Functional analysis of a coding variant in ZC3HC1 at 7q32.2 associated with protection against Coronary Artery Disease (CAD)

Tara A. Linseman

A thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for the MSc degree in Biochemistry, Specialisation in Human and Molecular Genetics

Biochemistry, Microbiology and Immunology
Faculty of Medicine
University of Ottawa

Abstract

Functional analysis of a coding variant in ZC3HC1 at 7q32.2 associated with protection against Coronary Artery Disease (CAD)

By Tara A. Linseman

Coronary artery disease (CAD), characterized by the narrowing of coronary arteries through the complex manifestation and development of atherosclerosis, is a complex disease and one of the leading causes of death worldwide. Both genetic and environmental factors are believed to contribute equally to the risk of CAD. Recently, a study identified a non-synonymous coding variant, rs11556924, (MAF, 0.38) in *ZC3HC1* associated with protection against CAD ($p= 9.8x10^{-18}$; OR= 0.90). NIPA, encoded by *ZC3HC1*, is a characterized F-Box protein and regulator of cell cycle. Since the amino acid change (Arg363His) is in a conserved region of NIPA and is predicted to have functional effects (Polyphen-2), this study aimed at understanding the functional implications of this amino acid change on NIPA and cell cycle regulation. Here we are able to effectively show a) allele specific differences in mRNA expression in whole blood, b) a slight structural difference between NIPA363Arg and NIPA363His variants, c) proliferation rates of NIPA363Arg expressing cells were significantly increased, and d) phosphorylation of a critical serine residue in close proximity to aa.363 is not statistically different between the two variants. These results suggest that rs11556924 plays a direct role in development of CAD through its disruption of cell cycle regulation and NIPA mRNA availability. This study is the first to identify a molecular basis for the association of rs11556924 to CAD development.

Acknowledgements

First to my thesis supervisor, Dr. Ruth McPherson: thank you for your constant guidance, support and direction on my project. You gave me the freedom to entertain my curiosity in the field of cardiovascular genetics and allowed me to grow as a researcher.

To Dr. Alexandre Stewart, Dr. Patrick Burgon and Dr. Adam Rudner: thank you for your contributions and guidance and serving as members of my thesis advisory committee.

To the rest of the McPherson lab, especially Sebastien Soubeyrand, Paulina Lau, Adam Turner and Amy Martinuk: thank you for not only giving me advice and guidance on my project, but making science fun and my time in the lab unforgettable.

To my best friends Ami, Leah and Justyn: your constant support throughout these past two years was the key to my sanity. From the long talks to quick visits, you kept me grounded and gave me a push when I needed it.

Lastly to my parents, Mike and Wendy, my brother Brett and my sister Lauren: without your constant support and encouragement over the years I would not be where I am today. Thank you for always believing in me and allowing me to pursue my dreams.

Table of Contents

Abbreviations

ALCL	Anaplastic Large-cell Lymphoma
ALK	Anaplastic Lyphoma Kinase
APC/C	Aanaphase Promoting Complex/cyclosome
CAD	Coronary Artery Disease
CDK	cyclin dependant kinase 1
CI	confidence interval
cIMT	carotid-intima media thickness values
CIP	Calf intestinal phosphatase
Co-IP	Co-Immunoprecipitation
Cul1	cullin protein 1
CVD	Cardiovascular disease
ECM	extracellular matrix components
ERK2	Extracellular Signal-regulated Kinase 2
GWAS	Genome wide association studies
iPSC	induced pluripotent stem cell
K. lactis	*Kluyveromyces lactis*
LD	Linkage disequilibrium
LDL-c	low-density lipoprotein cholesterol
MAF	minor allele frequency
MI	myocardial infarction
NIPA-ALK	Nuclear Interacting Partner of ALK
NPM	nucleophosmin-ALK
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase chain reaction
PPIA	Peptidylprolyl isomerase A
RIPA	Radioimmunoprecipitation assay buffer
Roc1	RING-finger protein 1

SCF	Skp1, cullin, F-Box protein
SD	standard deviation
SE	standard error
Skp1	S-phase kinase-associated protein 1
SNP	single nucleotide polymorphisms
VSMC	vascular smooth muscle cells

List of Figures

List of Tables

1 Introduction

1.1 Coronary Artery Disease

Cardiovascular disease (CVD) is the number one cause of death worldwide, contributing to about 31% of all deaths[1]. Coronary artery disease (CAD), in particular, contributes to about 42% of CVD related deaths (7.4 million deaths worldwide from CAD in 2012)[1]. CAD results from the narrowing of coronary arteries through the complex manifestation and development of atherosclerosis.

