 transfected and non-transfected SK-N-BE(2) cells.

It is possible that this aggregation is another localization pattern of DNMBP in which the truncated protein produced by variant 5 does not localize to. However, staining for endogenous DNMBP in non-transfected cells did not show aggregations of DNMBP in the cytoplasm. The aggregations were also not observed in every Caco-2 and HEK293 cell transfected with DNMBP or a variant, but were seen in almost every SK-N-BE-2 cell transfected. Otani et al. (*117*) observed similar aggregations of DNMBP in the cytoplasm along with a co-localizing partner ZO-1 in the cytoplasm when both proteins were transfected into Caco-2 cells. When *Cecr2* was transfected into Caco-2 cells, aggregates of the transfected protein are also observed in the cytoplasm of 10-16% of cells even though CECR2 is expressed predominately in the nucleus. Thus, the excess amount of protein due to transfection is likely the cause of the formation of aggregates. This does not explain why Otani et al. (*117*) did not see aggregations when a control protein, β-catenin, was transfected into cells. It also does not explain why aggregates were not observed in cells transfected with variant 5 since Western blot analysis showed expression of variant 5 is higher than other variants. It may be that a region of the C-terminus end of DNMBP

is necessary for aggregation. It would be interesting to produce a DNMBP protein missing a portion of the N-terminus to see if aggregates are formed when transfected.

Another possibility is that these aggregates are inclusion bodies of protein. Inclusion bodies usually form when proteins are misfolded, leading to hydrophobic domains that are usually within a protein structure being exposed to solvents (*204*). It is possible that overexpressing DNMBP through transfection overwhelms cellular machinery, preventing proper folding of transfected protein. Furthermore, the BAR domain of DNMBP is insoluble (*96*), which may lead to inclusion body formation. This would explain why variant 5 does not show aggregates since the protein was truncated before the BAR domain. Other studies have observed inclusion bodies after protein transfection (*204–207*), though the aggregates I observed look different from inclusion bodies shown in these other studies. Also, less than 20% of transfected Caco-2 and HEK293 cells formed these aggregates. Thus, if these are actually inclusion bodies forming due to cellular machinery being overwhelmed, I would expect to see the inclusions in a larger majority of the cells transfected.

Since these aggregates did not show up in very many cells, it is unlikely that they would be disruptive to co-immunoprecipitation analyses or the CDC42 activation assay. However, the aggregates were very frequent in the SK-N-BE(2) cells. If the aggregates reduce the amount of available transfected protein in the cell, it may be the cause of inconsistencies in the dense-core vesicle analysis. To determine the aggregates form do to transfected protein overload, the amount of plasmid DNA could be reduced. Co-transfecting plasmids containing protein-folding machinery with DNMBP plasmids could also reduce the formation of inclusions.

4.4: Overall Conclusions and Future Directions

This research focused on the putative modifier gene *Dnmbp* as a susceptibility gene for neural tube defects. Previous studies have suggested that *Dnmbp* is one of the genes responsible for modifying the *Cecr2* exencephaly penetrance differences observed in BALB/cCrl and FVB/N mouse embryos. One mouse variant of *Dnmbp* and eight human variants from cranial NTD human fetuses were studied in cultured cell lines. I attempted to examine any abnormalities in function and localization of each variant.

The protein-protein interaction analysis using co-immunoprecipitations was technically inconsistent, thus, it may be necessary to examine interactions of each variant with another method. It would be worth reinvestigating protein-protein interactions with a different experiment such as FRET, Two-Hybrid, or BioID to determine if any variants are unable to interact with DNMBP partners. Due to the truncation in variant 5, it most likely would not interact with many of the actin regulatory proteins such as MENA, VASP, and N-WASP that bind to the C-terminal end, and thus variant 5 could act as a control. It is possible that other partners of DNMBP, besides MENA and VASP, form more stable interactions with DNMBP. Therefore, it would be worth attempting co-immunoprecipitations with another known interacting partner of DNMBP.

The G-LISA CDC42 Activation Assay did not show changes in CDC42 activity in cells when excess DNMBP was expressed. Preliminary results were not promising; thus, two alternative approaches were taken to determine if *Cdc42* was associated to NTD susceptibility through *Dnmbp*. Firstly, since increases in CDC42 activation leads to an increase in dense core vesicles, DNMBP variants were transfected into SK-N-BE(2) to determine if an increase in vesicles could be measured. An increase in vesicles was quantified by measuring mean fluoresce

intensity of each cell since individual vesicles were not always visible. Mean fluorescence intensities were very variable, thus results from this assay were inconclusive. It may be beneficial to re-attempt the assay with a different neural cell line, one that has a higher transfection efficiency and shows distinguishable dense-core vesicles. There is also the possibility that the transfected wildtype DNMBP protein is not functioning as GEF. It would be necessary to perform the GEF assay if there is more evidence of this.

It would be beneficial to further investigate aggregations that are forming in cells transfected with DNMBP. In Caco-2 and HEK293 cells, the frequency of aggregations was relatively low (less than 20%), thus, they likely do not effect assays being conducted. However, in SK-N-BE(2) cells, aggregates were seen in almost every transfected cell. If the aggregates are inclusion bodies, they could reduce the amount of available transfected DNMBP in the cell, thereby effecting assay results.

Secondly, a mouse model with a 50% *Cdc42* deficiency and homozygous *Cecr2* mutation was produced on the resistant FVB/N background to determine if there was an increase in exencephaly penetrance. To date, exencephaly has not observed in embryos with a combination of *Cdc42* and *Cecr2* deficiencies, although the experiment is ongoing

Since mice with a 50% *Cdc42* deficiency and homozygous *Cecr2* mutation did not show exencephaly, *Dnmbp* is likely not modifying *Cecr2* through *Cdc42*. One also needs to confirm that CDC42 activation is decreased in *Cdc42* heterozygous deletion mutants relative to wildtype embryos. This is necessary because as mentioned, even though there is less CDC42 being expressed, the embryo may be able to compensate by activating a higher percentage of CDC42, in which case there would be no abnormalities downstream of CDC42. This could be done using the G-LISA CDC42 Activation Assay with neurulating embryos. The assay may be more

accurate for this experiment since embryos are not being manipulated, unlike the attempts made

using this assay with cell lines. If a lower activation of CDC42 is measured in heterozygous

Cdc42 deletion embryos relative to wildtype embryos, it would support the idea that *Cdc42* is not

involved in the modification of *Cecr2*. Thus, it may be more beneficial to focus on studying the

effects DNMBP variants have on functions other than CDC42 activation, namely in actin

regulation during neurulation. This assay can also be conducted on FVB/N and BALB/cCrl

wildtype and *Cecr2* mutant embryos to determine if there is any change in CDC42 activation

between the strains.

If CDC42 activation is not lowered in heterozygous *Cdc42* deletion embryos, the mouse

model does not tell us if *Cdc42* is involved in modification of *Cecr2*. In this case, it would be

ideal to go back and optimize the G-LISA CDC42 Activation Assay to test the function of each

DNMBP variant in cells since it provides a direct measurement of CDC42 activation by DNMBP

variants and is more advantageous than other less direct experiments such as the GEF assay or

dense-core vesicle assay. Optimizing serum starvation of cells to determine when CDC42

activation is at the lowest level would be ideal. I would also recommend optimizing transfection

timing to find the moment when CDC42 activation is highest after transfection. This would

allow for the measurement of significant changes in CDC42 activation.

A more robust mouse model to determine if *Cdc42* is involved in the susceptibility to

NTDs with *Cecr2* would be one that contains a heterozygous mutant that lacks the RhoGEF

domain of DNMBP in conjunction with a homozygous *Cecr2* mutation on the FVB/N resistant

background. This would lead to a 50% reduction in the DNMBP's function as a CDC42 activator

that may be occurring due to deleterious variants with mutations in the RhoGEF domain. Thus, it

should only theoretically effect activation of CDC42.

Localization analyses in Caco-2 cells indicated that variant 5 (R1024X) was not localizing to the cell membrane of epithelial cells as expected, which may be necessary for the bending of the neural tube. Thus, in this study I have demonstrated that one human variant of *DNMBP* has abnormal characteristics that could be involved in the development of exencephaly and that *Dnmbp* may not be modifying *Cecr2* through *Cdc42*.

An alternative to using cell-based assays for the analysis of *Dnmbp* variants is performing *in vivo* experiments in zebrafish. With zebrafish, the homologous gene of interest can be knocked down with antisense morpholino oligonucleotides (*208*). Abnormal phenotypes are then examined, followed by rescue experiments using the human wildtype homologue of the gene. If the human wildtype gene is able to rescue some or all of the abnormal phenotypes, variants of the human gene can then be expressed to determine if they too are able to rescue the abnormal phenotypes. If a variant is not able to rescue abnormal phenotypes at a level similar to the wildtype human gene, the variant would be considered to be functioning abnormally. This methodology has been used previously to analyze human variants from NTD patients. For instance, knockdown of the gene *tri* in zebrafish leads to abnormal convergent extension, which is rescued by the human *VANGL1* gene (*209*). Two human variants of VANGL1 were not able to rescue the abnormal convergent extension phenotype during development in zebrafish – a process necessary for vertebrate neurulation. Similarly, *DNMBP* variants can be examined in zebrafish. Knockdown of *dnmbp* in zebrafish embryos has been shown to result in downward-curled tails, ciliary defects, small eyes, pericardial edema, hydrocephalus, and abdominal fluid accumulation (*137*). Wildtype mouse *Dnmbp* was able to rescue these phenotypes. Thus, since we know knockdown of *Dnmbp* in zebrafish leads to abnormal phenotypes, zebrafish would be an appropriate alternative for the analysis of human and mouse DNMBP variants.

Once functional analyses of individual variants in cell lines have shown that a few DNMBP variants lead to abnormal phenotypes, the next step would be to analyze DNMBP deficiencies in mouse models. Since a DNMBP knockout mouse line does not currently exist, one would need to be produced. It is possible that knockout of *Dnmbp* can itself lead to NTDs or be lethal before neurulation. In this case, it would be necessary to perform conditional knockouts under the control of tissue-specific drivers. The best route of action to take would be introduce a *Dnmbp* mutation on an FVB/N background and then cross on a *Cecr2* mutation to create *Cecr2* mutant homozygotes. If *Dnmbp* truly is a modifier of *Cecr2*, an increase in exencephaly penetrance should be observed. From there, variants of *Dnmbp* that were shown to be deleterious through functional and/or localization analyses, such as variant 5, could be introduced into a similar mouse model. This would help in the further characterization of human *Dnmbp* variants that are thought lead to the susceptibility of NTDs. A *Dnmbp* variant showing abnormal localization or function in cells is not enough evidence to prove that it is involved in NTDs in humans or mice. However, having a mouse model with irregularities in *Dnmbp* leading to an NTD in conjunction with *Cecr2* mutations would provide much more support for human and mice variants of *DNMBP* being susceptibility factors. It is possible that *Dnmbp* is not a modifier of *Cecr2* in mice, in which case we would not see an increase in exencephaly when *Dnmbp* is knocked out in addition to a *Cecr2* mutation on the FVB/N background.

In conclusion, the research presented here attempted to validate the gene *Dnmbp* as an NTD susceptibility gene in humans and modifier of *Cecr2* in mice. At least one human variant found in a cranial NTD proband localizes abnormally, providing some evidence for *DNMBP's* association to human NTDs. Mouse models support that *Dnmbp* is likely not modifying *Cecr2* through *Cdc42* activation. Further functional analysis of *DNMBP* variants as well as variants of

the other top candidate modifier genes of *Cecr2* would be beneficial as they would not only help

elucidate the complexity of neurulation, but they might also one day be used in screens for neural

tube defect susceptibility.

Literature Cited

1. N. D. E. Greene, A. J. Copp, Neural Tube Defects. *Annu. Rev. Neurosci.* **37**, 221–242 (2014).

2. E. R. Detrait *et al.*, Human neural tube defects: developmental biology, epidemiology, and genetics. *Neurotoxicol. Teratol.* **27**, 515–24 (2005).

3. S. I. Wilson, T. Edlund, Neural induction: toward a unifying mechanism. *Nat. Neurosci. Suppl.* **4**, 1161–1168 (2001).

4. V. Tropepe *et al.*, Direct Neural Fate Specification from Embryonic Stem Cells. *Neuron*. **30**, 65–78 (2001).

5. Hemmati-Brivanlou, D. Melton, A. Hemmati-brivanlou, D. Melton, Vertebrate embryonic cells will become nerve cells unless told otherwise. *Cell*. **88**, 13–17 (1997).

6. P. A. Wilson, A. Hemmati-Brivanlou, Induction of epidermis and inhibition of neural fate by Bmp-4. *Nature*. **376** (1995), pp. 331–333.

7. H.-K. Lee, H.-S. Le, S. A. Moody, Neural transcription factors: from embryos to neural stem cells. *Mol. Cells*. **37**, 705–12 (2014).

8. A. K. Groves, C. Labonne, Setting appropriate boundaries: Fate, patterning and competence at the neural plate border. *Dev. Biol.* **389**, 39–49 (2014).

9. Y. Sasai, Identifying the Missing Links: Genes that Connect Neural Induction and Primary Neurogenesis in Vertebrate Embryos. *Neuron*. **21**, 455–458 (1998).

10. C. D. Rogers, S. A. Moody, E. S. Casey, Neural induction and factors that stabilize a neural fate. *Birth Defects Res. Part C Embryo Today*. **87**, 249–262 (2009).

11. G. Sheng, M. Dos Reis, C. D. Stern, Churchill, a Zinc Finger Transcriptional Activator, Regulates the Transition between Gastrulation and Neurulation. *Cell*. **115**, 603–613 (2003).

12. K. R. Nitta, K. Tanegashima, S. Takahashi, M. Asashima, XSIP1 is essential for early neural gene expression and neural differentiation by suppression of BMP signaling. *Dev. Biol.* **275**, 258–267 (2004).

13. A. J. Copp, N. D. E. Greene, J. N. Murdoch, The genetic basis of mammalian neurulation. *Nat. Rev. Genet.* **4**, 784–793 (2003).

14. J. J. Wilde, J. R. Petersen, L. Niswander, Genetic, Epigenetic, and Environmental Contributions to Neural Tube Closure. *Annu. Rev. Genet.* **48**, 583–611 (2014).

15.   P. Ybot-Gonzalez, P. Cogram, D. Gerrelli, A. J. Copp, Sonic hedgehog and the molecular regulation of mouse neural tube closure. *Development*. **129**, 2507–2517 (2002).

16.   A. J. Copp, N. D. E. Greene, Neural tube defects - disorders of neurulation and related embryonic processes. *WIREs Dev. Biol.* **2**, 213–227 (2013).

17.   N. D. E. Greene, A. J. Copp, Development of the vertebrate central nervous system: formation of the neural tube. *Prenat. Diagn.* **29**, 303–311 (2009).

18.   P. Ybot-Gonzalez *et al.*, Neural plate morphogenesis during mouse neurulation is regulated by antagonism of Bmp signalling. *Development*. **134**, 3203–3211 (2007).

19.   C. Pyrgaki, P. Trainor, A. K. Hadjantonakis, L. Niswander, Dynamic imaging of mammalian neural tube closure. *Dev. Biol.* **344**, 941–947 (2010).

20.   D. M. Juriloff, M. J. Harris, C. Tom, K. B. Macdonald, Normal mouse strains differ in the site of initiation of closure of the cranial neural tube. *Teratology*. **44**, 225–233 (1991).

21.   L. E. Mitchell, Epidemiology of neural tube defects. *Am. J. Med. Genet. Part C*. **135C**, 88–94 (2005).

22.   A. J. Copp, N. D. E. Greene, Genetics and development of neural tube defects. *J. Pathol.* **220**, 217–230 (2010).

23.   C. A. Moore, in *Neural Tube Defects: From Origin to Treatment*, D. Wyszynski, Ed. (Oxford University Press, Inc., Oxford, 2006), pp. 66–75.

24.   L. V. Goodrich, L. Milenković, K. M. Higgins, M. P. Scott, Altered neural cell fates and medulloblastoma in mouse patched mutants. *Science*. **277**, 1109–1113 (1997).

25.   J. N. Murdoch, A. Copp, The relationship between Sonic hedgehog signalling , cilia and neural tube defects. *Birth Defects Res. (Part A)*. **88**, 633–652 (2010).

26.   D. M. Juriloff, M. J. Harris, A consideration of the evidence that genetic defects in planar cell polarity contribute to the etiology of human neural tube defects. *Birth Defects Res. (Part A)*. **94**, 824–840 (2012).

27.   J. N. Murdoch *et al.*, Circletail, a new mouse mutant with severe neural tube defects: chromosomal localization and interaction with the loop-tail mutation. *Genomics*. **78**, 55–63 (2001).

28.   X. Lu *et al.*, PTK7/CCK-4 is a novel regulator of planar cell polarity in vertebrates. *Nature*. **430**, 93–98 (2004).

29.   MRC Vitamin Study Research Group, "Prevention of neural tube defects: results of the Medical Research Council Vitamin Study" (1991).

30.   M. E. Youngblood *et al.*, 2012 Update on global prevention of folic acid-preventable spina bifida and anencephaly. *Birth Defects Res. (Part A).* **97**, 658–663 (2013).

31.   S. Zabihi, M. R. Loeken, Understanding diabetic teratogenesis: where are we now and where are we going? *Birth Defects Res. (Part A).* **88**, 779–790 (2010).

32.   K. J. Stothard, P. W. Tennant, R. Bell, J. Rankin, Maternal Overweight and Obesity and the Risk of Congenital Anomalies. *Surv. Anesthesiol.* **53**, 218–219 (2009).

33.   C. D. Chambers, Risks of hyperthermia associated with hot tub or spa use by pregnant women. *Birth Defects Res. (Part A).* **76**, 569–573 (2006).

34.   J. W. Dreier, A.-M. N. Andersen, G. Berg-Beckhoff, Systematic review and meta-analyses: fever in pregnancy and health impacts in the offspring. *Pediatrics.* **133**, e674-688 (2014).

35.   K. S. Au, A. Ashley-Koch, H. Northrup, Epidemiologic and genetic aspects of spina bifida and other neural tube defects. *Dev. Disabil. Res. Rev.* **16**, 6–15 (2010).

36.   W. Satoh, M. Matsuyama, H. Takemura, S. Aizawa, A. Shimono, Sfrp1, Sfrp2, and Sfrp5 regulate the Wnt/Beta-catenin and the planar cell polarity pathways during early trunk formation in mouse. *Genesis.* **46**, 92–103 (2008).

37.   A. E. Beaudin *et al.*, Shmt1 and de novo thymidylate biosynthesis underlie folate-responsive neural tube defects in mice. *Am. J. Clin. Nutr.* **93**, 789–798 (2011).

38.   I. E. Zohn, Mouse as a model for multifactorial inheritance of neural tube defects. *Birth Defects Res. Part C - Embryo Today.* **96**, 193–205 (2012).

39.   M. Kousi, N. Katsanis, Genetic modifiers and oligogenic inheritance. *Cold Spring Harb. Perspect. Med.* **5**, 1–22 (2015).

40.   B. A. Hamilton, B. D. Yu, Modifier genes and the plasticity of genetic networks in mice. *PLoS Genet.* **8**, e1002644 (2012).

41.   J. H. Nadeau, Modifier genes in mice and humans. *Nat. Rev. Genet. Genet.* **2**, 165–174 (2001).

42.   K. Noben-Trauth, Q. Y. Zheng, K. R. Johnson, P. M. Nishina, Mdfw: a Deafness Susceptibility Locus That Interacts With Deaf Waddler (Dfw). *Genomics.* **44**, 266–72 (1997).

43. V. A. Street, J. W. McKee-Johnson, R. C. Fonseca, B. L. Tempel, K. Noben-Trauth, Mutations in a plasma membrane Ca2+-ATPase gene cause deafness in deafwaddler mice. *Nat. Genet.* **19**, 390–394 (1998).

44. K. Noben-Trauth, Q. Y. Zheng, K. R. Johnson, Association of cadherin 23 with polygenic inheritance and genetic modification of sensorineural hearing loss. *Nat. Genet.* **4**, 21–23 (2011).

45. C. Colmenares *et al.*, Loss of the SKI proto-oncogene in individuals affected with 1p36 deletion syndrome is predicted by strain-dependent defects in Ski-/- mice. *Nat. Genet.* **30**, 106–109 (2002).

46. E. Andersson *et al.*, Genetic interaction between Lrp6 and Wnt5a during mouse development. *Dev. Dyn.* **239**, 237–245 (2010).

47. G. S. Banting *et al.*, CECR2, a protein involved in neurulation, forms a novel chromatin remodeling complex with SNF2L. *Hum. Mol. Genet.* **14**, 513–524 (2005).

48. J. Böhm *et al.*, Sall1, sall2, and sall4 are required for neural tube closure in mice. *Am. J. Pathol.* **173**, 1455–1463 (2008).

49. V. Bryja *et al.*, The Extracellular Domain of Lrp5/6 Inhibits Noncanonical Wnt Signaling In Vivo. *Seikagaku*. **20**, 924–936 (2009).

50. M. C. Lemos *et al.*, Genetic background influences embryonic lethality and the occurrence of neural tube defects in Men1 null mice: Relevance to genetic modifiers. *J. Endocrinol.* **203**, 133–142 (2009).

51. J. Mao *et al.*, The iron exporter ferroportin 1 is essential for development of the mouse embryo, forebrain patterning and neural tube closure. *Development*. **137**, 3079–3088 (2010).

52. J. A. McMahon *et al.*, and patterning of the neural tube and somite Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. *GENES Dev.* **12**, 1438–1452 (1998).

53. T. Takeuchi, M. Kojima, K. Nakajima, S. Kondo, jumonji gene is essential for the neurulation and cardiac development of mouse embryos with a C3H/He background. *Mech. Dev.* **86**, 29–38 (1999).

54. I. E. Zohn *et al.*, The flatiron mutation in mouse ferroportin acts as a dominant negative to cause ferroportin disease. *Blood*. **109**, 4174–4180 (2007).

55. R. Y. M. Leduc, P. Singh, H. E. Mcdermid, Genetic Backgrounds and Modifier Genes of NTD Mouse Models: An Opportunity for Greater Understanding of the Multifactorial Etiology of Neural Tube Defects. *Birth Defects Res. (Part A)*, 1–14 (2016).

56. S. C. P. De Castro *et al.*, Lamin B1 Polymorphism Influences Morphology of the Nuclear Envelope, Cell Cycle Progression, and Risk of Neural Tube Defects in Mice. *PLoS Genet.* **8**, e1003059 (2012).

57. M. M. Lakkis, J. A. Golden, K. S. O'Shea, J. A. Epstein, Neurofibromin deficiency in mice causes exencephaly and is a modifier for Splotch neural tube defects. *Dev. Biol.* **212**, 80–92 (1999).

58. J. Dixon, M. J. Dixon, Genetic Background Has a Major Effect on the Penetrance and Severity of Craniofacial Defects in Mice Heterozygous for the Gene Encoding the Nucleolar Protein Treacle. *Dev. Dyn.* **229**, 907–914 (2004).

59. C. Kuang *et al.*, Intragenic deletion of Tgif causes defects in brain development. *Hum. Mol. Genet.* **15**, 3508–3519 (2006).

60. A. Fleming, A. J. Copp, A genetic risk factor for mouse neural tube defects: defining the embryonic basis. *Hum. Mol. Genet.* **9**, 575–581 (2000).

61. D. D. Anderson, P. J. Stover, SHMT1 and SHMT2 are functionally redundant in nuclear de novo thymidylate biosynthesis. *PLoS One.* **4**, e5839 (2009).

62. A. J. MacFarlane *et al.*, Cytoplasmic serine hydroxymethyltransferase regulates the metabolic partitioning of methylenetetrahydrofolate but is not essential in mice. *J. Biol. Chem.* **283**, 25846–25853 (2008).

63. P. G. Matteson *et al.*, The orphan G protein-coupled receptor, Gpr161, encodes the vacuolated lens locus and controls neurulation and lens development. *Proc. Natl. Acad. Sci.* **105**, 2088–2093 (2008).

64. R. Korstanje *et al.*, Quantitative trait loci affecting phenotypic variation in the vacuolated lens mouse mutant, a multigenic mouse model of neural tube defects. *Physiol. Genomics.* **35**, 296–304 (2008).

65. S. B. Ting *et al.*, Inositol- and folate-resistant neural tube defects in mice lacking the epithelial-specific factor Grhl-3. *Nat. Med.* **9**, 1513–1519 (2003).

66. P. E. Neumann *et al.*, Multifactorial inheritance of neural tube defects: localization of the major gene and recognition of modifiers in ct mutant mice. *Nature.* **6**, 362–369 (1994).

67.   Q. Zhao, R. R. Behringer, B. de Crombrugghe, Prenatal folic acid treatment suppresses acrania and meroanencephaly in mice mutant for the Cart-1 homeobox gene. *Nat. Genet.* **14**, 353–356 (1996).

68.   W. Lin *et al.*, Proper expression of the Gcn5 histone acetyltransferase is required for neural tube closure in mouse embryos. *Dev. Dyn.* **237**, 928–940 (2008).

69.   B. I. Li *et al.*, The orphan GPCR, Gpr161, regulates the retinoic acid and canonical Wnt pathways during neurulation. *Dev. Biol.* **402**, 17–31 (2015).

70.   K. Hahm *et al.*, Defective neural tube closure and anteroposterior patterning in mice lacking the LIM protein LMO4 or its interacting partner Deaf-1. *Mol. Cell. Biol.* **24**, 2074–82 (2004).

71.   T. P. Yao *et al.*, Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell.* **93**, 361–372 (1998).

72.   V. P. Sah *et al.*, A subset of p53-deficient embryos exhibit exencephaly. *Nat. Genet.* **10**, 175–180 (1995).

73.   Z. A. Abdelhamed *et al.*, Variable expressivity of ciliopathy neurological phenotypes that encompass Meckel - Gruber syndrome and Joubert syndrome is caused by complex de-regulated ciliogenesis, Shh and wnt signalling defects. *Hum. Mol. Genet.* **22**, 1358–1372 (2013).

74.   K. Niederreither *et al.*, Genetic evidence that oxidative derivatives of retinoic acid are not involved in retinoid signaling during mouse development. *Nat. Genet.* **31**, 84–88 (2002).

75.   P. J. Thompson, K. A. Norton, F. H. Niri, C. E. Dawe, H. E. McDermid, CECR2 is involved in spermatogenesis and forms a complex with SNF2H in the testis. *J. Mol. Biol.* **415**, 793–806 (2012).

76.   N. A. Fairbridge *et al.*, Cecr2 mutations causing exencephaly trigger misregulation of mesenchymal/ectodermal transcription factors. *Birth Defects Res. (Part A)*. **88**, 619–25 (2010).

77.   S. K. Lee, E. J. Park, H. S. Lee, Y. S. Lee, J. Kwon, Genome-wide screen of human bromodomain-containing proteins identifies Cecr2 as a novel DNA damage response protein. *Mol. Cells*. **34**, 85–91 (2012).

78.   O. Fedorov *et al.*, [1,2,4]Triazolo[4,3-a]phthalazines: Inhibitors of diverse bromodomains. *J. Med. Chem.* **57**, 462–476 (2014).

79. R. Sanchez, M.-M. Zhou, The role of human bromodomains in chromatin biology and gene transcription. *Curr. Opin. Drug Discov. Devel.* **12**, 659–665 (2009).

80. P. Filippakopoulos, S. Knapp, The bromodomain interaction module. *FEBS Lett.* **586**, 2692–2704 (2012).

81. T. Doerks, R. Copley, P. Bork, DDT - A novel domain in different transcription and chromosome remodeling factors. *Trends Biochem. Sci.* **26**, 145–146 (2001).

82. L. Aravind, D. Landsman, AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res.* **26**, 4413–4421 (1998).

83. J. Dong *et al.*, SLIDE, The protein interacting domain of imitation switch remodelers, binds DDT-domain proteins of different subfamilies in chromatin remodeling complexes. *J. Integr. Plant Biol.* **55**, 928–937 (2013).

84. C. E. Dawe, M. K. Kooistra, N. A. Fairbridge, A. C. Pisio, H. E. McDermid, Role of chromatin remodeling gene Cecr2 in neurulation and inner ear development. *Dev. Dyn.* **240**, 372–383 (2011).

85. J. Chen, G. Morosan-Puopolo, F. Dai, J. Wang, B. Brand-Saberi, Molecular cloning of chicken Cecr2 and its expression during chicken embryo development. *Int. J. Dev. Biol.* **54**, 925–929 (2010).

86. K. S. K. Fong *et al.*, A mutation disrupting TET1 activity alters the expression of genes critical for neural tube closure in the tuft mouse. *Dis. Model. Mech.* **9**, 585–596 (2016).

87. M. M. Dawlaty *et al.*, Combined Deficiency of Tet1 and Tet2 Causes Epigenetic Abnormalities but Is Compatible with Postnatal Development. *Dev. Cell.* **24**, 310–323 (2013).

88. K. S. K. Fong *et al.*, Craniofacial features resembling frontonasal dysplasia with a tubulonodular interhemispheric lipoma in the adult 3H1 tuft mouse. *Birth Defects Res. (Part A).* **94**, 102–113 (2012).

89. C. Pyrgaki, A. Liu, L. Niswander, Grainyhead-like 2 regulates neural tube closure and adhesion molecule expression during neural fold fusion. *Dev. Biol.* **353**, 38–49 (2011).

90. M. K. Kooistra *et al.*, Strain-specific modifier genes of Cecr2-associated exencephaly in mice: genetic analysis and identification of differentially expressed candidate genes. *Physiol. Genomics.* **44**, 35–46 (2012).

91.   C. E. Davidson, Q. Li, G. A. Churchill, L. R. Osborne, H. E. McDermid, Modifier locus for exencephaly in Cecr2 mutant mice is syntenic to the 10q25.3 region associated with neural tube defects in humans. *Physiol. Genomics*. **31**, 244–51 (2007).

92.   R. Y. M. Leduc, Role of Cecr2 and Candidate Modifier Genes of Cecr2 in the Neural Tube Defect Exencephaly. *Thesis* (2015).

93.   P. Kumar, S. Henikoff, P. C. Ng, Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009).

94.   E. V. Davydov *et al.*, Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput. Biol.* **6** (2010).

95.   I. A. Adzhubei *et al.*, A method and server for predicting damaging missense mutations. *Nat. Methods*. **7**, 248–249 (2010).

96.   M. A. Salazar *et al.*, Tuba, a novel protein containing bin/amphiphysin/Rvs and Dbl homology domains, links dynamin to regulation of the actin cytoskeleton. *J. Biol. Chem.* **278**, 49031–49043 (2003).

97.   G. Cestra, A. Kwiatkowski, M. Salazar, F. Gertler, P. De Camilli, Tuba, A GEF for CDC42, links dynamin to actin regulatory proteins. *Methods Enzymol.* **404**, 537–545 (2006).

98.   G. C. Ochoa *et al.*, A functional link between dynamin and the actin cytoskeleton at podosomes. *J. Cell Biol.* **150**, 377–389 (2000).

99.   A. Roux, Reaching a consensus on the mechanism of dynamin? *F1000Prime Rep.* **6** (2014).

100.   J. Cherfils, P. Chardin, GEFs: Structural basis for their activation of small GTP-binding proteins. *Trends Biochem. Sci.* **24**, 306–311 (1999).

101.   S. J. Heasman, A. J. Ridley, Mammalian Rho GTPases: new insights into their functions from in vivo studies. *Nat. Rev. Mol. Cell Biol.* **9**, 690–701 (2008).

102.   A. Eberth, M. R. Ahmadian, In vitro GEF and GAP assays. *Curr. Protoc. Cell Biol.* **14.9** (2009).

103.   K. L. Rossman, C. J. Der, J. Sondek, GEF means Go: turning on Rho GTPases with guanine nucleotide-exchange factors. *Nat. Rev. Mol. Cell Biol.* **6**, 167–180 (2005).

104.   B. Habermann, The BAR-domain family of proteins: a case of bending and binding? *EMBO Rep.* **5**, 250–255 (2004).

105. L. Lu, H. Horstmann, C. Ng, W. Hong, Regulation of Golgi structure and function by ARF-like protein 1 (Arl1). *J. Cell Sci.* **114**, 4543–55 (2001).

106. O. H. Shin, J. H. Exton, Differential binding of arfaptin 2/POR1 to ADP-ribosylation factors and Rac1. *Biochem. Biophys. Res. Commun.* **285**, 1267–1273 (2001).

107. M. J. Bottomley, P. L. Surdo, P. C. Driscoll, Endocytosis: How dynamin sets vesicles PHree! *Curr. Biol.* **9**, R301–R304 (1999).

108. S. Tortorella, T. C. Karagiannis, Transferrin receptor-mediated endocytosis: A useful target for cancer therapy. *J. Membr. Biol.* **247**, 291–307 (2014).

109. S. J. McClure, P. J. Robinson, Dynamin, endocytosis and intracellular signalling. *Mol. Membr. Biol.* **13**, 189–215 (1996).

110. P. Wigge, Y. Vallis, H. T. McMahon, Inhibition of receptor-mediated endocytosis by the amphiphysin SH3 domain. *Curr. Biol.* **7**, 554–60 (1997).

111. N. J. Nosworthy *et al.*, Actin binding proteins: regulation of cytoskeletal microfilaments. *Physiol. Rev.* **83**, 433–473 (2003).

112. S. R. Gross, Actin binding proteins: Their ups and downs in metastatic life. *Cell Adhes. Migr.* **7**, 199–213 (2013).

113. A. V. Kwiatkowski, F. B. Gertler, J. J. Loureiro, Function and regulation of Ena/VASP proteins. *Trends Cell Biol.* **13**, 386–392 (2003).

114. J. E. Bear, F. B. Gertler, Ena/VASP: towards resolving a pointed controversy at the barbed end. *J. Cell Sci.* **122**, 1947–1953 (2009).

115. J. E. Bear *et al.*, Negative regulation of fibroblast motility by Ena/VASP proteins. *Cell*. **101**, 717–728 (2000).

116. Y. Oda, T. Otani, J. Ikenouchi, M. Furuse, Tricellulin regulates junctional tension of epithelial cells at tricellular contacts via Cdc42. *J. Cell Sci.* **127**, 4201–4212 (2014).

117. T. Otani, T. Ichii, S. Aono, M. Takeichi, Cdc42 GEF Tuba regulates the junctional configuration of simple epithelial cells. *J. Cell Biol.* **175**, 135–146 (2006).

118. D. C. Y. Phua *et al.*, ZO-1 and ZO-2 are required for extra-embryonic endoderm integrity, primitive ectoderm survival and normal cavitation in embryoid bodies derived from mouse embryonic stem cells. *PLoS One*. **9**, e99532 (2014).

119. A. S. Fanning *et al.*, The Tight Junction Protein ZO-1 Establishes a Link between the Transmembrane Protein Occludin and the Actin Cytoskeleton. *J. Biol. Chem.* **273**, 29745–29753 (1998).

120. N. Nakamura, Emerging new roles of GM130, a cis-Golgi matrix protein, in higher order cell functions. *J. Pharmacol. Sci.* **112**, 255–264 (2010).

121. A. Kodani, C. Sutterlin, The Golgi Protein GM130 Regulates Centrosome Morphology and Function. *Mol. Biol. Cell*. **19**, 745–753 (2008).

122. A. Kodani, I. Kristensen, L. Huang, C. Su, GM130-dependent Control of Cdc42 Activity at the Golgi Regulates Centrosome Organization. *Mol. Biol. Cell*. **20**, 1192–1200 (2009).

123. S. Sinha, W. Yang, Cellular signaling for activation of Rho GTPase Cdc42. *Cell. Signal.* **20**, 1927–1934 (2008).

124. L. E. Arias-Romero, J. Chernoff, Targeting Cdc42 in cancer. *Expert Opin. Ther. Targets*. **17**, 1263–1273 (2013).

125. L. Yang, L. Wang, Y. Zheng, Gene Targeting of Cdc42 and Cdc42GAP Affirms the Critical Involvement of Cdc42 in Filopodia Induction, Directed Migration, and Proliferation in Primary Mouse Embryonic Fibroblasts. *Mol. Biol. Cell*. **17**, 4675–4685 (2006).

126. S. Cotteret, J. Chernoff, The evolutionary history of effectors downstream of Cdc42 and Rac. *Genome Biol.* **3** (2002).

127. S. Y. Choi *et al.*, Cdc42 Deficiency Causes Ciliary Abnormalities and Cystic Kidneys. *J. Am. Soc. Nephrol.* **24**, 1435–1450 (2013).

128. R. Nagahama *et al.*, Rho GTPase protein Cdc42 is critical for postnatal cartilage development. *Biochem. Biophys. Res. Commun.* **470**, 813–817 (2016).

129. T. K. Ito *et al.*, A crucial role for CDC42 in senescence-associated inflammation and atherosclerosis. *PLoS One*. **9**, e102186 (2014).

130. Y. Liu *et al.*, Inactivation of Cdc42 in neural crest cells causes craniofacial and cardiovascular morphogenesis defects. *Dev. Biol.* **383**, 239–252 (2013).

131. Y. Qin, W. H. Meisen, Y. Hao, I. G. Macara, Tuba, a Cdc42 GEF, is required for polarized spindle orientation during epithelial cyst formation. *J. Cell Biol.* **189**, 661–669 (2010).

132.	E. M. Kovacs, S. Verma, S. G. Thomas, A. S. Yap, Tuba and N-WASP function cooperatively to position the central lumen during epithelial cyst morphogenesis. *Cell Adh. Migr.* **5**, 344–350 (2011).

133.	A. B. Jaffe, N. Kaji, J. Durgan, A. Hall, Cdc42 controls spindle orientation to position the apical surface during epithelial morphogenesis. *J. Cell Biol.* **183**, 625–633 (2008).

134.	E. S. Fleming *et al.*, Planar spindle orientation and asymmetric cytokinesis in the mouse small intestine. *J. Histochem. Cytochem.* **55**, 1173–1180 (2007).

135.	Z. Hao *et al.*, Impaired maturation of large dense-core vesicles in muted-deficient adrenal chromaffin cells. *J. Cell Sci.* **128**, 1365–1374 (2015).

136.	S.-C. Tien *et al.*, The Shp2-induced epithelial disorganization defect is reversed by HDAC6 inhibition independent of Cdc42. *Nat. Commun.* **7**, 10420 (2016).

137.	J.-I. Baek *et al.*, Dynamin Binding Protein (Tuba) Deficiency Inhibits Ciliogenesis and Nephrogenesis *in Vitro* and *in Vivo*. *J. Biol. Chem.* **291**, 8632–8643 (2016).

138.	S. Fan *et al.*, Polarity Proteins Control Ciliogenesis via Kinesin Motor Interactions. *Curr. Biol.* **14**, 1451–1461 (2004).

139.	X. Zuo, B. Fogelgren, J. H. Lipschutz, The small GTPase Cdc42 is necessary for primary ciliogenesis in renal tubular epithelial cells. *J. Biol. Chem.* **286**, 22469–77 (2011).

140.	R. D. Mullins, How WASP-family proteins and the Arp2/3 complex convert intracellular signals into cytoskeletal structures. *Curr. Opin. Cell Biol.* **12**, 91–96 (2000).

141.	N. K. Hussain *et al.*, Endocytic protein intersectin-l regulates actin assembly via Cdc42 and N-WASP. *Nat. Cell Biol.* **3**, 927–932 (2001).

142.	E. M. Kovacs, R. S. Makar, F. B. Gertler, Tuba stimulates intracellular N-WASP-dependent actin assembly. *J. Cell Sci.* **119**, 2715–2726 (2006).

143.	M. Sato *et al.*, The small GTPase Cdc42 modulates the number of exocytosis-competent dense-core vesicles in PC12 cells. *Biochem. Biophys. Res. Commun.* **420**, 417–421 (2012).

144.	R. D'Alessandro, J. Meldolesi, Expression and function of the dense-core vesicle membranes are governed by the transcription repressor REST. *FEBS Lett.* **587**, 1915–1922 (2013).

145.	D. Giner *et al.*, Vesicle movements are governed by the size and dynamics of F-actin cytoskeletal structures in bovine chromaffin cells. *Neuroscience.* **146**, 659–669 (2007).

146. L. M. Lanier *et al.*, Mena Is Required for Neurulation and Commissure Formation. *Neuron*. **22**, 313–325 (1999).

147. A. S. Menzies *et al.*, Mena and vasodilator-stimulated phosphoprotein are required for multiple actin-dependent processes that shape the vertebrate nervous system. *J. Neurosci.* **24**, 8029–8038 (2004).

148. A. J. Koleske *et al.*, Essential roles for the Abl and Arg tyrosine kinases in neurulation. *Neuron*. **21**, 1259–1272 (1998).

149. M. R. Brouns *et al.*, The adhesion signaling molecule p190 RhoGAP is required for morphogenetic processes in neural development. *Development*. **127**, 4891–4903 (2000).

150. D. J. Stumpo, C. B. Bock, J. S. Tuttle, P. J. Blackshear, MARCKS deficiency in mice leads to abnormal brain development and perinatal death. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 944–948 (1995).

151. H. Luo *et al.*, Disruption of palladin results in neural tube closure defects in mice. *Mol. Cell. Neurosci.* **29**, 507–515 (2005).

152. J. D. Hildebrand, P. Soriano, Shroom, a PDZ Domain–Containing Actin-Binding Protein, Is Required for Neural Tube Morphogenesis in Mice. *Cell*. **99**, 485–497 (1999).

153. A. S. Rakeman, K. V. Anderson, Axis specification and morphogenesis in the mouse embryo require Nap1, a regulator of WAVE-mediated actin branching. *Development*. **133**, 3075–3083 (2006).

154. C. B. Gurniak, E. Perlas, W. Witke, The actin depolymerizing factor n-cofilin is essential for neural tube morphogenesis and neural crest cell migration. *Dev. Biol.* **278**, 231–241 (2005).