Atherosclerosis is characterized by thickening of arterial walls and the development of cholesterol rich lesions in the arterial intima, resulting in restricted blood flow though the lumen[2]. Prolonged blood flow restriction to major organs can result is both acute and chronic organ damage. If the cholesterol rich lesions or atherosclerotic plaque become unstable, they can rupture releasing thrombotic material into the lumen[3]. Thrombotic material may become lodged in arterial branches blocking blood flow to critical organs such as the heart or brain causing a myocardial infarction (MI) or stroke respectively[4].

Atherosclerotic plaque development is often a long and complex process. Fatty streaks are the beginning stages of plaque development and are often found in the arteries of children and adolescents. The fatty streaks continue to develop over their lifetime eventually maturing into advanced atherosclerotic plaques. The formation of fatty streaks begins with the build-up of low-density lipoprotein cholesterol (LDL-c) in the intima space where they are oxidized and modified

triggering an inflammatory response[5]. Macrophages respond to the inflammatory signals, migrating to the location of LDL-c accumulation and begin phagocytosis. Unregulated and proliferative macrophage phagocytosis can eventually result in foam cell formation[6]. Foam cells are enlarged, cholesterol rich macrophages that are unable to completely breakdown cholesterol molecules. The accumulation of foam cells within the intima of arteries is what forms the basis of the fatty streaks[7]. However it is primarily the prolonged inflammatory state that causes the fatty streaks to progress into advanced lesions.

Many factors contribute to the chronic inflammatory state of atherosclerotic plaque, including buildup of foam cells, continuous uptake of cholesterol into the inflamed intima space, cholesterol crystallization and the improper clearance of apoptotic debris[8]. The prolonged inflammatory state of the inner plaque recruits vascular smooth muscle cells (VSMCs) to proliferate and migrate to the intima. Once there, VSMCs produce and promote generation of extracellular matrix components (ECM) which are used to develop a fibrous cap under the endothelium of the lumen[4]. The fibrous cap provides stability to the plaque and acts as a barrier for the lesion. If the cap becomes unstable or “vulnerable”, the plaque is at risk for rupture resulting in released thrombotic material causing an MI or stroke. Along with a thin fibrous cap, vulnerable plaques also have characteristically large necrotic cores and high inflammatory states[4].

Stability of the fibrous cap in advanced lesions is reliant upon the recruitment and proliferation of VSMC[9]. Growth factors and pro-inflammatory cytokines induce VSMC proliferation in the media and promote their migration to the intima, resulting in the formation of the protective cap[10]. Recent studies have added to the growing body of evidence suggesting that not only do VSMCs proliferate during atherosclerotic plaque development, but macrophages proliferate at these sites as well. Local macrophage proliferation can maintain accumulation in inflammatory tissues and is aberrant in advance atherosclerotic lesions[11]. It has recently been shown by Robbins *et al.* (2013) [12] that macrophage proliferation contributes to over 80% of local macrophage accumulation. Through the use of 5-Fu inhibitor of cell cycle progression, Robbins *et al.* (2013) were able to decrease the proliferation of macrophages ad ultimately show progression of atherosclerosis in a CAD mouse model[12]. In another study Tang *et al.* (2015) were able to efficiently reduce plaque inflammation in apolipoprotein E deficient mice (*Apoe*$^{-/-}$) through local inhibition of plaque macrophage proliferation[13]. These studies add to the established theory that proliferation, particularly VSMCs and plaque macrophages, is a critical component of atherosclerosis development.

1.2 Genome-wide Association Studies

Both genetic and environmental risk factors contribute to the development of CAD. Although environmental risks such as smoking, diet and exercise all contribute to the development CAD, it has been well established that approximately 40-60% of

CAD risk can be attributed to genetic factors[14]. Understanding the contribution of genetic burden on CAD risk is a complex process.

Many single gene mutations have been identified through the use of family linkage studies and trio sequencing comparisons. The majority of these mutations are transmitted through Mendelian inheritance patterns and often carry a large genetic burden[15]. Despite their immense benefit for detecting causes of Mendelian-inherited diseases, linkage and family studies are significantly limited when it comes to understanding the genetic risk of complex diseases such as CAD[15]. However, in recent years many tools were developed in order to better understand and identify genetic risk burden for complex diseases.

Genome-wide association studies (GWAS), made possible by the development of technologies such as microarray chips which probe for single nucleotide polymorphisms (SNPs) spanning the genome, are based off the common disease-common variant hypothesis; the idea that common diseases, such as CAD, are caused by the accumulation and compounding effects of many common genetic variants (minor allele frequency (MAF) ≥ 5%), and not by single more rare variants (MAF <1%)[16]. Under this hypothesis, many variants with low penetrance are predicted to synergistically contribute to overall development of complex diseases. GWAS aim to identify genetic variants contributing to complex diseases by comparing genetic variation between healthy control individuals and disease patients[17]. In order to identify variants GWAS rely heavily on arrays which capture large portions of

common variation in the human genome. These arrays contain ~1 million SNP probes[16]. With the human genome containing close to 10 million SNPs, arrays are limited in their ability to assess complete variation. Thus, the arrays contain tag SNPs — location markers to identify regions of the genome which are associated with disease[18]. With this platform, GWAS use hypothesis-free association testing between SNPs across the genome and disease of interest[16]. Due to large sample sizes and millions of statistical association testing, corrected thresholds must be used in order to determine true significant associations and reduce false positives. Using the Bonferroni correction for multiple testing, a corrected p-value of $p<5.0x10^{-8}$ is used as the threshold of genome wide significant associations.

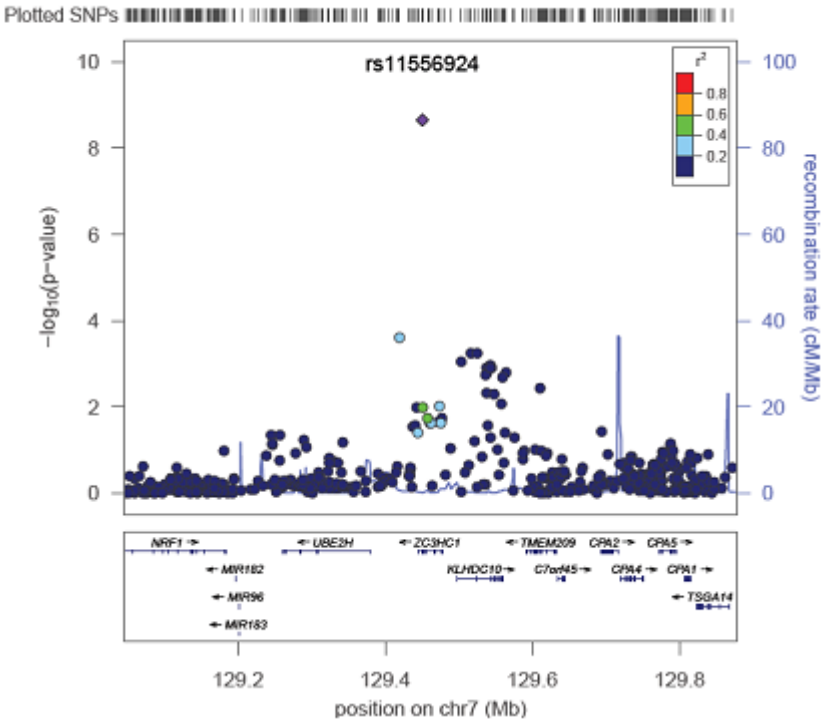
SNPs found to be significant through association testing may not directly affect the development of disease, but often identify regions of the genome that play a role in disease development. This is often the case in GWAS, with most of the significant SNPs falling within non-coding regions of the genome. Both statistical and biological methods have been developed in order to identify these regions of significance using SNPs[19]. Linkage disequilibrium (LD) is the concept that two markers (SNPs) found geographically close enough are statistically more likely to be inherited together than to be associated through genetic recombination[16,20]. SNPs in LD with one another therefore can identify loci (regions of the genome) which may harbour genes or regions of interest associated with disease development. Sometimes a SNP found to be associated with disease lies within a protein-coding region. Although rare, protein-coding SNPs are ideal for functional analysis due to the relative ease of evaluating

gene expression and protein implications rather than trying to understand the relevance of non-coding regions.