155. M. E. Durkin *et al.*, DLC-1, a Rho GTPase-activating protein with tumor suppressor function, is essential for embryonic development. *FEBS Lett.* **579**, 1191–1196 (2005).

156. W. Xu, H. Baribault, Vinculin knockout results in heart and brain defects during embryonic development. *Development*. **125**, 327–37 (1998).

157. G. Morriss-Kay, F. Tuckett, The role of microfilaments in cranial neurulation in rat embryos: effects of short-term exposure to cytochalasin D. *J. Embryol. Exp. Morphol.* **88**, 333–348 (1985).

158. P. Ybot-Gonzalez, A. J. Copp, Bending of the neural plate during mouse spinal neurulation is independent of actin microfilaments. *Dev. Dyn.* **215**, 273–283 (1999).

159. M. Suzuki, H. Morita, N. Ueno, Molecular mechanisms of cell shape changes that contribute to vertebrate neural tube closure. *Dev. Growth Differ.* **54**, 266–276 (2012).

160. S. L. Haigo, J. D. Hildebrand, R. M. Harland, J. B. Wallingford, Shroom Induces Apical Constriction and Is Required for Hingepoint Formation during Neural Tube Closure. *Curr. Biol.* **13**, 2125–2137 (2003).

161. M. Krause, E. W. Dent, J. E. Bear, J. J. Loureiro, F. B. Gertler, ENA/VASP Proteins: Regulators of the Actin Cytoskeleton and Cell Migration. *Annu. Rev. Cell Dev. Biol.* **19**, 541–564 (2003).

162. F. Drees, F. B. Gertler, Ena/VASP: proteins at the tip of the nervous system. *Curr. Opin. Neurobiol.* **18**, 53–59 (2008).

163. A. M. Chou, K. P. Sem, G. D. Wright, T. Sudhaharan, S. Ahmed, Dynamin1 is a novel target for IRSp53 protein and works with mammalian enabled (Mena) protein and Eps8 to regulate filopodial dynamics. *J. Biol. Chem.* **289**, 24383–24396 (2014).

164. S. Krugmann *et al.*, Cdc42 induces filopodia by promoting the formation of an IRSp53:Mena complex. *Curr. Biol.* **11**, 1645–1655 (2001).

165. B. L. Kim *et al.*, The Cdc42 effector IRSp53 generates filopodia by coupling membrane protrusion with actin dynamics. *J. Biol. Chem.* **283**, 20454–20472 (2008).

166. F. Castellano *et al.*, Inducible recruitment of Cdc42 or WASP to a cell-surface receptor triggers actin polymerization and filopodium formation. *Curr. Biol.* **9**, 351–360 (1999).

167. F. Chen *et al.*, Cdc42 is required for PIP(2)-induced actin polymerization and early development but not for cell viability. *Curr. Biol.* **10**, 758–765 (2000).

168. P. M. Duquette, N. Lamarche-Vane, Rho GTPases in embryonic development. *Small GTPases*. **5**, e972857-1–9 (2014).

169. A. Rolo *et al.*, Regulation of cell protrusions by small GTPases during fusion of the neural folds. *Elife*. **5**, 1–24 (2016).

170. E. K. Kieserman, J. B. Wallingford, In vivo imaging reveals a role for Cdc42 in spindle positioning and planar orientation of cell divisions during vertebrate neural tube closure. *J. Cell Sci.* **122**, 2481–2490 (2009).

171. A. Hall, Rho GTPases and the actin cytoskeleton. *Science (80-.).* **279**, 509–514 (1998).

172. S.-T. Sit, E. Manser, Rho GTPases and their role in organizing the actin cytoskeleton. *J. Cell Sci.* **124**, 679–683 (2011).

173. S. Etienne-Manneville, A. Hall, Integrin-mediated activation of Cdc42 controls cell polarity in migrating astrocytes through PKCζ. *Cell*. **106**, 489–498 (2001).

174. M. Fukata *et al.*, Rac1 and Cdc42 capture microtubules through IQGAP1 and CLIP-170. *Cell*. **109**, 873–885 (2002).

175. E. Aaku-Saraste, A. Hellwig, W. B. Huttner, Loss of occludin and functional tight junctions, but not ZO-1, during neural tube closure - remodeling of the neuroepithelium prior to neurogenesis. *Dev. Biol.* **180**, 664–679 (1996).

176. T. Katsuno *et al.*, Deficiency of Zonula Occludens-1 Causes Embryonic Lethal Phenotype Associated with Defected Yolk Sac Angiogenesis and Apoptosis of Embryonic Cells. *Mol. Biol. Cell*. **19**, 2465–2475 (2008).

177. K. Umeda *et al.*, ZO-1 and ZO-2 Independently Determine Where Claudins Are Polymerized in Tight-Junction Strand Formation. *Cell*. **126**, 741–754 (2006).

178. J. Xu *et al.*, Early embryonic lethality of mice lacking ZO-2, but Not ZO-3, reveals critical and nonredundant roles for individual zonula occludens proteins in mammalian development. *Mol. Cell. Biol.* **28**, 1669–1678 (2008).

179. J.-Y. Métais, C. Navarro, M.-J. Santoni, S. Audebert, J.-P. Borg, hScrib interacts with ZO-2 at the cell-cell junctions of epithelial cells. *FEBS Lett.* **579**, 3725–3730 (2005).

180. J. N. Murdoch *et al.*, Disruption of scribble (Scrb1) causes severe neural tube defects in the circletail mouse. *Hum. Mol. Genet.* **12**, 87–98 (2003).

181. A. Landy, Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annu. Rev. Biochem.* **58**, 913–949 (1989).

182. J. L. Hartley, G. F. Temple, M. A. Brasch, DNA Cloning Using In Vitro Site-Specific Recombination. *Genome Res.* **10**, 1788–1795 (2000).

183. T. Koressaar, M. Remm, Enhancements and modifications of primer design program Primer3. *Bioinformatics*. **23**, 1289–1291 (2007).

184. A. R. Niederriter *et al.*, In vivo modeling of the morbid human genome using Danio rerio. *J. Vis. Exp.* **78**, e50338 (2013).

185. L. Zheng, U. Baumann, J.-L. Reymond, An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **32** (2004).

186. A. E. Carpenter *et al.*, CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7** (2006).

187. M. Lakso *et al.*, Efficient in vivo manipulation of mouse genomic sequences at the zygote stage. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 5860–5 (1996).

188. G. Rigi *et al.*, Comparison of the extracellular full-length and truncated recombinant protein A production in Escherichia coli BL21 (DE3). *J. Paramed. Sci.* **6**, 35–45 (2015).

189. X. Y. Yang *et al.*, Mutations in the COPII vesicle component gene SEC24B are associated with human neural tube defects. *Hum. Mutat.* **34**, 1094–1101 (2013).

190. C. Wansleeben *et al.*, Planar cell polarity defects and defective Vangl2 trafficking in mutants for the COPII gene Sec24b. *Development*. **137**, 1067–1073 (2010).

191. A. K. Kenworthy, Imaging protein-protein interactions using fluorescence resonance energy transfer microscopy. *Methods*. **24**, 289–296 (2001).

192. Y. Luo, A. Batalao, H. Zhou, L. Zhu, Mammalian two-hybrid system: A complementary approach to the yeast two hybrid system. *Biotechniques*. **22**, 350–352 (1997).

193. K. J. Roux, D. I. Kim, B. Burke, BioID: A screen for protein-protein interactions. *Curr. Protoc. Protein Sci.* **19.23** (2013).

194. D. I. Kim *et al.*, An improved smaller biotin ligase for BioID proximity labeling. *Mol. Biol. Cell*. **27**, 1188–1196 (2016).

195. A. A. Mehus, R. H. Anderson, K. J. Roux, *BioID Identification of Lamin-Associated Proteins* (Elsevier Inc., ed. 1, 2016), vol. 569.

196. J. Melendez, M. Grogg, Y. Zheng, Signaling role of Cdc42 in regulating mammalian physiology. *J. Biol. Chem.* **286**, 2375–2381 (2011).

197. Y. Shi *et al.*, Identification of novel rare mutations of DACT1 in human neural tube defects. *Hum. Mutat.* **33**, 1450–1455 (2012).

198. R. Suriben, S. Kivimäe1, D. A. Fisher, R. T. Moon, B. N. R. Cheyette, Posterior Malformations in Dact1 mutant mice arise through misregulated Vangl2 at the Primitive Streak. *Nat. Genet.* **41**, 977–985 (2009).

199. J. Wen *et al.*, Loss of Dact1 disrupts planar cell polarity signaling by altering dishevelled activity and leads to posterior malformation in mice. *J. Biol. Chem.* **285**, 11023–11030 (2010).

200. S. Pirkmajer, A. V. Chibalin, Serum starvation: caveat emptor. *Am. J. Cell Physiol.* **301**, C272–C279 (2011).

201.	A. Robinson *et al.*, Mutations in the planar cell polarity genes CELSR1 and SCRIB are associated with the severe neural tube defect craniorachischisis. *Hum. Mutat.* **33**, 440–447 (2012).

202.	J. A. Curtin *et al.*, Mutation of Celsr1 Disrupts Planar Polarity of Inner Ear Hair Cells and Causes Severe Neural Tube Defects in the Mouse. *Curr. Biol.* **13**, 1129–1133 (2003).

203.	L. Polle, L. A. Rigano, R. Julian, K. Ireton, W. D. Schubert, Structural details of human tuba recruitment by InlC of Listeria monocytogenes elucidate bacterial cell-cell spreading. *Structure*. **22**, 304–14 (2014).

204.	R. S. Rajan, M. E. Illing, N. F. Bence, R. R. Kopito, Specificity in intracellular protein aggregation and inclusion body formation. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 13060–5 (2001).

205.	C. O'Farrell *et al.*, Transfected synphilin-1 forms cytoplasmic inclusions in HEK293 cells. *Mol. Brain Res.* **97**, 94–102 (2001).

206.	T. J. Broering, J. S. L. Parker, P. L. Joyce, J. Kim, M. L. Nibert, Mammalian Reovirus Nonstructural Protein NS Forms Large Inclusions and Colocalizes with Reovirus Microtubule-Associated Protein 2 in Transfected Cells. *J. Virol.* **76**, 8285–8297 (2002).

207.	K. Bersuker, M. Brandeis, R. R. Kopito, Protein misfolding specifies recruitment to cytoplasmic inclusion bodies. *J. Cell Biol.* **213**, 229–241 (2016).

208.	V. M. Bedell, S. E. Westcot, S. C. Ekker, Lessons from morpholino-based screening in zebrafish. *Brief. Funct. Genomics*. **10**, 181–188 (2011).

209.	A. Reynolds *et al.*, VANGL1 rare variants associated with neural tube defects affect convergent extension in zebrafish. *Mech. Dev.* **127**, 385–392 (2010).

Appendices

Appendix A: Wildtype and variant DNMBP-pENTR^{TM}221 Construct Sequences.

The *DNMBP* gene is in black while light grey represents the vector sequence. attL recombination

sites are underlined. Mutations introduced through site-directed mutagenesis are in bold.

>Wildtype-DNMBP (ENSG00000107554)

```
ATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGG
GCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCAC
CGACAAACAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGC
CTGGCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAAA
ACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTT
CGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTT
CGCCGCCACCATGGAGGCTGGCTCAGTGGTTCGAGCCATTTTTGACTTCTGCCCTAGCGTATCA
GAAGAACTGCCGCTCTTTGTGGGAGATATTATTGAGGTGCTGGCAGTGGTGGATGAATTTTGGC
TTCTAGGAAAGAAGGAGGATGTTACAGGACAATTCCCCAGCAGTTTTGTGGAAATTGTGACCAT
TCCCAGTCTAAAAGAGGGAGAGAGGCTGTTTGTGTGCATTTGTGAATTCACATCCCAAGAGTTG
GATAATCTTCCCCTCCATCGAGGTGACCTGGTGATTCTCGATGGCATTCCCACTGCAGGCTGGC
TGCAGGGCCGAAGCTGCTGGGGCGCACGGGGCTTCTTCCCATCTTCATGTGTCCGCGAGCTCTG
CCTCTCCTCACAGAGCCGGCAGTGGCACTCCCAGAGCGCCCTGTTTCAGATTCCGGAATATTCC
ATGGGACAAGCCCGGGCCCTAATGGGGCTTTCTGCTCAGCTGGATGAAGAGCTGGACTTCAGGG
AAGGGGATGTGATCACCATTATTGGAGTTCCTGAACCAGGCTGGTTTGAAGGGGAGTTAGAGGG
CCGAAGAGGCATTTTTCCAGAAGGTTTTGTAGAGCTGTTGGGGCCCCTGAGGACTGTGGATGAG
TCAGTAAGTTCTGGAAATCAAGATGACTGCATTGTTAATGGTGAAGTAGATACCCCTGTAGGAG
AAGAAGAGATAGGGCCGGATGAGGATGAGGAGGAGCCAGGGACCTATGGGGTCGCCCTGTACAG
ATTCCAAGCCCTGGAGCCAAATGAGCTGGATTTCGAGGTCGGGGATAAAATCCGAATTCTGGCG
ACCTTGGAAGATGGCTGGCTGGAAGGATCCCTGAAGGGCAGGACAGGCATCTTTCCTTACCGGT
TTGTGAAATTATGTCCTGACACACGGGTGGAGGAAACCATGGCTCTGCCCCAGGAAGGCAGCCT
TGCCAGGATCCCGGAAACTTCTTTGGATTGTTTGGAGAACACCTTAGGAGTAGAGGAACAAAGA
CATGAAACCAGTGACCATGAGGCCGAGGAGCCTGACTGCATTATTTCTGAAGCACCCACTTCTC
CCCTCGGTCATCTGACTTCAGAGTATGACACAGACAGAAACTCTTATCAGGACGAGGACACCGC
AGGAGGGCCCCCGAGAAGCCCAGGCGTGGAGTGGGAAATGCCTCTTGCCACAGACTCTCCCACA
TCTGACCCTACAGAAGTAGTCAATGGTATTTCCTCCCAACCTCAGGTCCCTTTTCATCCCAACT
TGCAGAAAAGCCAGTATTATTCTACAGTGGGAGGGAGCCACCCGCACTCAGAACAGTACCCCGA
CCTTCTTCCCCTAGAAGCAAGGACTAGAGACTATGCCAGCCTACCTCCCAAAAGAATGTATTCC
CAGCTAAAAACTCTTCAGAAGCCAGTGCTCCCTCTTTACAGGGGCTCTTCTGTTTCAGCTTCAA
GGGTAGTCAAACCCAGACAATCAAGTCCTCAGCTCCACAACCTAGCAAGTTATACTAAAAAGCA
CCACACGTCCAGTGTTTACTCCATCTCAGAGAGATTGGAGATGAAGCCTGGTCCGCAAGCCCAA
GGGCTTGTTATGGAAGCAGCAACACATTCACAGGGAGACGGCAGCACTGACCTGGACTCGAAGC
TGACACAACAGCTGATCGAGTTTGAGAAGAGCTTGGCAGGGCCCGGCACAGAGCCAGATAAAAT
TTTACGCCACTTTTCAATCATGGACTTTAACTCTGAGAAGGATATTGTCCGAGGTTCCTCAAAG
TTAATCACCGAGCAGGAGCTGCCGGAAAGGAGAAAGGCCCTAAGGCCACCGCCACCTCGTCCCT
GTACTCCGGTATCCACTTCTCCCCATTTGCTGGTTGACCAGAACCTAAAACCTGCACCACCCTT
GGTGGTGCGACCCTCTCGCCCAGCTCCCCTGCCTCCCTCAGCACAGCAGAGAACGAATGCGGTA
TCCCCCAAGCTCCTATCTCGACACCGTCCTACCTGTGAGACCTTAGAAAAGGAGGGCCCTGGTC
ATATGGGAAGGAGTCTGGACCAGACCTCCCCATGCCCCTTAGTGCTGGTGAGGATTGAGGAAAT
```

GGAGCGGGACTTGGATATGTACAGTAGAGCTCAAGAAGAGCTAAACCTCATGCTGGAGGAGAAG
CAGGATGAATCATCAAGAGCAGAGACCCTCGAGGATCTCAAGTTCTGTGAAAGTAACATTGAAA
GTTTGAATATGGAACTCCAGCAACTAAGAGAAATGACGCTCCTCTCCTCCCAGTCTTCATCACT
GGTGGCCCCTTCTGGGTCTGTGTCTGCCGAAAATCCAGAGCAGAGGATGCTGGAGAAGAGAGCC
AAGGTCATAGAAGAACTTCTTCAGACAGAAAGAGACTACATTCGGGATCTGGAAATGTGTATTG
AGCGGATCATGGTACCCATGCAGCAGGCACAGGTACCAAACATTGATTTTGAGGGACTTTTTGG
AAATATGCAGATGGTGATTAAGGTCTCGAAGCAATTATTGGCTGCTCTGGAAATCAGCGATGCT
GTAGGACCTGTGTTTCTTGGTCACCGGGATGAGCTTGAGGGAACATACAAGATTTACTGCCAGA
ATCATGATGAGGCCATTGCGCTGCTTGAAATCTACGAGAAGGATGAGAAGATCCAGAAGCATCT
TCAGGACTCCTTGGCAGATCTGAAGAGCCTATACAACGAATGGGGATGCACAAATTATATTAAC
CTGGGCTCCTTCCTCATCAAACCAGTACAGAGAGTAATGCGTTACCCGCTGTTGCTAATGGAGT
TGCTGAATTCCACCCCAGAATCCCACCCAGATAAAGTGCCTTTAACCAATGCAGTCCTTGCGGT
CAAGGAAATCAACGTTAACATTAATGAATATAAACGGCGAAAGGACCTGGTCCTCAAGTACCGT
AAGGGTGATGAAGATAGCCTTATGGAGAAAATTTCCAAACTGAACATCCACTCCATCATCAAGA
AATCCAACCGAGTTAGCAGTCACCTGAAGCATCTCACTGGCTTTGCTCCTCAGATAAAAGATGA
AGTATTTGAAGAAACAGAAAAAAACTTCCGAATGCAAGAAGATTGATTAAGTCTTTTATCCGA
GACCTGTCTCTCTACCTCCAGCACATCCGGGAGTCCGCATGTGTGAAAGTGGTGGCTGCTGTGA
GCATGTGGGATGTGTGCATGGAGAGAGGACACCGGGACCTGGAGCAGTTTGAGAGGGTGCATCG
CTACATCAGTGACCAGCTCTTCACAAACTTTAAGGAGAGGACAGAGCGGCTTGTCATCTCCCCC
TTAAATCAGTTACTGAGCATGTTTACAGGGCCCCATAAGCTGGTACAGAAACGCTTTGACAAGC
TCCTGGACTTCTATAACTGTACAGAACGGGCAGAAAAGCTAAAGGACAAGAAGACCCTGGAGGA
GCTGCAGTCGGCCCGGAACAACTATGAGGCCCTGAATGCACAGCTGCTGGATGAGCTGCCCAAG
TTCCACCAGTACGCCCAGGGCCTCTTCACCAACTGTGTCCACGGCTATGCTGAAGCCCACTGTG
ACTTTGTGCACCAGGCTCTGGAGCAATTAAAGCCACTGCTTTCGTTACTCAAAGTGGCTGGCAG
AGAGGGAAACCTTATTGCCATCTTCCACGAAGAGCACAGCAGAGTTCTGCAGCAACTCCAGGTT
TTTACCTTCTTCCCGGAGTCTCTTCCAGCTACCAAGAAGCCATTTGAGAGGAAAACCATTGACC
GCCAGTCTGCTCGAAAGCCACTCCTGGGCCTGCCAAGTTACATGCTACAGTCAGAAGAACTCCG
GGCCTCCCTCCTGGCCAGGTATCCCCCTGAAAAACTCTTCCAGGCAGAACGGAACTTCAATGCT
GCTCAAGACTTGGATGTCTCACTTTTGGAAGGTGACCTGGTGGGTGTGATTAAGAAAAAAGACC
CCATGGGCAGCCAGAACCGCTGGCTGATTGACAATGGAGTCACCAAAGGCTTCGTGTACAGCTC
TTTCCTAAAGCCCTACAATCCTCGCCGCAGCCACTCCGATGCCTCCGTGGGTAGCCACTCCTCC
ACAGAGTCTGAGCACGGCAGCTCCTCCCCCAGGTTCCCACGCCAGAACAGCGGCAGCACCCTGA
CCTTCAACCCCAGCAGCATGGCTGTATCCTTTACCTCGGGGTCTTGCCAGAAGCAGCCTCAAGA
TGCATCTCCTCCGCCAAAAGAATGTGACCAAGGAACTCTCAGTGCATCCCTAAATCCGAGTAAT
TCAGAGAGTAGTCCTTCCAGATGCCCTTCAGACCCAGACTCCACCTCCCAGCCAAGGTCAGGGG
ACTCTGCAGATGTAGCTAGAGATGTAAAGCAACCCACTGCCACGCCGAGGAGCTACCGGAACTT
CAGGCATCCAGAAATAGTTGGCTACTCCGTACCAGGACGAAATGGGCAAAGTCAAGACCTCGTC
AAAGGATGTGCAAGAACAGCCCAGGCTCCGGAAGACAGAAGTACAGAGCCAGATGGCAGTGAGG
CAGAAGGCAACCAGGTCTATTTTGCTGTCTACACCTTCAAGGCACGAAACCCAAATGAGCTGAG
CGTGTCAGCCAATCAGAAACTCAAGATCCTCGAGTTTAAAGATGTTACAGGAAATACAGAGTGG
TGGTTAGCTGAGGTTAACGGGAAGAAGGGCTACGTTCCCTCCAATTATATCCGCAAAACCGAGT
ACACCTGAGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTG
TTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATATCCCC
TATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATC
TCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACAT
AAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAAT
TCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTG