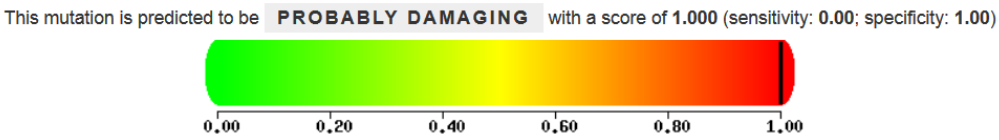
1.3 ZC3HC1 Locus for CAD

In a recent GWAS meta-analysis[21], 13 new CAD susceptibility loci were identified. One of the top hits in this study and the only coding SNP was rs11556924 found on chromosome 7 at q32.2. This SNP encodes a non-synonymous mutation in ZC3HC1 (p.Arg363His). ZC3HC1 encodes an E3 ligase NIPA (Nuclear Interacting Partner of Anaplastic Lyphoma Kinase). Using 1000 Genomes data to analyze the LD structure of the region, no other variants were identified with greater biological plausibility or effect size (**Figure 1a**). The minor allele (T) frequency of rs11556924 is 0.38 and an odds ratio for CAD of 0.90 ($p=2.4x\ 10^{-17}$). This region of the genome is found to be highly conserved (**Figure 1c**). Through the use of Polyphen2, a prediction software for functional implications of variants, the p.Arg363His variant was predicted to be damaging (**Figure 1b**)[22].

A)


Plotted SNPs
rs11556924
r²
0.8
0.6
0.4
0.2
−log10(p−value)
recombination rate (cM/Mb)
10
8
6
4
2
0
100
80
60
40
20
0
NRF1
UBE2H
2C3HC1
TMEM309
CPA2
CPA5
MIR182
KLHDC10
C7orf45
CPA4
CPA1
MIR96
TSGA14
MIR183
129.2
129.4
129.6
129.8
position on chr7 (Mb)

B)


This mutation is predicted to be PROBABLY DAMAGING with a score of 1.000 (sensitivity: 0.00; specificity: 1.00)
0.00
0.20
0.40
0.60
0.80
1.00

C)


354
359
363
370
395
Human 352 T R S W D S S S P V D R P E P E A A S P T T R T R P V T R S M G T G D T P G L - - E V P S S P L R K A K R 402
M. musculus 351 T R S W E S S S P V D R P E L E A A S P T T R S R P V T R S M G T G D S A G V - - E V P S S P L R R T K R 401
X. laevis 332 T R S S D L L S P A D - - - - - - - S E A V R N R P V T R S M G Q G E N T G L G N E L H S S P H R R A K R 377

Figure 1. Characteristics of rs11556924 locus. A regional plot of rs11556924 locus depicts other SNPs in the immediate region surrounding the lead SNP (A). No other SNP surrounding rs11556924 reached genome wide significance and is not in LD (as seen by the r^2 labelling). The amino acid change from an argenine to histidine is predicted by Polyphen2 to be damaging[22] (B). NIPA amino acid conservation across three species with the location of critical serine residues (residues marked with *) and R363H location (marked in red box) (C).

Since the initial report of rs11556924 association with CAD[21], two other studies have been done to evaluate the implications of this SNP on CAD associated traits in other cohorts. In the first study, López-Mejías *et al.* (2013) [23] showed that Northern Spaniards with rheumatoid arthritis and homozygous for the minor allele (TT), had significantly higher carotid-intima media thickness values (cIMT: a measurement for determining subclinical atherosclerosis) than patients homozygous for the CC allele. The authors indicate that their findings further support the role of rs1556924 in development of CAD. This is despite the fact that their results are in direct contradiction to reports that the C allele is the risk allele for CAD, reported in the Schunkert *et al* (2011) paper.

In a more recent study, the association between rs11556924 and hypertension was investigated in 50 year-old Finnish patients[24]. The study involved clinical diagnosed hypertension patients enrolled in the TAMRISK study — a long-term population health study in Finland. The authors showed that a CC genotype at rs11556924 locus was associated with hypertension whereas a CT and TT genotype showed no statistically significant associations. The authors therefore suggest that the rs11556924 risk allele (C), as identified by Schunkert *et al* (2011), also plays a role in hypertension.

Although both these studies are limited by statistical power and seek further genetic associations, they do provide small insights into the greater role of

rs11556924 in the development of CAD. To date no studies have been published on the molecular implications of rs11556924 or NIPA in relation to CAD development.

1.4 Nuclear Interacting Partner of Anaplastic Lymphoma Kinase (NIPA)

Anaplastic Lymphoma Kinase (ALK) is a tyrosine kinase found to be activated in Anaplastic Large-cell Lymphoma (ALCL), a form of Non-Hodgkin's Lymphoma. It is well known that chromosomal translocations in cancer cells, more specifically Lymphoma, can cause overexpression/aberrant activity of tyrosine kinases through abnormal auto-phosphorylation[25]. These translocations/mutations may cause downstream constitutive expression of signaling pathways and accelerate cell proliferation among other things. Oncogenic ALK fusion proteins, specifically nucleophosmin (NPM)-ALK, which is a result of t(2;5)(p233:q35) translocation, have been shown to increase oncogenic properties of cells during progression of lymphoma[25]. Identifying downstream interacting partners of NPM-ALK could lead to further understanding of its role during oncogenisis. Basserman *et al* (2003) set up a yeast-two hybrid screen targeted at detecting downstream interactors of NPM-ALK, effectively identifying NIPA (Nuclear Interacting Partner of ALK).

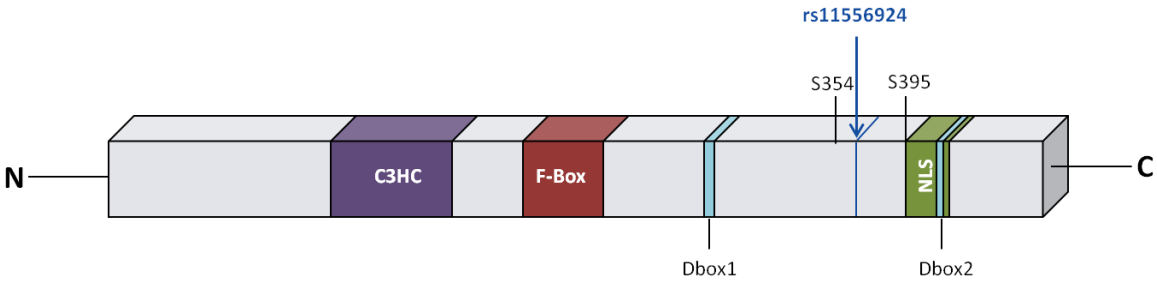

rs11556924
S354
S395
N
C3HC
F-Box
NLS
C
Dbox1
Dbox2

Figure 2. Characterized protein motifs within NIPA. NIPA is 60 kDa protein with multiple functional motifs that lie within it. As an E3 ligase, NIPA contains an F-Box binding domain near the N-terminus (amino acids 170-210)[26] and a nuclear localization signal between amino acids 396 and 402 located at the C-terminus[25] It also contains 2 D-Box motifs critical serine residues which are phosphorylated.

NIPA, encoded by *ZC3HC1* (7q32.2), is a 60 kDa protein which shares 85% amino acid homology with its mouse homologue mNIPA[25]. As an E3 ligase, NIPA

contains an F-Box binding domain near the N-terminus (amino acids 170-210)[26] and a nuclear localization signal between amino acids 396 and 402 located at the C-terminus[25] (**Figure 2**). The nuclear localization signal confines the expression of NIPA strictly to the nucleus[26]. After RNA expression analysis, it was found that NIPA is broadly expressed in human tissues with highest expression in the heart, skeletal muscle and testis[25].

1.5 NIPA's role in the regulation of cyclin B1 during cell cycle.

Cell cycle progression is a highly organized and regulated process. Levels of specific proteins are tightly regulated in order to either maintain cells at a particular stage or push them through into subsequent stages. Like many of these cell cycle proteins, NIPA levels oscillate throughout the course of cell cycle[27]. In growth-arrested cells ($G_{0/1}$) NIPA is found at minimal levels reaching highest concentration during S and G_2/M stages[26]. As cells progress into G_2 phase and eventually reach the G_2/M boundary, NIPA becomes phosphorylated at two critical serine residues (S354 & S359) by ERK2 (Extracellular Signal-Regulated Kinase 2)[26,28]. The phosphorylation of these serine residues and one other (S395) dictate the activity of NIPA and ultimately the timing of mitosis entry by the cell.

Entry into mitosis, like most cell cycle checkpoints, is very well regulated and is tightly controlled. Activation of the cyclin B1/CDK1 complex (CDK: cyclin dependant kinase 1), often referred to as the Mitosis promoting factor (MPF), and its accumulation in the nucleus triggers cells to enter mitosis[29]. During interphase cyclin

B1 is cytoplasmic, thus inactive[30]. It is only during mitotic progression that cyclin B1 is activated via entry into the nucleus[30–32]. If cyclin B1 prematurely enters into the nucleus, mitotic entry is initiated too early potentially perpetuating adverse effects elsewhere during mitosis[32]. NIPA plays a essential role in regulating the levels of cyclinB1 in the nucleus.