>N373K-DNMBP

ATGGAGGCTGGCTCAGTGGTTCGAGCCATTTTTGACTTCTGCCCTAGCGTATCA
GAAGAACTGCCGCTCTTTGTGGGAGATATTATTGAGGTGCTGGCAGTGGTGGATGAATTTTGGC
TTCTAGGAAAGAAGGAGGATGTTACAGGACAATTCCCCAGCAGTTTTGTGGAAATTGTGACCAT
TCCCAGTCTAAAAGAGGGAGAGAGGCTGTTTGTGTGCATTTGTGAATTCACATCCCAAGAGTTG
GATAATCTTCCCCTCCATCGAGGTGACCTGGTGATTCTCGATGGCATTCCCACTGCAGGCTGGC
TGCAGGGCCGAAGCTGCTGGGGCGCACGGGGCTTCTTCCCATCTTCATGTGTCCGCGAGCTCTG
CCTCTCCTCACAGAGCCGGCAGTGGCACTCCCAGAGCGCCCTGTTTCAGATTCCGGAATATTCC
ATGGGACAAGCCCGGGCCCTAATGGGGCTTTCTGCTCAGCTGGATGAAGAGCTGGACTTCAGGG
AAGGGGATGTGATCACCATTATTGGAGTTCCTGAACCAGGCTGGTTTGAAGGGGAGTTAGAGGG
CCGAAGAGGCATTTTTCCAGAAGGTTTTGTAGAGCTGTTGGGGCCCCTGAGGACTGTGGATGAG
TCAGTAAGTTCTGGAAATCAAGATGACTGCATTGTTAATGGTGAAGTAGATACCCCTGTAGGAG
AAGAAGAGATAGGGCCGGATGAGGATGAGGAGGAGCCAGGGACCTATGGGGTCGCCCTGTACAG
ATTCCAAGCCCTGGAGCCAAATGAGCTGGATTTCGAGGTCGGGGATAAAATCCGAATTCTGGCG
ACCTTGGAAGATGGCTGGCTGGAAGGATCCCTGAAGGGCAGGACAGGCATCTTTCCTTACCGGT
TTGTGAAATTATGTCCTGACACACGGGTGGAGGAAACCATGGCTCTGCCCCAGGAAGGCAGCCT
TGCCAGGATCCCGGAAACTTCTTTGGATTGTTTGGAGAACACCTTAGGAGTAGAGGAACAAAGA
CATGAAACCAGTGACCATGAGGCCGAGGAGCCTGACTGCATTATTTCTGAAGCACCCACTTCTC
CCCTCGGTCATCTGACTTCAGAGTATGACACAGACAGAAA**G**TCTTATCAGGACGAGGACACCGC
AGGAGGGCCCCCGAGAAGCCCAGGCGTGGAGTGGGAAATGCCTCTTGCCACAGACTCTCCCACA
TCTGACCCTACAGAAGTAGTCAATGGTATTTCCTCCCAACCTCAGGTCCCTTTTCATCCCAACT
TGCAGAAAAGCCAGTATTATTCTACAGTGGGAGGGAGCCACCCGCACTCAGAACAGTACCCCGA
CCTTCTTCCCCTAGAAGCAAGGACTAGAGACTATGCCAGCCTACCTCCCAAAAGAATGTATTCC
CAGCTAAAAACTCTTCAGAAGCCAGTGCTCCCTCTTTACAGGGGCTCTTCTGTTTCAGCTTCAA
GGGTAGTCAAACCCAGACAATCAAGTCCTCAGCTCCACAACCTAGCAAGTTATACTAAAAAGCA
CCACACGTCCAGTGTTTACTCCATCTCAGAGAGATTGGAGATGAAGCCTGGTCCGCAAGCCCAA
GGGCTTGTTATGGAAGCAGCAACACATTCACAGGGAGACGGCAGCACTGACCTGGACTCGAAGC
TGACACAACAGCTGATCGAGTTTGAGAAGAGCTTGGCAGGGCCCGGCACAGAGCCAGATAAAAT
TTTACGCCACTTTTCAATCATGGACTTTAACTCTGAGAAGGATATTGTCCGAGGTTCCTCAAAG
TTAATCACCGAGCAGGAGCTGCCGGAAAGGAGAAAGGCCCTAAGGCCACCGCCACCTCGTCCCT
GTACTCCGGTATCCACTTCTCCCCATTTGCTGGTTGACCAGAACCTAAAACCTGCACCACCCTT
GGTGGTGCGACCCTCTCGCCCAGCTCCCCTGCCTCCCTCAGCACAGCAGAGAACGAATGCGGTA
TCCCCCAAGCTCCTATCTCGACACCGTCCTACCTGTGAGACCTTAGAAAAGGAGGGCCCTGGTC
ATATGGGAAGGAGTCTGGACCAGACCTCCCCATGCCCCTTAGTGCTGGTGAGGATTGAGGAAAT
GGAGCGGGACTTGGATATGTACAGTAGAGCTCAAGAAGAGCTAAACCTCATGCTGGAGGAGAAG
CAGGATGAATCATCAAGAGCAGAGACCCTCGAGGATCTCAAGTTCTGTGAAAGTAACATTGAAA
GTTTGAATATGGAACTCCAGCAACTAAGAGAAATGACGCTCCTCTCCTCCCAGTCTTCATCACT

GGTGGCCCCTTCTGGGTCTGTGTCTGCCGAAAATCCAGAGCAGAGGATGCTGGAGAAGAGAGCC
AAGGTCATAGAAGAACTTCTTCAGACAGAAAGAGACTACATTCGGGATCTGGAAATGTGTATTG
AGCGGATCATGGTACCCATGCAGCAGGCACAGGTACCAAACATTGATTTTGAGGGACTTTTTGG
AAATATGCAGATGGTGATTAAGGTCTCGAAGCAATTATTGGCTGCTCTGGAAATCAGCGATGCT
GTAGGACCTGTGTTTCTTGGTCACCGGGATGAGCTTGAGGGAACATACAAGATTTACTGCCAGA
ATCATGATGAGGCCATTGCGCTGCTTGAAATCTACGAGAAGGATGAGAAGATCCAGAAGCATCT
TCAGGACTCCTTGGCAGATCTGAAGAGCCTATACAACGAATGGGGATGCACAAATTATATTAAC
CTGGGCTCCTTCCTCATCAAACCAGTACAGAGAGTAATGCGTTACCCGCTGTTGCTAATGGAGT
TGCTGAATTCCACCCCAGAATCCCACCCAGATAAAGTGCCTTTAACCAATGCAGTCCTTGCGGT
CAAGGAAATCAACGTTAACATTAATGAATATAAACGGCGAAAGGACCTGGTCCTCAAGTACCGT
AAGGGTGATGAAGATAGCCTTATGGAGAAAATTTCCAAACTGAACATCCACTCCATCATCAAGA
AATCCAACCGAGTTAGCAGTCACCTGAAGCATCTCACTGGCTTTGCTCCTCAGATAAAAGATGA
AGTATTTGAAGAAACAGAAAAAAACTTCCGAATGCAAGAAGATTGATTAAGTCTTTTATCCGA
GACCTGTCTCTCTACCTCCAGCACATCCGGGAGTCCGCATGTGTGAAAGTGGTGGCTGCTGTGA
GCATGTGGGATGTGTGCATGGAGAGAGGACACCGGGACCTGGAGCAGTTTGAGAGGGTGCATCG
CTACATCAGTGACCAGCTCTTCACAAACTTTAAGGAGAGGACAGAGCGGCTTGTCATCTCCCCC
TTAAATCAGTTACTGAGCATGTTTACAGGGCCCCATAAGCTGGTACAGAAACGCTTTGACAAGC
TCCTGGACTTCTATAACTGTACAGAACGGGCAGAAAAGCTAAAGGACAAGAAGACCCTGGAGGA
GCTGCAGTCGGCCCGGAACAACTATGAGGCCCTGAATGCACAGCTGCTGGATGAGCTGCCCAAG
TTCCACCAGTACGCCCAGGGCCTCTTCACCAACTGTGTCCACGGCTATGCTGAAGCCCACTGTG
ACTTTGTGCACCAGGCTCTGGAGCAATTAAAGCCACTGCTTTCGTTACTCAAAGTGGCTGGCAG
AGAGGGAAACCTTATTGCCATCTTCCACGAAGAGCACAGCAGAGTTCTGCAGCAACTCCAGGTT
TTTACCTTCTTCCCGGAGTCTCTTCCAGCTACCAAGAAGCCATTTGAGAGGAAAACCATTGACC
GCCAGTCTGCTCGAAAGCCACTCCTGGGCCTGCCAAGTTACATGCTACAGTCAGAAGAACTCCG
GGCCTCCCTCCTGGCCAGGTATCCCCCTGAAAAACTCTTCCAGGCAGAACGGAACTTCAATGCT
GCTCAAGACTTGGATGTCTCACTTTTGGAAGGTGACCTGGTGGGTGTGATTAAGAAAAAAGACC
CCATGGGCAGCCAGAACCGCTGGCTGATTGACAATGGAGTCACCAAAGGCTTCGTGTACAGCTC
TTTCCTAAAGCCCTACAATCCTCGCCGCAGCCACTCCGATGCCTCCGTGGGTAGCCACTCCTCC
ACAGAGTCTGAGCACGGCAGCTCCTCCCCCAGGTTCCCACGCCAGAACAGCGGCAGCACCCTGA
CCTTCAACCCCAGCAGCATGGCTGTATCCTTTACCTCGGGGTCTTGCCAGAAGCAGCCTCAAGA
TGCATCTCCTCCGCCAAAAGAATGTGACCAAGGAACTCTCAGTGCATCCCTAAATCCGAGTAAT
TCAGAGAGTAGTCCTTCCAGATGCCCTTCAGACCCAGACTCCACCTCCCAGCCAAGGTCAGGGG
ACTCTGCAGATGTAGCTAGAGATGTAAAGCAACCCACTGCCACGCCGAGGAGCTACCGGAACTT
CAGGCATCCAGAAATAGTTGGCTACTCCGTACCAGGACGAAATGGGCAAAGTCAAGACCTCGTC
AAAGGATGTGCAAGAACAGCCCAGGCTCCGGAAGACAGAAGTACAGAGCCAGATGGCAGTGAGG
CAGAAGGCAACCAGGTCTATTTTGCTGTCTACACCTTCAAGGCACGAAACCCAAATGAGCTGAG
CGTGTCAGCCAATCAGAAACTCAAGATCCTCGAGTTTAAAGATGTTACAGGAAATACAGAGTGG
TGGTTAGCTGAGGTTAACGGGAAGAAGGGCTACGTTCCCTCCAATTATATCCGCAAAACCGAGT
ACACCTGAGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTG
TTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATATCCCC
TATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATC
TCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACAT
AAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAAT
TCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTG
CGACAATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGG
TAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCT
CTTCCGACCATCAAGCATTTTATCCGT

>P820L-DNMBP

ATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGG
GCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCAC
CGACAAACAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGC
CTGGCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAAA
ACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTT
CGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTT
CGCCGCCACCATGGAGGCTGGCTCAGTGGTTCGAGCCATTTTTGACTTCTGCCCTAGCGTATCA
GAAGAACTGCCGCTCTTTGTGGGAGATATTATTGAGGTGCTGGCAGTGGTGGATGAATTTTGGC
TTCTAGGAAAGAAGGAGGATGTTACAGGACAATTCCCCAGCAGTTTTGTGGAAATTGTGACCAT
TCCCAGTCTAAAAGAGGGAGAGAGGCTGTTTGTGTGCATTTGTGAATTCACATCCCAAGAGTTG
GATAATCTTCCCCTCCATCGAGGTGACCTGGTGATTCTCGATGGCATTCCCACTGCAGGCTGGC
TGCAGGGCCGAAGCTGCTGGGGCGCACGGGGCTTCTTCCCATCTTCATGTGTCCGCGAGCTCTG
CCTCTCCTCACAGAGCCGGCAGTGGCACTCCCAGAGCGCCCTGTTTCAGATTCCGGAATATTCC
ATGGGACAAGCCCGGGCCCTAATGGGGCTTTCTGCTCAGCTGGATGAAGAGCTGGACTTCAGGG
AAGGGGATGTGATCACCATTATTGGAGTTCCTGAACCAGGCTGGTTTGAAGGGGAGTTAGAGGG
CCGAAGAGGCATTTTTCCAGAAGGTTTTGTAGAGCTGTTGGGGCCCCTGAGGACTGTGGATGAG
TCAGTAAGTTCTGGAAATCAAGATGACTGCATTGTTAATGGTGAAGTAGATACCCCTGTAGGAG
AAGAAGAGATAGGGCCGGATGAGGATGAGGAGGAGCCAGGGACCTATGGGGTCGCCCTGTACAG
ATTCCAAGCCCTGGAGCCAAATGAGCTGGATTTCGAGGTCGGGGATAAAATCCGAATTCTGGCG
ACCTTGGAAGATGGCTGGCTGGAAGGATCCCTGAAGGGCAGGACAGGCATCTTTCCTTACCGGT
TTGTGAAATTATGTCCTGACACACGGGTGGAGGAAACCATGGCTCTGCCCCAGGAAGGCAGCCT
TGCCAGGATCCCGGAAACTTCTTTGGATTGTTTGGAGAACACCTTAGGAGTAGAGGAACAAAGA
CATGAAACCAGTGACCATGAGGCCGAGGAGCCTGACTGCATTATTTCTGAAGCACCCACTTCTC
CCCTCGGTCATCTGACTTCAGAGTATGACACAGACAGAAACTCTTATCAGGACGAGGACACCGC
AGGAGGGCCCCCGAGAAGCCCAGGCGTGGAGTGGGAAATGCCTCTTGCCCACAGACTCTCCCACA
TCTGACCCTACAGAAGTAGTCAATGGTATTTCCTCCCAACCTCAGGTCCCTTTTCATCCCAACT
TGCAGAAAGCCAGTATTATTCTACAGTGGGAGGGAGCCACCCGCACTCAGAACAGTACCCCGA
CCTTCTTCCCCTAGAAGCAAGGACTAGAGACTATGCCAGCCTACCTCCCAAAAGAATGTATTCC
CAGCTAAAAACTCTTCAGAAGCCAGTGCTCCCTCTTTACAGGGGCTCTTCTGTTTCAGCTTCAA
GGGTAGTCAAACCCAGACAATCAAGTCCTCAGCTCCACAACCTAGCAAGTTATACTAAAAAGCA
CCACACGTCCAGTGTTTACTCCATCTCAGAGAGATTGGAGATGAAGCCTGGTCCGCAAGCCCAA
GGGCTTGTTATGGAAGCAGCAACACATTCACAGGGAGACGGCAGCACTGACCTGGACTCGAAGC
TGACACAACAGCTGATCGAGTTTGAGAAGAGCTTGGCAGGGCCCGGCACAGAGCCAGATAAAAT
TTTACGCCACTTTTCAATCATGGACTTTAACTCTGAGAAGGATATTGTCCGAGGTTCCTCAAAG
TTAATCACCGAGCAGGAGCTGCCGGAAAGGAGAAAGGCCCTAAGGCCACCGCCACCTCGTCCCT
GTACTCCGGTATCCACTTCTCCCCATTTGCTGGTTGACCAGAACCTAAAACCTGCACCACCCTT
GGTGGTGCGACCCTCTCGCCCAGCTCCCCTGCCTCCCTCAGCACAGCAGAGAACGAATGCGGTA
TCCCCCAAGCTCCTATCTCGACACCGTCCTACCTGTGAGACCTTAGAAAAGGAGGGCCCTGGTC
ATATGGGAAGGAGTCTGGACCAGACCTCCCCATGCCCCTTAGTGCTGGTGAGGATTGAGGAAAT
GGAGCGGGACTTGGATATGTACAGTAGAGCTCAAGAAGAGCTAAACCTCATGCTGGAGGAGAAG
CAGGATGAATCATCAAGAGCAGAGACCCTCGAGGATCTCAAGTTCTGTGAAAGTAACATTGAAA
GTTTGAATATGGAACTCCAGCAACTAAGAGAAATGACGCTCCTCTCCTCCCAGTCTTCATCACT
GGTGGCCCCTTCTGGGTCTGTGTCTGCCGAAAATCCAGAGCAGAGGATGCTGGAGAAGAGAGCC
AAGGTCATAGAAGAACTTCTTCAGACAGAAAGAGACTACATTCGGGATCTGGAAATGTGTATTG
AGCGGATCATGGTACCCATGCAGCAGGCACAGGTAC**T**AAACATTGATTTTGAGGGACTTTTTGG
AAATATGCAGATGGTGATTAAGGTCTCGAAGCAATTATTGGCTGCTCTGGAAATCAGCGATGCT

GTAGGACCTGTGTTTCTTGGTCACCGGGATGAGCTTGAGGGAACATACAAGATTTACTGCCAGA
ATCATGATGAGGCCATTGCGCTGCTTGAAATCTACGAGAAGGATGAGAAGATCCAGAAGCATCT
TCAGGACTCCTTGGCAGATCTGAAGAGCCTATACAACGAATGGGGATGCACAAATTATATTAAC
CTGGGCTCCTTCCTCATCAAACCAGTACAGAGAGTAATGCGTTACCCGCTGTTGCTAATGGAGT
TGCTGAATTCCACCCCAGAATCCCACCCAGATAAAGTGCCTTTAACCAATGCAGTCCTTGCGGT
CAAGGAAATCAACGTTAACATTAATGAATATAAACGGCGAAAGGACCTGGTCCTCAAGTACCGT
AAGGGTGATGAAGATAGCCTTATGGAGAAAATTTCCAAACTGAACATCCACTCCATCATCAAGA
AATCCAACCGAGTTAGCAGTCACCTGAAGCATCTCACTGGCTTTGCTCCTCAGATAAAAGATGA
AGTATTTGAAGAAACAGAAAAAAACTTCCGAATGCAAGAAGATTGATTAAGTCTTTTATCCGA
GACCTGTCTCTCTACCTCCAGCACATCCGGGAGTCCGCATGTGTGAAAGTGGTGGCTGCTGTGA
GCATGTGGGATGTGTGCATGGAGAGAGGACACCGGGACCTGGAGCAGTTTGAGAGGGTGCATCG
CTACATCAGTGACCAGCTCTTCACAAACTTTAAGGAGAGGACAGAGCGGCTTGTCATCTCCCCC
TTAAATCAGTTACTGAGCATGTTTACAGGGCCCCATAAGCTGGTACAGAAACGCTTTGACAAGC
TCCTGGACTTCTATAACTGTACAGAACGGGCAGAAAAGCTAAAGGACAAGAAGACCCTGGAGGA
GCTGCAGTCGGCCCGGAACAACTATGAGGCCCTGAATGCACAGCTGCTGGATGAGCTGCCCAAG
TTCCACCAGTACGCCCAGGGCCTCTTCACCAACTGTGTCCACGGCTATGCTGAAGCCCACTGTG
ACTTTGTGCACCAGGCTCTGGAGCAATTAAAGCCACTGCTTTCGTTACTCAAAGTGGCTGGCAG
AGAGGGAAACCTTATTGCCATCTTCCACGAAGAGCACAGCAGAGTTCTGCAGCAACTCCAGGTT
TTTACCTTCTTCCCGGAGTCTCTTCCAGCTACCAAGAAGCCATTTGAGAGGAAAACCATTGACC
GCCAGTCTGCTCGAAAGCCACTCCTGGGCCTGCCAAGTTACATGCTACAGTCAGAAGAACTCCG
GGCCTCCCTCCTGGCCAGGTATCCCCCTGAAAAACTCTTCCAGGCAGAACGGAACTTCAATGCT
GCTCAAGACTTGGATGTCTCACTTTTGGAAGGTGACCTGGTGGGTGTGATTAAGAAAAAAGACC
CCATGGGCAGCCAGAACCGCTGGCTGATTGACAATGGAGTCACCAAAGGCTTCGTGTACAGCTC
TTTCCTAAAGCCCTACAATCCTCGCCGCAGCCACTCCGATGCCTCCGTGGGTAGCCACTCCTCC
ACAGAGTCTGAGCACGGCAGCTCCTCCCCCAGGTTCCCACGCCAGAACAGCGGCAGCACCCTGA
CCTTCAACCCCAGCAGCATGGCTGTATCCTTTACCTCGGGGTCTTGCCAGAAGCAGCCTCAAGA
TGCATCTCCTCCGCCAAAAGAATGTGACCAAGGAACTCTCAGTGCATCCCTAAATCCGAGTAAT
TCAGAGAGTAGTCCTTCCAGATGCCCTTCAGACCCAGACTCCACCTCCCAGCCAAGGTCAGGGG
ACTCTGCAGATGTAGCTAGAGATGTAAAGCAACCCACTGCCACGCCGAGGAGCTACCGGAACTT
CAGGCATCCAGAAATAGTTGGCTACTCCGTACCAGGACGAAATGGGCAAAGTCAAGACCTCGTC
AAAGGATGTGCAAGAACAGCCCAGGCTCCGGAAGACAGAAGTACAGAGCCAGATGGCAGTGAGG
CAGAAGGCAACCAGGTCTATTTTGCTGTCTACACCTTCAAGGCACGAAACCCAAATGAGCTGAG
CGTGTCAGCCAATCAGAAACTCAAGATCCTCGAGTTTAAAGATGTTACAGGAAATACAGAGTGG
TGGTTAGCTGAGGTTAACGGGAAGAAGGGCTACGTTCCCTCCAATTATATCCGCAAAACCGAGT
ACACCTGAGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTG
TTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATATCCCC
TATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATC
TCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACAT
AAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAAT
TCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTG
CGACAATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGG
TAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCT
CTTCCGACCATCAAGCATTTTATCCGT