As a characterized F-Box protein, NIPA is a key component of the SCF (Skp1, Cullin, F-Box protein)-type E3 ligase, SCF^{NIPA} [26]. More specifically NIPA is able to bind to Skp1 (S-phase kinase-associated protein 1) of SCF^{NIPA} through NIPA's F-Box motif (amino acids 170-210). Through Co-Immunoprecipitaiton (Co-IP) experiments, Basserman *et al* was able to show that NIPA co-purified with Roc1 (RING-finger protein 1), Skp1 and Cul1 (cullin protein 1) which are the main SCF complex subunits. Furthermore they were able to effectively show that NIPA binds to nuclear cyclin B1 during interphase, but not during G_2/M transition, strongly suggesting that NIPA and nuclear cyclin B1 only interact when SCF^{NIPA} is intact[26].

NIPA, as an F-Box protein, plays a crucial role in the formation of the SCF^{NIPA} complex. However the SCF complex is not specific to NIPA and degradation of cyclin B1. The SCF complex is a multi-functional ubiquitin ligase protein complex that has the ability to specifically target a wide array of proteins often involved in cell cycle[26]. The complex consists of a cullin protein (cul1-5, cul7) that acts as a scaffolding protein to hold Skp1 on one end of the scaffold and a RING protein and E2 ligase on the other[33,34]. This complex organization allows for a claw-like conformation change,

permitting the RING/E2 ligase to add ubiquitin molecules to the protein recruited by Skp1[35]. SCF complex's ability to target an array of proteins is facilitated through the interchangeable F-Box proteins. This family of proteins contain a unique F-Box domain that allows them to bind directly to Skp1[36]. Furthermore, F-Box proteins contain a substrate-specific binding domain directly carboxyl-terminal to the F-Box domain[37]. This domain is responsible for targeting proteins for ubiquination and ultimately degradation. NIPA contains an F-Box domain and has been shown to specifically target cyclin B1 for ubiquination.

Investigation into the interaction of SCF^{NIPA} and cyclin B1 found that ubiquination of cyclin B1 was restricted to interphase and more interestingly, only associated with the un-phosphorylated S354 form of NIPA[38]. Upon mutation of S354 ($NIPA^{S354A}$), $SCF^{NIPA\text{-}S354A}$ is constitutively active and targets nuclear cyclin B1 for ubiquination well into G_2/M[38]. This suggests that the timely phosphorylation of S354 plays a critical role in the assembly and activity of SCF^{NIPA}.

The interaction between NIPA and cyclin B1 extends past interphase. It was effectively shown that upon the deactivation of NIPA through ERK2 phosphorylation of S354 and S359, NIPA is further phosphorylated by the cyclin B1-Cdk1 complex[28]. This additional phosphorylation by the MPF is a secondary measure to ensure the complete inactivation of NIPA, ultimately preventing further degradation of cyclin B1.

1.6 Oscillation of NIPA throughout cell cycle.

Throughout cell cycle the levels of NIPA fluctuate with the help of many protein complexes. During late G2 phase, a phosphorylated form of NIPA begins to appear in the cell, with amounts peaking at the G2/M boundary. As the cell passes through M and progresses into G1 phase, phosphorylated NIPA seems to disappear, returning later in G1. Upon use of cycloheximide, von Klitzing *et al.* (2010) were able to show that the appearance of NIPA is late G1 was due to the synthesis of new NIPA protein rather than de-phosphorylation of existing NIPA[39]. They were also able to effectively demonstrate that phosphorylated NIPA is degraded by the APC/C^{cdh1} (anaphase promoting complex or cyclosome) through NIPA-cdch1 interactions and the stabilization of phosphorylated NIPA following cdch1 knock down[39]. Furthermore NIPA was found to be protected from APC/C^{cdh1} degradation through its interaction with skp1; this interaction prevents phosphorylation by ERK2[39].

Through the cumulative work of many groups, it has been clearly demonstrated that NIPA's activity during cell cycle is strictly regulated through its interactions with various complexes and the phosphorylation of its three critical serine residues (S354, S359 and S395). It can therefore be inferred that NIPA plays a crucial role during cell cycle progression. **Figure 3** summarizes the role of NIPA in the regulation of cyclin B1 and overall cell cycle progression.