>M831T-DNMBP

ATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGG
GCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCAC
CGACAAACAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGC
CTGGCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAAA
ACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTT
CGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTT
CGCCGCCACCATGGAGGCTGGCTCAGTGGTTCGAGCCATTTTTGACTTCTGCCCTAGCGTATCA
GAAGAACTGCCGCTCTTTGTGGGAGATATTATTGAGGTGCTGGCAGTGGTGGATGAATTTTGGC
TTCTAGGAAAGAAGGAGGATGTTACAGGACAATTCCCCAGCAGTTTTGTGGAAATTGTGACCAT
TCCCAGTCTAAAAGAGGGAGAGAGGCTGTTTGTGTGCATTTGTGAATTCACATCCCAAGAGTTG
GATAATCTTCCCCTCCATCGAGGTGACCTGGTGATTCTCGATGGCATTCCCACTGCAGGCTGGC
TGCAGGGCCGAAGCTGCTGGGGCGCACGGGGCTTCTTCCCATCTTCATGTGTCCGCGAGCTCTG
CCTCTCCTCACAGAGCCGGCAGTGGCACTCCCAGAGCGCCCTGTTTCAGATTCCGGAATATTCC
ATGGGACAAGCCCGGGCCCTAATGGGGCTTTCTGCTCAGCTGGATGAAGAGCTGGACTTCAGGG
AAGGGGATGTGATCACCATTATTGGAGTTCCTGAACCAGGCTGGTTTGAAGGGGAGTTAGAGGG
CCGAAGAGGCATTTTTCCAGAAGGTTTTGTAGAGCTGTTGGGGCCCCTGAGGACTGTGGATGAG
TCAGTAAGTTCTGGAAATCAAGATGACTGCATTGTTAATGGTGAAGTAGATACCCCTGTAGGAG
AAGAAGAGATAGGGCCGGATGAGGATGAGGAGGAGCCAGGGACCTATGGGGTCGCCCTGTACAG
ATTCCAAGCCCTGGAGCCAAATGAGCTGGATTTCGAGGTCGGGGATAAAATCCGAATTCTGGCG
ACCTTGGAAGATGGCTGGCTGGAAGGATCCCTGAAGGGCAGGACAGGCATCTTTCCTTACCGGT
TTGTGAAATTATGTCCTGACACACGGGTGGAGGAAACCATGGCTCTGCCCCAGGAAGGCAGCCT
TGCCAGGATCCCGGAAACTTCTTTGGATTGTTTGGAGAACACCTTAGGAGTAGAGGAACAAAGA
CATGAAACCAGTGACCATGAGGCCGAGGAGCCTGACTGCATTATTTCTGAAGCACCCACTTCTC
CCCTCGGTCATCTGACTTCAGAGTATGACACAGACAGAAACTCTTATCAGGACGAGGACACCGC
AGGAGGGCCCCCGAGAAGCCCAGGCGTGGAGTGGGAAATGCCTCTTGCCACAGACTCTCCCACA
TCTGACCCTACAGAAGTAGTCAATGGTATTTCCTCCCAACCTCAGGTCCCTTTTCATCCCAACT
TGCAGAAAGCCAGTATTATTCTACAGTGGGAGGGAGCCACCCGCACTCAGAACAGTACCCCGA
CCTTCTTCCCCTAGAAGCAAGGACTAGAGACTATGCCAGCCTACCTCCCAAAAGAATGTATTCC
CAGCTAAAAACTCTTCAGAAGCCAGTGCTCCCTCTTTACAGGGGCTCTTCTGTTTCAGCTTCAA
GGGTAGTCAAACCCAGACAATCAAGTCCTCAGCTCCACAACCTAGCAAGTTATACTAAAAAGCA
CCACACGTCCAGTGTTTACTCCATCTCAGAGAGATTGGAGATGAAGCCTGGTCCGCAAGCCCAA
GGGCTTGTTATGGAAGCAGCAACACATTCACAGGGAGACGGCAGCACTGACCTGGACTCGAAGC
TGACACAACAGCTGATCGAGTTTGAGAAGAGCTTGGCAGGGCCCGGCACAGAGCCAGATAAAAT
TTTACGCCACTTTTCAATCATGGACTTTAACTCTGAGAAGGATATTGTCCGAGGTTCCTCAAAG
TTAATCACCGAGCAGGAGCTGCCGGAAAGGAGAAAGGCCCTAAGGCCACCGCCACCTCGTCCCT
GTACTCCGGTATCCACTTCTCCCCATTTGCTGGTTGACCAGAACCTAAAACCTGCACCACCCTT
GGTGGTGCGACCCTCTCGCCCAGCTCCCCTGCCTCCCTCAGCACAGCAGAGAACGAATGCGGTA
TCCCCCAAGCTCCTATCTCGACACCGTCCTACCTGTGAGACCTTAGAAAAGGAGGGCCCTGGTC
ATATGGGAAGGAGTCTGGACCAGACCTCCCCATGCCCCTTAGTGCTGGTGAGGATTGAGGAAAT
GGAGCGGGACTTGGATATGTACAGTAGAGCTCAAGAAGAGCTAAACCTCATGCTGGAGGAGAAG
CAGGATGAATCATCAAGAGCAGAGACCCTCGAGGATCTCAAGTTCTGTGAAAGTAACATTGAAA
GTTTGAATATGGAACTCCAGCAACTAAGAGAAATGACGCTCCTCTCCTCCCAGTCTTCATCACT
GGTGGCCCCTTCTGGGTCTGTGTCTGCCGAAAATCCAGAGCAGAGGATGCTGGAGAAGAGAGCC
AAGGTCATAGAAGAACTTCTTCAGACAGAAAGAGACTACATTCGGGATCTGGAAATGTGTATTG
AGCGGATCATGGTACCCATGCAGCAGGCACAGGTACCAAACATTGATTTTGAGGGACTTTTTGG
AAATA**C**GCAGATGGTGATTAAGGTCTCGAAGCAATTATTGGCTGCTCTGGAAATCAGCGATGCT

150

```
GTAGGACCTGTGTTTCTTGGTCACCGGGATGAGCTTGAGGGAACATACAAGATTTACTGCCAGA
ATCATGATGAGGCCATTGCGCTGCTTGAAATCTACGAGAAGGATGAGAAGATCCAGAAGCATCT
TCAGGACTCCTTGGCAGATCTGAAGAGCCTATACAACGAATGGGGATGCACAAATTATATTAAC
CTGGGCTCCTTCCTCATCAAACCAGTACAGAGAGTAATGCGTTACCCGCTGTTGCTAATGGAGT
TGCTGAATTCCACCCCAGAATCCCACCCAGATAAAGTGCCTTTAACCAATGCAGTCCTTGCGGT
CAAGGAAATCAACGTTAACATTAATGAATATAAACGGCGAAAGGACCTGGTCCTCAAGTACCGT
AAGGGTGATGAAGATAGCCTTATGGAGAAAATTTCCAAACTGAACATCCACTCCATCATCAAGA
AATCCAACCGAGTTAGCAGTCACCTGAAGCATCTCACTGGCTTTGCTCCTCAGATAAAAGATGA
AGTATTTGAAGAAACAGAAAAAAACTTCCGAATGCAAGAAGATTGATTAAGTCTTTTATCCGA
GACCTGTCTCTCTACCTCCAGCACATCCGGGAGTCCGCATGTGTGAAAGTGGTGGCTGCTGTGA
GCATGTGGGATGTGTGCATGGAGAGAGGACACCGGGACCTGGAGCAGTTTGAGAGGGTGCATCG
CTACATCAGTGACCAGCTCTTCACAAACTTTAAGGAGAGGACAGAGCGGCTTGTCATCTCCCCC
TTAAATCAGTTACTGAGCATGTTTACAGGGCCCCATAAGCTGGTACAGAAACGCTTTGACAAGC
TCCTGGACTTCTATAACTGTACAGAACGGGCAGAAAAGCTAAAGGACAAGAAGACCCTGGAGGA
GCTGCAGTCGGCCCGGAACAACTATGAGGCCCTGAATGCACAGCTGCTGGATGAGCTGCCCAAG
TTCCACCAGTACGCCCAGGGCCTCTTCACCAACTGTGTCCACGGCTATGCTGAAGCCCACTGTG
ACTTTGTGCACCAGGCTCTGGAGCAATTAAAGCCACTGCTTTCGTTACTCAAAGTGGCTGGCAG
AGAGGGAAACCTTATTGCCATCTTCCACGAAGAGCACAGCAGAGTTCTGCAGCAACTCCAGGTT
TTTACCTTCTTCCCGGAGTCTCTTCCAGCTACCAAGAAGCCATTTGAGAGGAAAACCATTGACC
GCCAGTCTGCTCGAAAGCCACTCCTGGGCCTGCCAAGTTACATGCTACAGTCAGAAGAACTCCG
GGCCTCCCTCCTGGCCAGGTATCCCCCTGAAAAACTCTTCCAGGCAGAACGGAACTTCAATGCT
GCTCAAGACTTGGATGTCTCACTTTTGGAAGGTGACCTGGTGGGTGTGATTAAGAAAAAAGACC
CCATGGGCAGCCAGAACCGCTGGCTGATTGACAATGGAGTCACCAAAGGCTTCGTGTACAGCTC
TTTCCTAAAGCCCTACAATCCTCGCCGCAGCCACTCCGATGCCTCCGTGGGTAGCCACTCCTCC
ACAGAGTCTGAGCACGGCAGCTCCTCCCCCAGGTTCCCACGCCAGAACAGCGGCAGCACCCTGA
CCTTCAACCCCAGCAGCATGGCTGTATCCTTTACCTCGGGGTCTTGCCAGAAGCAGCCTCAAGA
TGCATCTCCTCCGCCAAAAGAATGTGACCAAGGAACTCTCAGTGCATCCCTAAATCCGAGTAAT
TCAGAGAGTAGTCCTTCCAGATGCCCTTCAGACCCAGACTCCACCTCCCAGCCAAGGTCAGGGG
ACTCTGCAGATGTAGCTAGAGATGTAAAGCAACCCACTGCCACGCCGAGGAGCTACCGGAACTT
CAGGCATCCAGAAATAGTTGGCTACTCCGTACCAGGACGAAATGGGCAAAGTCAAGACCTCGTC
AAAGGATGTGCAAGAACAGCCCAGGCTCCGGAAGACAGAAGTACAGAGCCAGATGGCAGTGAGG
CAGAAGGCAACCAGGTCTATTTTGCTGTCTACACCTTCAAGGCACGAAACCCAAATGAGCTGAG
CGTGTCAGCCAATCAGAAACTCAAGATCCTCGAGTTTAAAGATGTTACAGGAAATACAGAGTGG
TGGTTAGCTGAGGTTAACGGGAAGAAGGGCTACGTTCCCTCCAATTATATCCGCAAAACCGAGT
ACACCTGAGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTG
TTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATATCCCC
TATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATC
TCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACAT
AAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAAT
TCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTG
CGACAATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGG
TAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCT
CTTCCGACCATCAAGCATTTTATCCGT
```

ATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGG
GCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCAC
CGACAAACAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGC
CTGGCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAAA
ACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTT
CGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTT
CGCCGCCACCATGGAGGCTGGCTCAGTGGTTCGAGCCATTTTTGACTTCTGCCCTAGCGTATCA
GAAGAACTGCCGCTCTTTGTGGGAGATATTATTGAGGTGCTGGCAGTGGTGGATGAATTTTGGC
TTCTAGGAAAGAAGGAGGATGTTACAGGACAATTCCCCAGCAGTTTTGTGGAAATTGTGACCAT
TCCCAGTCTAAAAGAGGGAGAGAGGCTGTTTGTGTGCATTTGTGAATTCACATCCCAAGAGTTG
GATAATCTTCCCCTCCATCGAGGTGACCTGGTGATTCTCGATGGCATTCCCACTGCAGGCTGGC
TGCAGGGCCGAAGCTGCTGGGGCGCACGGGGCTTCTTCCCATCTTCATGTGTCCGCGAGCTCTG
CCTCTCCTCACAGAGCCGGCAGTGGCACTCCCAGAGCGCCCTGTTTCAGATTCCGGAATATTCC
ATGGGACAAGCCCGGGCCCTAATGGGGCTTTCTGCTCAGCTGGATGAAGAGCTGGACTTCAGGG
AAGGGGATGTGATCACCATTATTGGAGTTCCTGAACCAGGCTGGTTTGAAGGGGAGTTAGAGGG
CCGAAGAGGCATTTTTCCAGAAGGTTTTGTAGAGCTGTTGGGGCCCCTGAGGACTGTGGATGAG
TCAGTAAGTTCTGGAAATCAAGATGACTGCATTGTTAATGGTGAAGTAGATACCCCTGTAGGAG
AAGAAGAGATAGGGCCGGATGAGGATGAGGAGGAGCCAGGGACCTATGGGGTCGCCCTGTACAG
ATTCCAAGCCCTGGAGCCAAATGAGCTGGATTTCGAGGTCGGGGATAAAATCCGAATTCTGGCG
ACCTTGGAAGATGGCTGGCTGGAAGGATCCCTGAAGGGCAGGACAGGCATCTTTCCTTACCGGT
TTGTGAAATTATGTCCTGACACACGGGTGGAGGAAACCATGGCTCTGCCCCAGGAAGGCAGCCT
TGCCAGGATCCCGGAAACTTCTTTGGATTGTTTGGAGAACACCTTAGGAGTAGAGGAACAAAGA
CATGAAACCAGTGACCATGAGGCCGAGGAGCCTGACTGCATTATTTCTGAAGCACCCACTTCTC
CCCTCGGTCATCTGACTTCAGAGTATGACACAGACAGAAACTCTTATCAGGACGAGGACACCGC
AGGAGGGCCCCCGAGAAGCCCAGGCGTGGAGTGGGAAATGCCTCTTGCCACAGACTCTCCCACA
TCTGACCCTACAGAAGTAGTCAATGGTATTTCCTCCCAACCTCAGGTCCCTTTTCATCCCAACT
TGCAGAAAAGCCAGTATTATTCTACAGTGGGAGGGAGCCACCCGCACTCAGAACAGTACCCCGA
CCTTCTTCCCCTAGAAGCAAGGACTAGAGACTATGCCAGCCTACCTCCCAAAAGAATGTATTCC
CAGCTAAAAACTCTTCAGAAGCCAGTGCTCCCTCTTTACAGGGGCTCTTCTGTTTCAGCTTCAA
GGGTAGTCAAACCCAGACAATCAAGTCCTCAGCTCCACAACCTAGCAAGTTATACTAAAAAGCA
CCACACGTCCAGTGTTTACTCCATCTCAGAGAGATTGGAGATGAAGCCTGGTCCGCAAGCCCAA
GGGCTTGTTATGGAAGCAGCAACACATTCACAGGGAGACGGCAGCACTGACCTGGACTCGAAGC
TGACACAACAGCTGATCGAGTTTGAGAAGAGCTTGGCAGGGCCCGGCACAGAGCCAGATAAAAT
TTTACGCCACTTTTCAATCATGGACTTTAACTCTGAGAAGGATATTGTCCGAGGTTCCTCAAAG
TTAATCACCGAGCAGGAGCTGCCGGAAAGGAGAAAGGCCCTAAGGCCACCGCCACCTCGTCCCT
GTACTCCGGTATCCACTTCTCCCCATTTGCTGGTTGACCAGAACCTAAAACCTGCACCACCCTT
GGTGGTGCGACCCTCTCGCCCAGCTCCCCTGCCTCCCTCAGCACAGCAGAGAACGAATGCGGTA
TCCCCCAAGCTCCTATCTCGACACCGTCCTACCTGTGAGACCTTAGAAAAGGAGGGCCCTGGTC
ATATGGGAAGGAGTCTGGACCAGACCTCCCCATGCCCCTTAGTGCTGGTGAGGATTGAGGAAAT
GGAGCGGGACTTGGATATGTACAGTAGAGCTCAAGAAGAGCTAAACCTCATGCTGGAGGAGAAG
CAGGATGAATCATCAAGAGCAGAGACCCTCGAGGATCTCAAGTTCTGTGAAAGTAACATTGAAA
GTTTGAATATGGAACTCCAGCAACTAAGAGAAATGACGCTCCTCTCCTCCCAGTCTTCATCACT
GGTGGCCCCTTCTGGGTCTGTGTCTGCCGAAAATCCAGAGCAGAGGATGCTGGAGAAGAGAGCC
AAGGTCATAGAAGAACTTCTTCAGACAGAAAGAGACTACATTCGGGATCTGGAAATGTGTATTG
AGCGGATCATGGTACCCATGCAGCAGGCACAGGTACCAAACATTGATTTTGAGGGACTTTTTGG
AAATACGCAGATGGTGATTAAGGTCTCGAAGCAATTATTGGCTGCTCTGGAAATCAGCGATGCT

GTAGGACCTGTGTTTCTTGGTCACCGGGATGAGCTTGAGGGAACATACAAGATTTACTGCCAGA
ATCATGATGAGGCCATTGCGCTGCTTGAAATCTACGAGAAGGATGAGAAGATCCAGAAGCATCT
TCAGGACTCCTTGGCAGATCTGAAGAGCCTATACAACGAATGGGGATGCACAAATTATATTAAC
CTGGGCTCCTTCCTCATCAAACCAGTACAGAGAGTAATGCGTTACCCGCTGTTGCTAATGGAGT
TGCTGAATTCCACCCCAGAATCCCACC**T**AGATAAAGTGCCTTTAACCAATGCAGTCCTTGCGGT
CAAGGAAATCAACGTTAACATTAATGAATATAAACGGCGAAAGGACCTGGTCCTCAAGTACCGT
AAGGGTGATGAAGATAGCCTTATGGAGAAAATTTCCAAACTGAACATCCACTCCATCATCAAGA
AATCCAACCGAGTTAGCAGTCACCTGAAGCATCTCACTGGCTTTGCTCCTCAGATAAAAGATGA
AGTATTTGAAGAAACAGAAAAAAACTTCCGAATGCAAGAAAGATTGATTAAGTCTTTTATCCGA
GACCTGTCTCTCTACCTCCAGCACATCCGGGAGTCCGCATGTGTGAAAGTGGTGGCTGCTGTGA
GCATGTGGGATGTGTGCATGGAGAGAGGACACCGGGACCTGGAGCAGTTTGAGAGGGTGCATCG
CTACATCAGTGACCAGCTCTTCACAAACTTTAAGGAGAGGACAGAGCGGCTTGTCATCTCCCCC
TTAAATCAGTTACTGAGCATGTTTACAGGGCCCCATAAGCTGGTACAGAAACGCTTTGACAAGC
TCCTGGACTTCTATAACTGTACAGAACGGGCAGAAAAGCTAAAGGACAAGAAGACCCTGGAGGA
GCTGCAGTCGGCCCGGAACAACTATGAGGCCCTGAATGCACAGCTGCTGGATGAGCTGCCCAAG
TTCCACCAGTACGCCCAGGGCCTCTTCACCAACTGTGTCCACGGCTATGCTGAAGCCCACTGTG
ACTTTGTGCACCAGGCTCTGGAGCAATTAAAGCCACTGCTTTCGTTACTCAAAGTGGCTGGCAG
AGAGGGAAACCTTATTGCCATCTTCCACGAAGAGCACAGCAGAGTTCTGCAGCAACTCCAGGTT
TTTACCTTCTTCCCGGAGTCTCTTCCAGCTACCAAGAAGCCATTTGAGAGGAAAACCATTGACC
GCCAGTCTGCTCGAAAGCCACTCCTGGGCCTGCCAAGTTACATGCTACAGTCAGAAGAACTCCG
GGCCTCCCTCCTGGCCAGGTATCCCCCTGAAAAACTCTTCCAGGCAGAACGGAACTTCAATGCT
GCTCAAGACTTGGATGTCTCACTTTTGGAAGGTGACCTGGTGGGTGTGATTAAGAAAAAAGACC
CCATGGGCAGCCAGAACCGCTGGCTGATTGACAATGGAGTCACCAAAGGCTTCGTGTACAGCTC
TTTCCTAAAGCCCTACAATCCTCGCCGCAGCCACTCCGATGCCTCCGTGGGTAGCCACTCCTCC
ACAGAGTCTGAGCACGGCAGCTCCTCCCCCAGGTTCCCACGCCAGAACAGCGGCAGCACCCTGA
CCTTCAACCCCAGCAGCATGGCTGTATCCTTTACCTCGGGGTCTTGCCAGAAGCAGCCTCAAGA
TGCATCTCCTCCGCCAAAAGAATGTGACCAAGGAACTCTCAGTGCATCCCTAAATCCGAGTAAT
TCAGAGAGTAGTCCTTCCAGATGCCCTTCAGACCCAGACTCCACCTCCCAGCCAAGGTCAGGGG
ACTCTGCAGATGTAGCTAGAGATGTAAAGCAACCCACTGCCACGCCGAGGAGCTACCGGAACTT
CAGGCATCCAGAAATAGTTGGCTACTCCGTACCAGGACGAAATGGGCAAAGTCAAGACCTCGTC
AAAGGATGTGCAAGAACAGCCCAGGCTCCGGAAGACAGAAGTACAGAGCCAGATGGCAGTGAGG
CAGAAGGCAACCAGGTCTATTTTGCTGTCTACACCTTCAAGGCACGAAACCCAAATGAGCTGAG
CGTGTCAGCCAATCAGAAACTCAAGATCCTCGAGTTTAAAGATGTTACAGGAAATACAGAGTGG
TGGTTAGCTGAGGTTAACGGGAAGAAGGGCTACGTTCCCTCCAATTATATCCGCAAAACCGAGT
ACACCTGAGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTG
TTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATATCCCC
TATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATC
TCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACAT
AAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAAT
TCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTG
CGACAATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGG
TAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCT
CTTCCGACCATCAAGCATTTTATCCGT