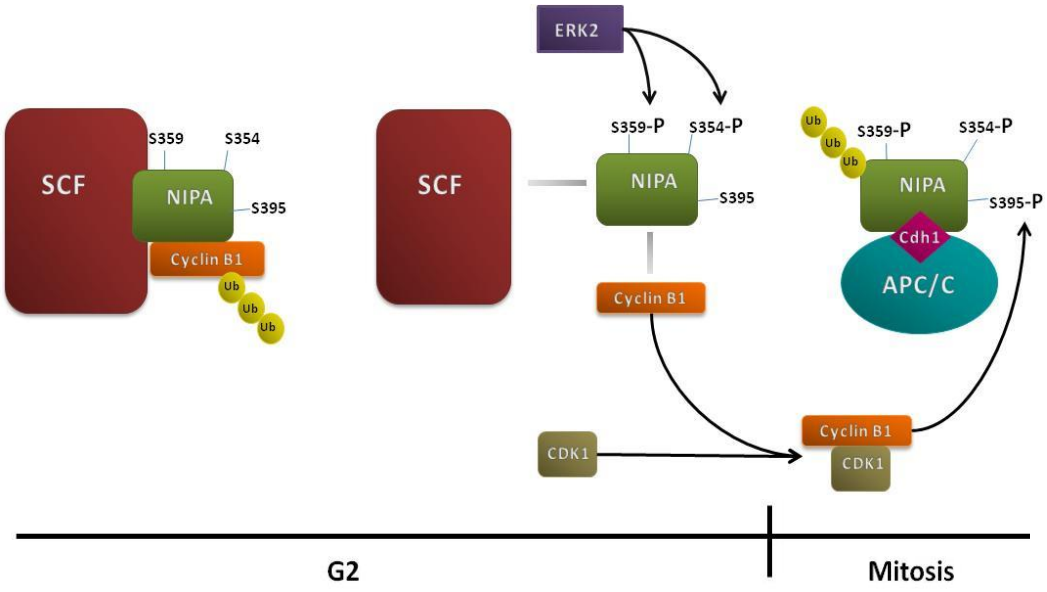

SCF
S359
S354
NIPA
S395
Cyclin B1
Ub
Ub
Ub
SCF
ERK2
S359-P
S354-P
NIPA
S395
Cyclin B1
CDK1
Ub
Ub
Ub
S359-P
S354-P
NIPA
S395-P
Cdh1
APC/C
Cyclin B1
CDK1
G2
Mitosis

Figure 3. Schematic representation of NIPA's role during cell cycle. During G2 NIPA is bound to the SCF complex via Skp1. SCF^{NIPA} specifically targets cyclinB1 for ubiquination and ultimately proteasomal degradation. During late G2/M transition NIPA is phosphorylated at S354 and S359 by ERK2, releasing it from the SCF complex. This enables cyclinB1 to form a complex with CDK1, which then further phosphorylates NIPA at S395. Phosphorylated NIPA is subsequently bound by the APC/C-cdh1 complex where it is ubiquinated, targeting it for proteasomal degradation. [image produced based on the experimental findings of the following papers[25,26,28,38–40]]

1.7 The effects of the R363H variant on the structure and activity of NIPA in the context of cell proliferation

Through the use of GWAS, a non-synonymous SNP rs11556924 was found to be associated with the development of CAD. This SNP is of particular importance due to its location in the protein coding region of a cell cycle regulatory protein NIPA. NIPA is a characterized F-Box protein and is involved in the regulation of cell cycle. Many proteins are involved in regulation of cell cycle and proliferation, each responsible for specific elements. It has been demonstrated previously in multiple studies that proliferation of both macrophages and VSMCs plays a crucial role in the development of advanced atherosclerotic plaque. The effect of rs11556924 (a non-synonymous SNP) on NIPA function and cell proliferation at a molecular level, have yet to be identified.

Hypothesis:

The Arg363His variant (rs11556924) associated with protection against CAD impedes the phosphorylation of critical serine residues on NIPA, preventing its disassociation from SCFNIPA complex; consequently cyclinB1 degradation is prolonged, ultimately leading to the attenuation of arterial smooth muscle cell proliferation and reduced atherosclerosis.

Specific Aims:

1) To determine the effect of the Arg363His variant on NIPA S354 phosphorylation. This will be achieved with the use of a phospho-specific antibody for NIPA- S354 to evaluate the ratios of phosphorylated to total NIPA between the variants.
2) To evaluate the effects of the NIPA-Arg363His variant on cell proliferation. Cell proliferation assays will be used to determine proliferative differences between cells expressing NIPA363Arg vs cells expressing NIPA363His.
3) To determine if the two variants are associated with different levels of mRNA expression. This will be assessed by q-PCR on RNA extracted from human whole blood samples using an allele-specific probe.
4) To evaluate further the difference observed in SDS-PAGE migration between variants to determine relevance to the structure and function of NIPA.