>R1024X-DNMBP

ATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGG
GCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCAC
CGACAAACAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGC
CTGGCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAAA
ACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTT
CGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTT
CGCCGCCACCATGGAGGCTGGCTCAGTGGTTCGAGCCATTTTTGACTTCTGCCCTAGCGTATCA
GAAGAACTGCCGCTCTTTGTGGGAGATATTATTGAGGTGCTGGCAGTGGTGGATGAATTTTGGC
TTCTAGGAAAGAAGGAGGATGTTACAGGACAATTCCCCAGCAGTTTTGTGGAAATTGTGACCAT
TCCCAGTCTAAAAGAGGGAGAGAGGCTGTTTGTGTGCATTTGTGAATTCACATCCCAAGAGTTG
GATAATCTTCCCCTCCATCGAGGTGACCTGGTGATTCTCGATGGCATTCCCACTGCAGGCTGGC
TGCAGGGCCGAAGCTGCTGGGGCGCACGGGGCTTCTTCCCATCTTCATGTGTCCGCGAGCTCTG
CCTCTCCTCACAGAGCCGGCAGTGGCACTCCCAGAGCGCCCTGTTTCAGATTCCGGAATATTCC
ATGGGACAAGCCCGGGCCCTAATGGGGCTTTCTGCTCAGCTGGATGAAGAGCTGGACTTCAGGG
AAGGGGATGTGATCACCATTATTGGAGTTCCTGAACCAGGCTGGTTTGAAGGGGAGTTAGAGGG
CCGAAGAGGCATTTTTCCAGAAGGTTTTGTAGAGCTGTTGGGGCCCCTGAGGACTGTGGATGAG
TCAGTAAGTTCTGGAAATCAAGATGACTGCATTGTTAATGGTGAAGTAGATACCCCTGTAGGAG
AAGAAGAGATAGGGCCGGATGAGGATGAGGAGGAGCCAGGGACCTATGGGGTCGCCCTGTACAG
ATTCCAAGCCCTGGAGCCAAATGAGCTGGATTTCGAGGTCGGGGATAAAATCCGAATTCTGGCG
ACCTTGGAAGATGGCTGGCTGGAAGGATCCCTGAAGGGCAGGACAGGCATCTTTCCTTACCGGT
TTGTGAAATTATGTCCTGACACACGGGTGGAGGAAACCATGGCTCTGCCCCAGGAAGGCAGCCT
TGCCAGGATCCCGGAAACTTCTTTGGATTGTTTGGAGAACACCTTAGGAGTAGAGGAACAAAGA
CATGAAACCAGTGACCATGAGGCCGAGGAGCCTGACTGCATTATTTCTGAAGCACCCACTTCTC
CCCTCGGTCATCTGACTTCAGAGTATGACACAGACAGAAACTCTTATCAGGACGAGGACACCGC
AGGAGGGCCCCCGAGAAGCCCAGGCGTGGAGTGGGAAATGCCTCTTGCCACAGACTCTCCCACA
TCTGACCCTACAGAAGTAGTCAATGGTATTTCCTCCCAACCTCAGGTCCCTTTTCATCCCAACT
TGCAGAAAGCCAGTATTATTCTACAGTGGGAGGGAGCCACCCGCACTCAGAACAGTACCCCGA
CCTTCTTCCCCTAGAAGCAAGGACTAGAGACTATGCCAGCCTACCTCCCAAAAGAATGTATTCC
CAGCTAAAAACTCTTCAGAAGCCAGTGCTCCCTCTTTACAGGGGCTCTTCTGTTTCAGCTTCAA
GGGTAGTCAAACCCAGACAATCAAGTCCTCAGCTCCACAACCTAGCAAGTTATACTAAAAAGCA
CCACACGTCCAGTGTTTACTCCATCTCAGAGAGATTGGAGATGAAGCCTGGTCCGCAAGCCCAA
GGGCTTGTTATGGAAGCAGCAACACATTCACAGGGAGACGGCAGCACTGACCTGGACTCGAAGC
TGACACAACAGCTGATCGAGTTTGAGAAGAGCTTGGCAGGGCCCGGCACAGAGCCAGATAAAAT
TTTACGCCACTTTTCAATCATGGACTTTAACTCTGAGAAGGATATTGTCCGAGGTTCCTCAAAG
TTAATCACCGAGCAGGAGCTGCCGGAAAGGAGAAAGGCCCTAAGGCCACCGCCACCTCGTCCCT
GTACTCCGGTATCCACTTCTCCCCATTTGCTGGTTGACCAGAACCTAAAACCTGCACCACCCTT
GGTGGTGCGACCCTCTCGCCCAGCTCCCCTGCCTCCCTCAGCACAGCAGAGAACGAATGCGGTA
TCCCCCAAGCTCCTATCTCGACACCGTCCTACCTGTGAGACCTTAGAAAGGAGGGCCCTGGTC
ATATGGGAAGGAGTCTGGACCAGACCTCCCCATGCCCCTTAGTGCTGGTGAGGATTGAGGAAAT
GGAGCGGGACTTGGATATGTACAGTAGAGCTCAAGAAGAGCTAAACCTCATGCTGGAGGAGAAG
CAGGATGAATCATCAAGAGCAGAGACCCTCGAGGATCTCAAGTTCTGTGAAAGTAACATTGAAA
GTTTGAATATGGAACTCCAGCAACTAAGAGAAATGACGCTCCTCTCCTCCCAGTCTTCATCACT
GGTGGCCCCTTCTGGGTCTGTGTCTGCCGAAAATCCAGAGCAGAGGATGCTGGAGAAGAGAGCC
AAGGTCATAGAAGAACTTCTTCAGACAGAAAGAGACTACATTCGGGATCTGGAAATGTGTATTG
AGCGGATCATGGTACCCATGCAGCAGGCACAGGTACCAAACATTGATTTTGAGGGACTTTTTGG
AAATATGCAGATGGTGATTAAGGTCTCGAAGCAATTATTGGCTGCTCTGGAAATCAGCGATGCT

GTAGGACCTGTGTTTCTTGGTCACCGGGATGAGCTTGAGGGAACATACAAGATTTACTGCCAGA
ATCATGATGAGGCCATTGCGCTGCTTGAAATCTACGAGAAGGATGAGAAGATCCAGAAGCATCT
TCAGGACTCCTTGGCAGATCTGAAGAGCCTATACAACGAATGGGGATGCACAAATTATATTAAC
CTGGGCTCCTTCCTCATCAAACCAGTACAGAGAGTAATGCGTTACCCGCTGTTGCTAATGGAGT
TGCTGAATTCCACCCCAGAATCCCACCCCGATAAAGTGCCTTTAACCAATGCAGTCCTTGCGGT
CAAGGAAATCAACGTTAACATTAATGAATATAAACGGCGAAAGGACCTGGTCCTCAAGTACCGT
AAGGGTGATGAAGATAGCCTTATGGAGAAAATTTCCAAACTGAACATCCACTCCATCATCAAGA
AATCCAACCGAGTTAGCAGTCACCTGAAGCATCTCACTGGCTTTGCTCCTCAGATAAAAGATGA
AGTATTTGAAGAAACAGAAAAAAACTTCCGAATGCAAGAAAGATTGATTAAGTCTTTTATC**T**GA
GACCTGTCTCTCTACCTCCAGCACATCCGGGAGTCCGCATGTGTGAAAGTGGTGGCTGCTGTGA
GCATGTGGGATGTGTGCATGGAGAGAGGACACCGGGACCTGGAGCAGTTTGAGAGGGTGCATCG
CTACATCAGTGACCAGCTCTTCACAAACTTTAAGGAGAGGACAGAGCGGCTTGTCATCTCCCCC
TTAAATCAGTTACTGAGCATGTTTACAGGGCCCCATAAGCTGGTACAGAAACGCTTTGACAAGC
TCCTGGACTTCTATAACTGTACAGAACGGGCAGAAAAGCTAAAGGACAAGAAGACCCTGGAGGA
GCTGCAGTCGGCCCGGAACAACTATGAGGCCCTGAATGCACAGCTGCTGGATGAGCTGCCCAAG
TTCCACCAGTACGCCCAGGGCCTCTTCACCAACTGTGTCCACGGCTATGCTGAAGCCCACTGTG
ACTTTGTGCACCAGGCTCTGGAGCAATTAAAGCCACTGCTTTCGTTACTCAAAGTGGCTGGCAG
AGAGGGAAACCTTATTGCCATCTTCCACGAAGAGCACAGCAGAGTTCTGCAGCAACTCCAGGTT
TTTACCTTCTTCCCGGAGTCTCTTCCAGCTACCAAGAAGCCATTTGAGAGGAAAACCATTGACC
GCCAGTCTGCTCGAAAGCCACTCCTGGGCCTGCCAAGTTACATGCTACAGTCAGAAGAACTCCG
GGCCTCCCTCCTGGCCAGGTATCCCCCTGAAAAACTCTTCCAGGCAGAACGGAACTTCAATGCT
GCTCAAGACTTGGATGTCTCACTTTTGGAAGGTGACCTGGTGGGTGTGATTAAGAAAAAAGACC
CCATGGGCAGCCAGAACCGCTGGCTGATTGACAATGGAGTCACCAAAGGCTTCGTGTACAGCTC
TTTCCTAAAGCCCTACAATCCTCGCCGCAGCCACTCCGATGCCTCCGTGGGTAGCCACTCCTCC
ACAGAGTCTGAGCACGGCAGCTCCTCCCCCAGGTTCCCACGCCAGAACAGCGGCAGCACCCTGA
CCTTCAACCCCAGCAGCATGGCTGTATCCTTTACCTCGGGGTCTTGCCAGAAGCAGCCTCAAGA
TGCATCTCCTCCGCCAAAAGAATGTGACCAAGGAACTCTCAGTGCATCCCTAAATCCGAGTAAT
TCAGAGAGTAGTCCTTCCAGATGCCCTTCAGACCCAGACTCCACCTCCCAGCCAAGGTCAGGGG
ACTCTGCAGATGTAGCTAGAGATGTAAAGCAACCCACTGCCACGCCGAGGAGCTACCGGAACTT
CAGGCATCCAGAAATAGTTGGCTACTCCGTACCAGGACGAAATGGGCAAAGTCAAGACCTCGTC
AAAGGATGTGCAAGAACAGCCCAGGCTCCGGAAGACAGAAGTACAGAGCCAGATGGCAGTGAGG
CAGAAGGCAACCAGGTCTATTTTGCTGTCTACACCTTCAAGGCACGAAACCCAAATGAGCTGAG
CGTGTCAGCCAATCAGAAACTCAAGATCCTCGAGTTTAAAGATGTTACAGGAAATACAGAGTGG
TGGTTAGCTGAGGTTAACGGGAAGAAGGGCTACGTTCCCTCCAATTATATCCGCAAAACCGAGT
ACACCTGAGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTG
TTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATATCCCC
TATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATC
TCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACAT
AAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAAT
TCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTG
CGACAATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGG
TAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCT
CTTCCGACCATCAAGCATTTTATCCGT

>R1154Q-DNMBP

ATGGAGGCTGGCTCAGTGGTTCGAGCCATTTTTGACTTCTGCCCTAGCGTATCA
GAAGAACTGCCGCTCTTTGTGGGAGATATTATTGAGGTGCTGGCAGTGGTGGATGAATTTTGGC
TTCTAGGAAAGAAGGAGGATGTTACAGGACAATTCCCCAGCAGTTTTGTGGAAATTGTGACCAT
TCCCAGTCTAAAAGAGGGAGAGAGGCTGTTTGTGTGCATTTGTGAATTCACATCCCAAGAGTTG
GATAATCTTCCCCTCCATCGAGGTGACCTGGTGATTCTCGATGGCATTCCCACTGCAGGCTGGC
TGCAGGGCCGAAGCTGCTGGGGCGCACGGGGCTTCTTCCCATCTTCATGTGTCCGCGAGCTCTG
CCTCTCCTCACAGAGCCGGCAGTGGCACTCCCAGAGCGCCCTGTTTCAGATTCCGGAATATTCC
ATGGGACAAGCCCGGGCCCTAATGGGGCTTTCTGCTCAGCTGGATGAAGAGCTGGACTTCAGGG
AAGGGGATGTGATCACCATTATTGGAGTTCCTGAACCAGGCTGGTTTGAAGGGGAGTTAGAGGG
CCGAAGAGGCATTTTTCCAGAAGGTTTTGTAGAGCTGTTGGGGCCCCTGAGGACTGTGGATGAG
TCAGTAAGTTCTGGAAATCAAGATGACTGCATTGTTAATGGTGAAGTAGATACCCCTGTAGGAG
AAGAAGAGATAGGGCCGGATGAGGATGAGGAGGAGCCAGGGACCTATGGGGTCGCCCTGTACAG
ATTCCAAGCCCTGGAGCCAAATGAGCTGGATTTCGAGGTCGGGGATAAAATCCGAATTCTGGCG
ACCTTGGAAGATGGCTGGCTGGAAGGATCCCTGAAGGGCAGGACAGGCATCTTTCCTTACCGGT
TTGTGAAATTATGTCCTGACACACGGGTGGAGGAAACCATGGCTCTGCCCCAGGAAGGCAGCCT
TGCCAGGATCCCGGAAACTTCTTTGGATTGTTTGGAGAACACCTTAGGAGTAGAGGAACAAAGA
CATGAAACCAGTGACCATGAGGCCGAGGAGCCTGACTGCATTATTTCTGAAGCACCCACTTCTC
CCCTCGGTCATCTGACTTCAGAGTATGACACAGACAGAAACTCTTATCAGGACGAGGACACCGC
AGGAGGGCCCCCGAGAAGCCCAGGCGTGGAGTGGGAAATGCCTCTTGCCACAGACTCTCCCACA
TCTGACCCTACAGAAGTAGTCAATGGTATTTCCTCCCAACCTCAGGTCCCTTTTCATCCCAACT
TGCAGAAAGCCAGTATTATTCTACAGTGGGAGGGAGCCACCCGCACTCAGAACAGTACCCCGA
CCTTCTTCCCCTAGAAGCAAGGACTAGAGACTATGCCAGCCTACCTCCCAAAAGAATGTATTCC
CAGCTAAAAACTCTTCAGAAGCCAGTGCTCCCTCTTTACAGGGGCTCTTCTGTTTCAGCTTCAA
GGGTAGTCAAACCCAGACAATCAAGTCCTCAGCTCCACAACCTAGCAAGTTATACTAAAAAGCA
CCACACGTCCAGTGTTTACTCCATCTCAGAGAGATTGGAGATGAAGCCTGGTCCGCAAGCCCAA
GGGCTTGTTATGGAAGCAGCAACACATTCACAGGGAGACGGCAGCACTGACCTGGACTCGAAGC
TGACACAACAGCTGATCGAGTTTGAGAAGAGCTTGGCAGGGCCCGGCACAGAGCCAGATAAAAT
TTTACGCCACTTTTCAATCATGGACTTTAACTCTGAGAAGGATATTGTCCGAGGTTCCTCAAAG
TTAATCACCGAGCAGGAGCTGCCGGAAAGGAGAAAGGCCCTAAGGCCACCGCCACCTCGTCCCT
GTACTCCGGTATCCACTTCTCCCCATTTGCTGGTTGACCAGAACCTAAAACCTGCACCACCCTT
GGTGGTGCGACCTCTCGCCCAGCTCCCCTGCCTCCCTCAGCACAGCAGAGAACGAATGCGGTA
TCCCCCAAGCTCCTATCTCGACACCGTCCTACCTGTGAGACCTTAGAAAAGGAGGGCCCTGGTC
ATATGGGAAGGAGTCTGGACCAGACCTCCCCATGCCCCTTAGTGCTGGTGAGGATTGAGGAAAT
GGAGCGGGACTTGGATATGTACAGTAGAGCTCAAGAAGAGCTAAACCTCATGCTGGAGGAGAAG
CAGGATGAATCATCAAGAGCAGAGACCCTCGAGGATCTCAAGTTCTGTGAAAGTAACATTGAAA
GTTTGAATATGGAACTCCAGCAACTAAGAGAAATGACGCTCCTCTCCTCCCAGTCTTCATCACT
GGTGGCCCCTTCTGGGTCTGTGTCTGCCGAAAATCCAGAGCAGAGGATGCTGGAGAAGAGAGCC
AAGGTCATAGAAGAACTTCTTCAGACAGAAAGAGACTACATTCGGGATCTGGAAATGTGTATTG
AGCGGATCATGGTACCCATGCAGCAGGCACAGGTACCAAACATTGATTTTGAGGGACTTTTTGG
AAATATGCAGATGGTGATTAAGGTCTCGAAGCAATTATTGGCTGCTCTGGAAATCAGCGATGCT

156

GTAGGACCTGTGTTTCTTGGTCACCGGGATGAGCTTGAGGGAACATACAAGATTTACTGCCAGA
ATCATGATGAGGCCATTGCGCTGCTTGAAATCTACGAGAAGGATGAGAAGATCCAGAAGCATCT
TCAGGACTCCTTGGCAGATCTGAAGAGCCTATACAACGAATGGGGATGCACAAATTATATTAAC
CTGGGCTCCTTCCTCATCAAACCAGTACAGAGAGTAATGCGTTACCCGCTGTTGCTAATGGAGT
TGCTGAATTCCACCCCAGAATCCCACCCCGATAAAGTGCCTTTAACCAATGCAGTCCTTGCGGT
CAAGGAAATCAACGTTAACATTAATGAATATAAACGGCGAAAGGACCTGGTCCTCAAGTACCGT
AAGGGTGATGAAGATAGCCTTATGGAGAAAATTTCCAAACTGAACATCCACTCCATCATCAAGA
AATCCAACCGAGTTAGCAGTCACCTGAAGCATCTCACTGGCTTTGCTCCTCAGATAAAAGATGA
AGTATTTGAAGAAACAGAAAAAAACTTCCGAATGCAAGAAGATTGATTAAGTCTTTTATCTG1
AGACCTGTCTCTCTACCTCCAGCACATCCGGGAGTCCGCATGTGTGAAAGTGGTGGCTGCTGTG
AGCATGTGGGATGTGTGCATGGAGAGAGGACACCGGGACCTGGAGCAGTTTGAGAGGGTGCATC
GCTACATCAGTGACCAGCTCTTCACAAACTTTAAGGAGAGGACAGAGCGGCTTGTCATCTCCCC
CTTAAATCAGTTACTGAGCATGTTTACAGGGCCCCATAAGCTGGTACAGAAACGCTTTGACAAG
CTCCTGGACTTCTATAACTGTACAGAACGGGCAGAAAAGCTAAAGGACAAGAAGACCCTGGAGG
AGCTGCAGTCGGCCC**A**GAACAACTATGAGGCCCTGAATGCACAGCTGCTGGATGAGCTGCCCAA
GTTCCACCAGTACGCCCAGGGCCTCTTCACCAACTGTGTCCACGGCTATGCTGAAGCCCACTGT
GACTTTGTGCACCAGGCTCTGGAGCAATTAAAGCCACTGCTTTCGTTACTCAAAGTGGCTGGCA
GAGAGGGAAACCTTATTGCCATCTTCCACGAAGAGCACAGCAGAGTTCTGCAGCAACTCCAGGT
TTTTACCTTCTTCCCGGAGTCTCTTCCAGCTACCAAGAAGCCATTTGAGAGGAAAACCATTGAC
CGCCAGTCTGCTCGAAAGCCACTCCTGGGCCTGCCAAGTTACATGCTACAGTCAGAAGAACTCC
GGGCCTCCCTCCTGGCCAGGTATCCCCCTGAAAAACTCTTCCAGGCAGAACGGAACTTCAATGC
TGCTCAAGACTTGGATGTCTCACTTTTGGAAGGTGACCTGGTGGGTGTGATTAAGAAAAAAGAC
CCCATGGGCAGCCAGAACCGCTGGCTGATTGACAATGGAGTCACCAAAGGCTTCGTGTACAGCT
CTTTCCTAAAGCCCTACAATCCTCGCCGCAGCCACTCCGATGCCTCCGTGGGTAGCCACTCCTC
CACAGAGTCTGAGCACGGCAGCTCCTCCCCCAGGTTCCCACGCCAGAACAGCGGCAGCACCCTG
ACCTTCAACCCCAGCAGCATGGCTGTATCCTTTACCTCGGGGTCTTGCCAGAAGCAGCCTCAAG
ATGCATCTCCTCCGCCAAAAGAATGTGACCAAGGAACTCTCAGTGCATCCCTAAATCCGAGTAA
TTCAGAGAGTAGTCCTTCCAGATGCCCTTCAGACCCAGACTCCACCTCCCAGCCAAGGTCAGGG
GACTCTGCAGATGTAGCTAGAGATGTAAAGCAACCCACTGCCACGCCGAGGAGCTACCGGAACT
TCAGGCATCCAGAAATAGTTGGCTACTCCGTACCAGGACGAAATGGGCAAAGTCAAGACCTCGT
CAAAGGATGTGCAAGAACAGCCCAGGCTCCGGAAGACAGAAGTACAGAGCCAGATGGCAGTGAG
GCAGAAGGCAACCAGGTCTATTTTGCTGTCTACACCTTCAAGGCACGAAACCCAAATGAGCTGA
GCGTGTCAGCCAATCAGAAACTCAAGATCCTCGAGTTTAAAGATGTTACAGGAAATACAGAGTG
GTGGTTAGCTGAGGTTAACGGGAAGAAGGGCTACGTTCCCTCCAATTATATCCGCAAAACCGAG
TACACCTGAG<u>ACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTT</u>
<u>GTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATATCCC</u>
CTATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAAT
CTCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACA
TAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAA
TTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGT
GCGACAATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAG
GTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCC
TCTTCCGACCATCAAGCATTTTATCCGT

>R1376H-DNMBP

ATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGG
GCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCAC
CGACAAACAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGC
CTGGCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAAA
ACGACGGCCAGTCTTAAGCTCGGGCCC<u>CAAATAATGATTTTATTTTGACTGATAGTGACCTGTT
CGTTGCAACAAATTGATGAGCAATGCT</u>TTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTT
CGCCGCCACCATGGAGGCTGGCTCAGTGGTTCGAGCCATTTTTGACTTCTGCCCTAGCGTATCA
GAAGAACTGCCGCTCTTTGTGGGAGATATTATTGAGGTGCTGGCAGTGGTGGATGAATTTTGGC
TTCTAGGAAAGAAGGAGGATGTTACAGGACAATTCCCCAGCAGTTTTGTGGAAATTGTGACCAT
TCCCAGTCTAAAAGAGGGAGAGAGGCTGTTTGTGTGCATTTGTGAATTCACATCCCAAGAGTTG
GATAATCTTCCCCTCCATCGAGGTGACCTGGTGATTCTCGATGGCATTCCCACTGCAGGCTGGC
TGCAGGGCCGAAGCTGCTGGGGCGCACGGGGCTTCTTCCCATCTTCATGTGTCCGCGAGCTCTG
CCTCTCCTCACAGAGCCGGCAGTGGCACTCCCAGAGCGCCCTGTTTCAGATTCCGGAATATTCC
ATGGGACAAGCCCGGGCCCTAATGGGGCTTTCTGCTCAGCTGGATGAAGAGCTGGACTTCAGGG
AAGGGGATGTGATCACCATTATTGGAGTTCCTGAACCAGGCTGGTTTGAAGGGGAGTTAGAGGG
CCGAAGAGGCATTTTTCCAGAAGGTTTTGTAGAGCTGTTGGGGCCCCTGAGGACTGTGGATGAG
TCAGTAAGTTCTGGAAATCAAGATGACTGCATTGTTAATGGTGAAGTAGATACCCCTGTAGGAG
AAGAAGAGATAGGGCCGGATGAGGATGAGGAGGAGCCAGGGACCTATGGGGTCGCCCTGTACAG
ATTCCAAGCCCTGGAGCCAAATGAGCTGGATTTCGAGGTCGGGGATAAAATCCGAATTCTGGCG
ACCTTGGAAGATGGCTGGCTGGAAGGATCCCTGAAGGGCAGGACAGGCATCTTTCCTTACCGGT
TTGTGAAATTATGTCCTGACACACGGGTGGAGGAAACCATGGCTCTGCCCCAGGAAGGCAGCCT
TGCCAGGATCCCGGAAACTTCTTTGGATTGTTTGGAGAACACCTTAGGAGTAGAGGAACAAAGA
CATGAAACCAGTGACCATGAGGCCGAGGAGCCTGACTGCATTATTTCTGAAGCACCCACTTCTC
CCCTCGGTCATCTGACTTCAGAGTATGACACAGACAGAAACTCTTATCAGGACGAGGACACCGC
AGGAGGGCCCCCGAGAAGCCCAGGCGTGGAGTGGGAAATGCCTCTTGCCACAGACTCTCCCACA
TCTGACCCTACAGAAGTAGTCAATGGTATTTCCTCCCAACCTCAGGTCCCTTTTCATCCCAACT
TGCAGAAAGCCAGTATTATTCTACAGTGGGAGGGAGCCACCCGCACTCAGAACAGTACCCCGA
CCTTCTTCCCCTAGAAGCAAGGACTAGAGACTATGCCAGCCTACCTCCCAAAAGAATGTATTCC
CAGCTAAAAACTCTTCAGAAGCCAGTGCTCCCTCTTTACAGGGGCTCTTCTGTTTCAGCTTCAA
GGGTAGTCAAACCCAGACAATCAAGTCCTCAGCTCCACAACCTAGCAAGTTATACTAAAAAGCA
CCACACGTCCAGTGTTTACTCCATCTCAGAGAGATTGGAGATGAAGCCTGGTCCGCAAGCCCAA
GGGCTTGTTATGGAAGCAGCAACACATTCACAGGGAGACGGCAGCACTGACCTGGACTCGAAGC
TGACACAACAGCTGATCGAGTTTGAGAAGAGCTTGGCAGGGCCCGGCACAGAGCCAGATAAAAT
TTTACGCCACTTTTCAATCATGGACTTTAACTCTGAGAAGGATATTGTCCGAGGTTCCTCAAAG
TTAATCACCGAGCAGGAGCTGCCGGAAAGGAGAAAGGCCCTAAGGCCACCGCCACCTCGTCCCT
GTACTCCGGTATCCACTTCTCCCCATTTGCTGGTTGACCAGAACCTAAAACCTGCACCACCCTT
GGTGGTGCGACCCTCTCGCCCAGCTCCCCTGCCTCCCTCAGCACAGCAGAGAACGAATGCGGTA
TCCCCCAAGCTCCTATCTCGACACCGTCCTACCTGTGAGACCTTAGAAAAGGAGGGCCCTGGTC
ATATGGGAAGGAGTCTGGACCAGACCTCCCCATGCCCCTTAGTGCTGGTGAGGATTGAGGAAAT
GGAGCGGGACTTGGATATGTACAGTAGAGCTCAAGAAGAGCTAAACCTCATGCTGGAGGAGAAG
CAGGATGAATCATCAAGAGCAGAGACCCTCGAGGATCTCAAGTTCTGTGAAAGTAACATTGAAA
GTTTGAATATGGAACTCCAGCAACTAAGAGAAATGACGCTCCTCTCCTCCCAGTCTTCATCACT
GGTGGCCCCTTCTGGGTCTGTGTCTGCCGAAAATCCAGAGCAGAGGATGCTGGAGAAGAGAGCC
AAGGTCATAGAAGAACTTCTTCAGACAGAAAGAGACTACATTCGGGATCTGGAAATGTGTATTG
AGCGGATCATGGTACCCATGCAGCAGGCACAGGTACCAAACATTGATTTTGAGGGACTTTTTGG
AAATATGCAGATGGTGATTAAGGTCTCGAAGCAATTATTGGCTGCTCTGGAAATCAGCGATGCT

GTAGGACCTGTGTTTCTTGGTCACCGGGATGAGCTTGAGGGAACATACAAGATTTACTGCCAGA
ATCATGATGAGGCCATTGCGCTGCTTGAAATCTACGAGAAGGATGAGAAGATCCAGAAGCATCT
TCAGGACTCCTTGGCAGATCTGAAGAGCCTATACAACGAATGGGGATGCACAAATTATATTAAC
CTGGGCTCCTTCCTCATCAAACCAGTACAGAGAGTAATGCGTTACCCGCTGTTGCTAATGGAGT
TGCTGAATTCCACCCCAGAATCCCACCCCGATAAAGTGCCTTTAACCAATGCAGTCCTTGCGGT
CAAGGAAATCAACGTTAACATTAATGAATATAAACGGCGAAAGGACCTGGTCCTCAAGTACCGT
AAGGGTGATGAAGATAGCCTTATGGAGAAAATTTCCAAACTGAACATCCACTCCATCATCAAGA
AATCCAACCGAGTTAGCAGTCACCTGAAGCATCTCACTGGCTTTGCTCCTCAGATAAAAGATGA
AGTATTTGAAGAAACAGAAAAAAACTTCCGAATGCAAGAAGATTGATTAAGTCTTTTATCTGA
GACCTGTCTCTCTACCTCCAGCACATCCGGGAGTCCGCATGTGTGAAAGTGGTGGCTGCTGTGA
GCATGTGGGATGTGTGCATGGAGAGAGGACACCGGGACCTGGAGCAGTTTGAGAGGGTGCATCG
CTACATCAGTGACCAGCTCTTCACAAACTTTAAGGAGAGGACAGAGCGGCTTGTCATCTCCCCC
TTAAATCAGTTACTGAGCATGTTTACAGGGCCCCATAAGCTGGTACAGAAACGCTTTGACAAGC
TCCTGGACTTCTATAACTGTACAGAACGGGCAGAAAAGCTAAAGGACAAGAAGACCCTGGAGGA
GCTGCAGTCGGCCCGGAACAACTATGAGGCCCTGAATGCACAGCTGCTGGATGAGCTGCCCAAG
TTCCACCAGTACGCCCAGGGCCTCTTCACCAACTGTGTCCACGGCTATGCTGAAGCCCACTGTG
ACTTTGTGCACCAGGCTCTGGAGCAATTAAAGCCACTGCTTTCGTTACTCAAAGTGGCTGGCAG
AGAGGGAAACCTTATTGCCATCTTCCACGAAGAGCACAGCAGAGTTCTGCAGCAACTCCAGGTT
TTTACCTTCTTCCCGGAGTCTCTTCCAGCTACCAAGAAGCCATTTGAGAGGAAAACCATTGACC
GCCAGTCTGCTCGAAAGCCACTCCTGGGCCTGCCAAGTTACATGCTACAGTCAGAAGAACTCCG
GGCCTCCCTCCTGGCCAGGTATCCCCCTGAAAAACTCTTCCAGGCAGAACGGAACTTCAATGCT
GCTCAAGACTTGGATGTCTCACTTTTGGAAGGTGACCTGGTGGGTGTGATTAAGAAAAAAGACC
CCATGGGCAGCCAGAACCGCTGGCTGATTGACAATGGAGTCACCAAAGGCTTCGTGTACAGCTC
TTTCCTAAAGCCCTACAATCCTCGCCGCAGCCACTCCGATGCCTCCGTGGGTAGCCACTCCTCC
ACAGAGTCTGAGCACGGCAGCTCCTCCCCCAGGTTCCCAC**A**CCAGAACAGCGGCAGCACCCTGA
CCTTCAACCCCAGCAGCATGGCTGTATCCTTTACCTCGGGGTCTTGCCAGAAGCAGCCTCAAGA
TGCATCTCCTCCGCCAAAAGAATGTGACCAAGGAACTCTCAGTGCATCCCTAAATCCGAGTAAT
TCAGAGAGTAGTCCTTCCAGATGCCCTTCAGACCCAGACTCCACCTCCCAGCCAAGGTCAGGGG
ACTCTGCAGATGTAGCTAGAGATGTAAAGCAACCCACTGCCACGCCGAGGAGCTACCGGAACTT
CAGGCATCCAGAAATAGTTGGCTACTCCGTACCAGGACGAAATGGGCAAAGTCAAGACCTCGTC
AAAGGATGTGCAAGAACAGCCCAGGCTCCGGAAGACAGAAGTACAGAGCCAGATGGCAGTGAGG
CAGAAGGCAACCAGGTCTATTTTGCTGTCTACACCTTCAAGGCACGAAACCCAAATGAGCTGAG
CGTGTCAGCCAATCAGAAACTCAAGATCCTCGAGTTTAAAGATGTTACAGGAAATACAGAGTGG
TGGTTAGCTGAGGTTAACGGGAAGAAGGGCTACGTTCCCTCCAATTATATCCGCAAAACCGAGT
ACACCTGAG<ins>ACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTG
TTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTG</ins>CCATCCAGCTGATATCCCC
TATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATC
TCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACAT
AAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAAT
TCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTG
CGACAATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGG
TAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCT
CTTCCGACCATCAAGCATTTTATCCGT

>C1413W-DNMBP

ATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGG
GCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCAC
CGACAAACAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGC
CTGGCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAAA
ACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTT
CGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTT
CGCCGCCACCATGGAGGCTGGCTCAGTGGTTCGAGCCATTTTTGACTTCTGCCCTAGCGTATCA
GAAGAACTGCCGCTCTTTGTGGGAGATATTATTGAGGTGCTGGCAGTGGTGGATGAATTTTGGC
TTCTAGGAAAGAAGGAGGATGTTACAGGACAATTCCCCAGCAGTTTTGTGGAAATTGTGACCAT
TCCCAGTCTAAAAGAGGGAGAGAGGCTGTTTGTGTGCATTTGTGAATTCACATCCCAAGAGTTG
GATAATCTTCCCCTCCATCGAGGTGACCTGGTGATTCTCGATGGCATTCCCACTGCAGGCTGGC
TGCAGGGCCGAAGCTGCTGGGGCGCACGGGGCTTCTTCCCATCTTCATGTGTCCGCGAGCTCTG
CCTCTCCTCACAGAGCCGGCAGTGGCACTCCCAGAGCGCCCTGTTTCAGATTCCGGAATATTCC
ATGGGACAAGCCCGGGCCCTAATGGGGCTTTCTGCTCAGCTGGATGAAGAGCTGGACTTCAGGG
AAGGGGATGTGATCACCATTATTGGAGTTCCTGAACCAGGCTGGTTTGAAGGGGAGTTAGAGGG
CCGAAGAGGCATTTTTCCAGAAGGTTTTGTAGAGCTGTTGGGGCCCCTGAGGACTGTGGATGAG
TCAGTAAGTTCTGGAAATCAAGATGACTGCATTGTTAATGGTGAAGTAGATACCCCTGTAGGAG
AAGAAGAGATAGGGCCGGATGAGGATGAGGAGGAGCCAGGGACCTATGGGGTCGCCCTGTACAG
ATTCCAAGCCCTGGAGCCAAATGAGCTGGATTTCGAGGTCGGGGATAAAATCCGAATTCTGGCG
ACCTTGGAAGATGGCTGGCTGGAAGGATCCCTGAAGGGCAGGACAGGCATCTTTCCTTACCGGT
TTGTGAAATTATGTCCTGACACACGGGTGGAGGAAACCATGGCTCTGCCCCAGGAAGGCAGCCT
TGCCAGGATCCCGGAAACTTCTTTGGATTGTTTGGAGAACACCTTAGGAGTAGAGGAACAAAGA
CATGAAACCAGTGACCATGAGGCCGAGGAGCCTGACTGCATTATTTCTGAAGCACCCACTTCTC
CCCTCGGTCATCTGACTTCAGAGTATGACACAGACAGAAACTCTTATCAGGACGAGGACACCGC
AGGAGGGCCCCCGAGAAGCCCAGGCGTGGAGTGGGAAATGCCTCTTGCCACAGACTCTCCCACA
TCTGACCCTACAGAAGTAGTCAATGGTATTTCCTCCCAACCTCAGGTCCCTTTTCATCCCAACT
TGCAGAAAAGCCAGTATTATTCTACAGTGGGAGGGAGCCACCCGCACTCAGAACAGTACCCCGA
CCTTCTTCCCCTAGAAGCAAGGACTAGAGACTATGCCAGCCTACCTCCCAAAAGAATGTATTCC
CAGCTAAAAACTCTTCAGAAGCCAGTGCTCCCTCTTTACAGGGGCTCTTCTGTTTCAGCTTCAA
GGGTAGTCAAACCCAGACAATCAAGTCCTCAGCTCCACAACCTAGCAAGTTATACTAAAAAGCA
CCACACGTCCAGTGTTTACTCCATCTCAGAGAGATTGGAGATGAAGCCTGGTCCGCAAGCCCAA
GGGCTTGTTATGGAAGCAGCAACACATTCACAGGGAGACGGCAGCACTGACCTGGACTCGAAGC
TGACACAACAGCTGATCGAGTTTGAGAAGAGCTTGGCAGGGCCCGGCACAGAGCCAGATAAAAT
TTTACGCCACTTTTTCAATCATGGACTTTAACTCTGAGAAGGATATTGTCCGAGGTTCCTCAAAG
TTAATCACCGAGCAGGAGCTGCCGGAAAGGAGAAAGGCCCTAAGGCCACCGCCACCTCGTCCCT
GTACTCCGGTATCCACTTCTCCCCATTTGCTGGTTGACCAGAACCTAAAACCTGCACCACCCTT
GGTGGTGCGACCCTCTCGCCCAGCTCCCCTGCCTCCCTCAGCACAGCAGAGAACGAATGCGGTA
TCCCCCAAGCTCCTATCTCGACACCGTCCTACCTGTGAGACCTTAGAAAAGGAGGGCCCTGGTC
ATATGGGAAGGAGTCTGGACCAGACCTCCCCATGCCCCTTAGTGCTGGTGAGGATTGAGGAAAT
GGAGCGGGACTTGGATATGTACAGTAGAGCTCAAGAAGAGCTAAACCTCATGCTGGAGGAGAAG
CAGGATGAATCATCAAGAGCAGAGACCCTCGAGGATCTCAAGTTCTGTGAAAGTAACATTGAAA
GTTTGAATATGGAACTCCAGCAACTAAGAGAAATGACGCTCCTCTCCTCCCAGTCTTCATCACT
GGTGGCCCCTTCTGGGTCTGTGTCTGCCGAAAATCCAGAGCAGAGGATGCTGGAGAAGAGAGCC
AAGGTCATAGAAGAACTTCTTCAGACAGAAAGAGACTACATTCGGGATCTGGAAATGTGTATTG
AGCGGATCATGGTACCCATGCAGCAGGCACAGGTACCAAACATTGATTTTGAGGGACTTTTTGG
AAATATGCAGATGGTGATTAAGGTCTCGAAGCAATTATTGGCTGCTCTGGAAATCAGCGATGCT

GTAGGACCTGTGTTTCTTGGTCACCGGGATGAGCTTGAGGGAACATACAAGATTTACTGCCAGA
ATCATGATGAGGCCATTGCGCTGCTTGAAATCTACGAGAAGGATGAGAAGATCCAGAAGCATCT
TCAGGACTCCTTGGCAGATCTGAAGAGCCTATACAACGAATGGGGATGCACAAATTATATTAAC
CTGGGCTCCTTCCTCATCAAACCAGTACAGAGAGTAATGCGTTACCCGCTGTTGCTAATGGAGT
TGCTGAATTCCACCCCAGAATCCCACCCCGATAAAGTGCCTTTAACCAATGCAGTCCTTGCGGT
CAAGGAAATCAACGTTAACATTAATGAATATAAACGGCGAAAGGACCTGGTCCTCAAGTACCGT
AAGGGTGATGAAGATAGCCTTATGGAGAAAATTTCCAAACTGAACATCCACTCCATCATCAAGA
AATCCAACCGAGTTAGCAGTCACCTGAAGCATCTCACTGGCTTTGCTCCTCAGATAAAAGATGA
AGTATTTGAAGAAACAGAAAAAAACTTCCGAATGCAAGAAGATTGATTAAGTCTTTTATCTGA
GACCTGTCTCTCTACCTCCAGCACATCCGGGAGTCCGCATGTGTGAAAGTGGTGGCTGCTGTGA
GCATGTGGGATGTGTGCATGGAGAGAGGACACCGGGACCTGGAGCAGTTTGAGAGGGTGCATCG
CTACATCAGTGACCAGCTCTTCACAAACTTTAAGGAGAGGACAGAGCGGCTTGTCATCTCCCCC
TTAAATCAGTTACTGAGCATGTTTACAGGGCCCCATAAGCTGGTACAGAAACGCTTTGACAAGC
TCCTGGACTTCTATAACTGTACAGAACGGGCAGAAAAGCTAAAGGACAAGAAGACCCTGGAGGA
GCTGCAGTCGGCCCGGAACAACTATGAGGCCCTGAATGCACAGCTGCTGGATGAGCTGCCCAAG
TTCCACCAGTACGCCCAGGGCCTCTTCACCAACTGTGTCCACGGCTATGCTGAAGCCCACTGTG
ACTTTGTGCACCAGGCTCTGGAGCAATTAAAGCCACTGCTTTCGTTACTCAAAGTGGCTGGCAG
AGAGGGAAACCTTATTGCCATCTTCCACGAAGAGCACAGCAGAGTTCTGCAGCAACTCCAGGTT
TTTACCTTCTTCCCGGAGTCTCTTCCAGCTACCAAGAAGCCATTTGAGAGGAAAACCATTGACC
GCCAGTCTGCTCGAAAGCCACTCCTGGGCCTGCCAAGTTACATGCTACAGTCAGAAGAACTCCG
GGCCTCCCTCCTGGCCAGGTATCCCCCTGAAAAACTCTTCCAGGCAGAACGGAACTTCAATGCT
GCTCAAGACTTGGATGTCTCACTTTTGGAAGGTGACCTGGTGGGTGTGATTAAGAAAAAAGACC
CCATGGGCAGCCAGAACCGCTGGCTGATTGACAATGGAGTCACCAAAGGCTTCGTGTACAGCTC
TTTCCTAAAGCCCTACAATCCTCGCCGCAGCCACTCCGATGCCTCCGTGGGTAGCCACTCCTCC
ACAGAGTCTGAGCACGGCAGCTCCTCCCCCAGGTTCCCACGCCAGAACAGCGGCAGCACCCTGA
CCTTCAACCCCAGCAGCATGGCTGTATCCTTTACCTCGGGGTCTTGCCAGAAGCAGCCTCAAGA
TGCATCTCCTCCGCCAAAGAATG**G**GACCAAGGAACTCTCAGTGCATCCCTAAATCCGAGTAAT
TCAGAGAGTAGTCCTTCCAGATGCCCTTCAGACCCAGACTCCACCTCCCAGCCAAGGTCAGGGG
ACTCTGCAGATGTAGCTAGAGATGTAAAGCAACCCACTGCCACGCCGAGGAGCTACCGGAACTT
CAGGCATCCAGAAATAGTTGGCTACTCCGTACCAGGACGAAATGGGCAAAGTCAAGACCTCGTC
AAAGGATGTGCAAGAACAGCCCAGGCTCCGGAAGACAGAAGTACAGAGCCAGATGGCAGTGAGG
CAGAAGGCAACCAGGTCTATTTTGCTGTCTACACCTTCAAGGCACGAAACCCAAATGAGCTGAG
CGTGTCAGCCAATCAGAAACTCAAGATCCTCGAGTTTAAAGATGTTACAGGAAATACAGAGTGG
TGGTTAGCTGAGGTTAACGGGAAGAAGGGCTACGTTCCCTCCAATTATATCCGCAAAACCGAGT
ACACCTGAGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTG
TTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATATCCCC
TATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATC
TCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACAT
AAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAAT
TCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTG
CGACAATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGG
TAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCT
CTTCCGACCATCAAGCATTTTATCCGT

>T1462M-DNMBP

ATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGG
GCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCAC
CGACAAACAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGC
CTGGCAGTTCCCTACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAAA
ACGACGGCCAGTCTTAAGCTCGGGCCCC<u>AAATAATGATTTTATTTTGACTGATAGTGACCTGTT
CGTTGCAACAAATTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTT</u>
CGCCGCCACCATGGAGGCTGGCTCAGTGGTTCGAGCCATTTTTGACTTCTGCCCTAGCGTATCA
GAAGAACTGCCGCTCTTTGTGGGAGATATTATTGAGGTGCTGGCAGTGGTGGATGAATTTTGGC
TTCTAGGAAAGAAGGAGGATGTTACAGGACAATTCCCCAGCAGTTTTGTGGAAATTGTGACCAT
TCCCAGTCTAAAAGAGGGAGAGAGGCTGTTTGTGTGCATTTGTGAATTCACATCCCAAGAGTTG
GATAATCTTCCCCTCCATCGAGGTGACCTGGTGATTCTCGATGGCATTCCCACTGCAGGCTGGC
TGCAGGGCCGAAGCTGCTGGGGCGCACGGGGCTTCTTCCCATCTTCATGTGTCCGCGAGCTCTG
CCTCTCCTCACAGAGCCGGCAGTGGCACTCCCAGAGCGCCCTGTTTCAGATTCCGGAATATTCC
ATGGGACAAGCCCGGGCCCTAATGGGGCTTTCTGCTCAGCTGGATGAAGAGCTGGACTTCAGGG
AAGGGGATGTGATCACCATTATTGGAGTTCCTGAACCAGGCTGGTTTGAAGGGGAGTTAGAGGG
CCGAAGAGGCATTTTTCCAGAAGGTTTTGTAGAGCTGTTGGGGCCCCTGAGGACTGTGGATGAG
TCAGTAAGTTCTGGAAATCAAGATGACTGCATTGTTAATGGTGAAGTAGATACCCCTGTAGGAG
AAGAAGAGATAGGGCCGGATGAGGATGAGGAGGAGCCAGGGACCTATGGGGTCGCCCTGTACAG
ATTCCAAGCCCTGGAGCCAAATGAGCTGGATTTCGAGGTCGGGGATAAAATCCGAATTCTGGCG
ACCTTGGAAGATGGCTGGCTGGAAGGATCCCTGAAGGGCAGGACAGGCATCTTTCCTTACCGGT
TTGTGAAATTATGTCCTGACACACGGGTGGAGGAAACCATGGCTCTGCCCCAGGAAGGCAGCCT
TGCCAGGATCCCGGAAACTTCTTTGGATTGTTTGGAGAACACCTTAGGAGTAGAGGAACAAAGA
CATGAAACCAGTGACCATGAGGCCGAGGAGCCTGACTGCATTATTTCTGAAGCACCCACTTCTC
CCCTCGGTCATCTGACTTCAGAGTATGACACAGACAGAAACTCTTATCAGGACGAGGACACCGC
AGGAGGGCCCCCGAGAAGCCCAGGCGTGGAGTGGGAAATGCCTCTTGCCACAGACTCTCCCACA
TCTGACCCTACAGAAGTAGTCAATGGTATTTCCTCCCAACCTCAGGTCCCTTTTCATCCCAACT
TGCAGAAAAGCCAGTATTATTCTACAGTGGGAGGGAGCCACCCGCACTCAGAACAGTACCCCGA
CCTTCTTCCCCTAGAAGCAAGGACTAGAGACTATGCCAGCCTACCTCCCAAAAGAATGTATTCC
CAGCTAAAAACTCTTCAGAAGCCAGTGCTCCCTCTTTACAGGGGCTCTTCTGTTTCAGCTTCAA
GGGTAGTCAAACCCAGACAATCAAGTCCTCAGCTCCACAACCTAGCAAGTTATACTAAAAAGCA
CCACACGTCCAGTGTTTACTCCATCTCAGAGAGATTGGAGATGAAGCCTGGTCCGCAAGCCCAA
GGGCTTGTTATGGAAGCAGCAACACATTCACAGGGAGACGGCAGCACTGACCTGGACTCGAAGC
TGACACAACAGCTGATCGAGTTTGAGAAGAGCTTGGCAGGGCCCGGCACAGAGCCAGATAAAAT
TTTACGCCACTTTTCAATCATGGACTTTAACTCTGAGAAGGATATTGTCCGAGGTTCCTCAAAG
TTAATCACCGAGCAGGAGCTGCCGGAAAGGAGAAAGGCCCTAAGGCCACCGCCACCTCGTCCCT
GTACTCCGGTATCCACTTCTCCCCATTTGCTGGTTGACCAGAACCTAAAACCTGCACCACCCTT
GGTGGTGCGACCCTCTCGCCCAGCTCCCCTGCCTCCCTCAGCACAGCAGAGAACGAATGCGGTA
TCCCCCAAGCTCCTATCTCGACACCGTCCTACCTGTGAGACCTTAGAAAAGGAGGGCCCTGGTC
ATATGGGAAGGAGTCTGGACCAGACCTCCCCATGCCCCTTAGTGCTGGTGAGGATTGAGGAAAT
GGAGCGGGACTTGGATATGTACAGTAGAGCTCAAGAAGAGCTAAACCTCATGCTGGAGGAGAAG
CAGGATGAATCATCAAGAGCAGAGACCCTCGAGGATCTCAAGTTCTGTGAAAGTAACATTGAAA
GTTTGAATATGGAACTCCAGCAACTAAGAGAAATGACGCTCCTCTCCTCCCAGTCTTCATCACT
GGTGGCCCCTTCTGGGTCTGTGTCTGCCGAAAATCCAGAGCAGAGGATGCTGGAGAAGAGAGCC
AAGGTCATAGAAGAACTTCTTCAGACAGAAAGAGACTACATTCGGGATCTGGAAATGTGTATTG
AGCGGATCATGGTACCCATGCAGCAGGCACAGGTACCAAACATTGATTTTGAGGGACTTTTTGG
AAATATGCAGATGGTGATTAAGGTCTCGAAGCAATTATTGGCTGCTCTGGAAATCAGCGATGCT

GTAGGACCTGTGTTTCTTGGTCACCGGGATGAGCTTGAGGGAACATACAAGATTTACTGCCAGA
ATCATGATGAGGCCATTGCGCTGCTTGAAATCTACGAGAAGGATGAGAAGATCCAGAAGCATCT
TCAGGACTCCTTGGCAGATCTGAAGAGCCTATACAACGAATGGGGATGCACAAATTATATTAAC
CTGGGCTCCTTCCTCATCAAACCAGTACAGAGAGTAATGCGTTACCCGCTGTTGCTAATGGAGT
TGCTGAATTCCACCCCAGAATCCCACCCCGATAAAGTGCCTTTAACCAATGCAGTCCTTGCGGT
CAAGGAAATCAACGTTAACATTAATGAATATAAACGGCGAAAGGACCTGGTCCTCAAGTACCGT
AAGGGTGATGAAGATAGCCTTATGGAGAAAATTTCCAAACTGAACATCCACTCCATCATCAAGA
AATCCAACCGAGTTAGCAGTCACCTGAAGCATCTCACTGGCTTTGCTCCTCAGATAAAAGATGA
AGTATTTGAAGAAACAGAAAAAAACTTCCGAATGCAAGAAGATTGATTAAGTCTTTTATCTGA
GACCTGTCTCTCTACCTCCAGCACATCCGGGAGTCCGCATGTGTGAAAGTGGTGGCTGCTGTGA
GCATGTGGGATGTGTGCATGGAGAGAGGACACCGGGACCTGGAGCAGTTTGAGAGGGTGCATCG
CTACATCAGTGACCAGCTCTTCACAAACTTTAAGGAGAGGACAGAGCGGCTTGTCATCTCCCCC
TTAAATCAGTTACTGAGCATGTTTACAGGGCCCCATAAGCTGGTACAGAAACGCTTTGACAAGC
TCCTGGACTTCTATAACTGTACAGAACGGGCAGAAAAGCTAAAGGACAAGAAGACCCTGGAGGA
GCTGCAGTCGGCCCGGAACAACTATGAGGCCCTGAATGCACAGCTGCTGGATGAGCTGCCCAAG
TTCCACCAGTACGCCCAGGGCCTCTTCACCAACTGTGTCCACGGCTATGCTGAAGCCCACTGTG
ACTTTGTGCACCAGGCTCTGGAGCAATTAAAGCCACTGCTTTCGTTACTCAAAGTGGCTGGCAG
AGAGGGAAACCTTATTGCCATCTTCCACGAAGAGCACAGCAGAGTTCTGCAGCAACTCCAGGTT
TTTACCTTCTTCCCGGAGTCTCTTCCAGCTACCAAGAAGCCATTTGAGAGGAAAACCATTGACC
GCCAGTCTGCTCGAAAGCCACTCCTGGGCCTGCCAAGTTACATGCTACAGTCAGAAGAACTCCG
GGCCTCCCTCCTGGCCAGGTATCCCCCTGAAAAACTCTTCCAGGCAGAACGGAACTTCAATGCT
GCTCAAGACTTGGATGTCTCACTTTTGGAAGGTGACCTGGTGGGTGTGATTAAGAAAAAAGACC
CCATGGGCAGCCAGAACCGCTGGCTGATTGACAATGGAGTCACCAAAGGCTTCGTGTACAGCTC
TTTCCTAAAGCCCTACAATCCTCGCCGCAGCCACTCCGATGCCTCCGTGGGTAGCCACTCCTCC
ACAGAGTCTGAGCACGGCAGCTCCTCCCCCAGGTTCCCACGCCAGAACAGCGGCAGCACCCTGA
CCTTCAACCCCAGCAGCATGGCTGTATCCTTTACCTCGGGGTCTTGCCAGAAGCAGCCTCAAGA
TGCATCTCCTCCGCCAAAAGAATGTGACCAAGGAACTCTCAGTGCATCCCTAAATCCGAGTAAT
TCAGAGAGTAGTCCTTCCAGATGCCCTTCAGACCCAGACTCCACCTCCCAGCCAAGGTCAGGGG
ACTCTGCAGATGTAGCTAGAGATGTAAAGCAACCCACTGCCA**T**GCCGAGGAGCTACCGGAACTT
CAGGCATCCAGAAATAGTTGGCTACTCCGTACCAGGACGAAATGGGCAAAGTCAAGACCTCGTC
AAAGGATGTGCAAGAACAGCCCAGGCTCCGGAAGACAGAAGTACAGAGCCAGATGGCAGTGAGG
CAGAAGGCAACCAGGTCTATTTTGCTGTCTACACCTTCAAGGCACGAAACCCAAATGAGCTGAG
CGTGTCAGCCAATCAGAAACTCAAGATCCTCGAGTTTAAAGATGTTACAGGAAATACAGAGTGG
TGGTTAGCTGAGGTTAACGGGAAGAAGGGCTACGTTCCCTCCAATTATATCCGCAAAACCGAGT
ACACCTGAG<u>ACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTG</u>
<u>TTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTG</u>CCATCCAGCTGATATCCCC
TATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATC
TCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACAT
AAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGGCCGCGATTAAAT
TCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTG
CGACAATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGG
TAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCT
CTTCCGACCATCAAGCATTTTATCCGT

Appendix B: Optimization of DNMBP Co-immunoprecipitation.

Table B.1: List of lysis buffers used for co-immunoprecipitations.

Varying concentrations of components were used to optimize cell lysis for co-immunoprecipitation. Common components included Tris-HCl (pH buffer system), NaCl (ionic strength/protein complex stabilization), Glycerol (stabilization), EDTA (protease inhibitor), and Detergent (solubilisation of cell membrane). A protease inhibitor cocktail (Sigma-Aldrich).

Lysis Buffer	Tris-HCl	NaCl	Glycerol	EDTA	Detergent	Other
ND-1	20mM (pH 8.0)	420mM	5%	2mM	1% IGEPAL	
ND-2	20mM (pH 8.0)	137mM	-	3mM	1% IGEPAL	
ND-3	20mM (pH 8.0)	150mM	10%	2mM	-	
ND-4	20mM (pH 8.0)	50mM	5%	2mM	1% IGEPAL	
ND-5	20mM (pH 8.0)	100mM	5%	2mM	1% IGEPAL	
ND-6	20mM (pH 8.0)	150mM	5%	2mM	1% IGEPAL	
ND-7	20mM (pH 8.0)	200mM	5%	2mM	1% IGEPAL	
ND-8	20mM (pH 8.0)	250mM	5%	2mM	1% IGEPAL	
ND-9	20mM (pH 8.0)	350mM	5%	2mM	1% IGEPAL	
ND-10 (Otani et al., 2006)	50mM (pH 7.5)	150mM	10%	1mM	0.5% IGEPAL	
ND-11	25mM (pH 7.5)	150mM	5%	1mM	1% IGEPAL	
PBS-TX (pH 7.4)	-	137mM	-	-	1% Triton X-100	2.7 mM KCl, 8 mM Na$_2$HPO$_4$, 2 mM KH$_2$PO$_4$
RIPA	50mM (pH 8.0)	150mM	-	1mM	1% IGEPAL	0.5% Sodium Deoxycholate, 0.1% SDS
Pierce™ Lysis Buffer	25mM (pH 7.4)	150mM	5%	1mM	1% NP-40	

Notes: Abbreviations: ND: non-denaturing, TX: Triton X-100, SDS: Sodium dodecyl sulfate

Table B.2: Co-immunoprecipitation reactions conducted. Co-immunoprecipitations were conducted on HEK293 or U-2 OS cells. Cells were lysed in the indicated lysis buffer and the co-immunoprecipitation reaction was carried out. Each reaction was attempted at least two times. Reactions recorded here were conducted by immunoprecipitating VASP and detecting for DNMBP. A successful result indicates immunoprecipitation led to co-immunoprecipitation of partner. Some reactions were repeated with immunoprecipitation of MENA.

Reaction	Lysis Buffer	Cells (transfected?)	IP buffer	Incubation	Washes	Results
Rxn1	ND-1	HEK293 (NT)	ND-1 with 150mM NaCl	Overnight	5X with ND-1 with 0.1% Triton X-100	Successful 2/10 attempts
Rxn2	ND-1	HEK293 (WT DNMBP)	ND-1 with 84mM NaCl	Overnight	5X with ND-1 with no detergent	Unsuccessful
Rxn3	ND-1	HEK293 or U-2 OS (WT DNMBP and VASP)	ND-1 with 84mM NaCl	Overnight	5X with ND-1 with 0.1% Triton X-100	Successful 2/7 attempts
Rxn4	ND-1	HEK293 (WT DNMBP and VASP)	ND-1 with 100mM NaCl	Overnight	5X with IP Buffer with 0.1% Triton X-100	Successful 1/3 attempts
Rxn5	ND-2	U-2 OS (WT DNMBP and VASP)	ND-2	Overnight	5X with ND-2	Unsuccessful
Rxn6	ND-2	U-2 OS (WT DNMBP and VASP)	ND-2 with 100mM NaCl	Overnight	5X with ND-2	Unsuccessful
Rxn7	ND-3	U-2 OS (WT DNMBP and VASP)	ND-3	Overnight	5X with ND-3	Unsuccessful

Table B.2 continued

Rxn8	ND-3	U-2 OS (WT DNMBP and VASP)	ND-3 with 100mM NaCl	Overnight	5X with ND-3	Unsuccessful
Rxn9	PBS-TX	U-2 OS (WT DNMBP and VASP)	PBS-TX	Overnight	3X with PBS-TX	Unsuccessful
Rxn10	RIPA	Fixed HEK293 (WT DNMBP and VASP)	RIPA	Overnight	5X with RIPA	Unsuccessful
Rxn11	ND-4	U-2 OS (WT DNMBP and VASP)	ND-4	Overnight	5X with ND-4	Unsuccessful
Rxn12	ND-5	U-2 OS (WT DNMBP and VASP)	ND-5	Overnight	5X with ND-5	Unsuccessful
Rxn13	ND-6	U-2 OS or HEK293 (WT DNMBP and VASP)	ND-6	Overnight	5X with ND-6	Unsuccessful
Rxn14	ND-7	HEK293 (WT DNMBP and VASP)	ND-7	Overnight	5X with ND-7	Unsuccessful
Rxn15	ND-8	HEK293 (WT DNMBP and VASP)	ND-8	Overnight	5X with ND-8	Unsuccessful
Rxn16	ND-9	HEK293 (WT DNMBP and VASP)	ND-9 with 84mM NaCl	Overnight	5X with ND-8 and 84mM with NaCl	Unsuccessful
Rxn17	ND-10	HEK293 (WT DNMBP and VASP)	ND-10	3 hours or overnight	3X with ND-10	Unsuccessful
Rxn18	ND-11	HEK293 (WT DNMBP and VASP)	ND-11	2 hours	3X with ND-11	Unsuccessful
Pierce™ Lysis	Pierce Lysis	U-2 OS or HEK293 (WT DNMBP and VASP)	Pierce Lysis	2 hours or Overnight	3X in Pierce™ Lysis	Unsuccessful

Notes:
Pierce™ IP Lysis Buffer (Pre-made) from Thermo Scientific
Abbreviations: Rxn: reaction, WT: wildtype, NT: non-transfected, 5X/3X: 5 times/3 times