Together, these experiments will provide further understanding of the molecular implications of rs11556924 and its association with CAD development.

2 Materials and Methods

2.1 Cell Maintenance

HeLa and HEK-293T cells were cultured and maintained at 5% CO_2 in Dulbecco's Modified Eagle Medium (Life Technologies; DMEM), containing 4.5 g/L glucose, and supplemented with 10% Fetal Bovine Serum (FBS), 1% L-glutamine and 1% penicillin-streptomycin (Life Technologies). HeLa and HEK-293T cells were grown until they reached 90% confluency, at which time they were plated at a minimum density of 25%.

2.2 Plasmid Preparation and Bacterial Culture

The following expression vectors were made previously in the lab by Sebastien Souberyand (Research Associate): pCDNA Flag-NIPA363Arg, pCDNA Flag-NIPA363His, pEYFP FP-NIPA363Arg and pEYFP FP-NIPA363His. The empty pCDNA and pEYFP vectors used as controls are un-altered, commercially found plasmids (CloneTech). All plasmid descriptions can be found in Appendix A.

Chemically competent *E.coli* DHα (NEB) were transformed by heat shock at 42°C and were incubated on a rotary shaken at 225rpm set at 37°C for 1 hour. The transformation reactions were then plated on LB-agar plates containing 100 µg/mL ampicillin and incubated overnight at 37°C. For all Midi-preps of plasmids, single colonies were picked from the LB-agar plates and inoculated into 100mL of LB medium containing 100 µg/mL ampicillin, and incubated overnight at 37°C and

225rpm. Plasmids were purified using the Qiagen Plasmid Midi Kit (Qiagen) following the manufacturer's instructions.

2.3 Western Blots

The following is the general protocol used for western blots. Any changes or specifics to the protocol will be outlined in the subsequent experimental procedures.

RIPA buffer (150mM NaCl, 50mM Tris-HCl pH 8.0, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS (sodium dodecyl sulphate)) supplemented with 1x PhoSTOP phosphatase inhibitor tablets (Roche) and 1x cOmplete EDTA-Free protease inhibitor tablets (Roche) was added to each well. Cells were scraped and added to a 1.5mL microcentrifuge tube. Cell lysates were spun on the Eppendorf Centrifuge 54515C at 4°C for 5 minutes at 14,000rpm to pellet cell debris, after which the supernatant was transfer to a new clean 1.5mL microcentrifuge tube. Protein concentrations were determined by BCA protein assay (ThermoFisher Scientific).

5X SDS-PAGE (Polyacrylamide Gel Electrophoresis) loading buffer and β-mercaptoethanol were added to 20 µg of protein extract (unless otherwise specified) for a final concentration of 1X and 5% (v/v), respectively. Samples were boiled for 5 minutes at 96°C and loaded onto an 8% polyacrylamide gel (4% stacking). Electrophoresis was conducted for approximately 45 minutes at 200v. The proteins were transferred to a 0.45 µm nitrocellulose membrane (BioRad) using the Trans-

Blot® Turbo™ transfer pack and Trans-Blot® Turbo™ transfer system (BioRad) set at 2.5 constant amps, 25 constant volts for 7 minutes.

The nitrocellulose membrane was blocked using a 5% milk- 1x PBS solution for 20 minutes at room temperature with gentle shaking. Blocking solution was washed off with 1x PBS followed by incubation with the primary antibodies overnight at 4°C. The membrane was washed for 4x5mins with 1xPBS and incubated in the dark with secondary antibodies for 1 hour with gentle shaking. The membrane was again washed 4x5mins with 1x PBS. The signals were detected using Odyssey® Infrared Imaging System (LI-COR®) and analysis was done using Odyssey® Analysis Software V3.0.21 (LI-COR®).

2.3.1 Migration Shift between NIPA363Arg and NIPA363His

HeLa cells were seeded into 6-well Costar cell culture plates at 25% density. They were grown to 80% confluency in order to transiently transfect them. Cells were transfected using GenJet™ transfection reagent (SignaGen Laboratories). Each well was transfected with a total of 0.05 µg, 0.2 µg or 0.5 µg of plasmid DNA (80% pCDNA Flag-NIPA363Arg or Flag-NIPA363His; 20% pEYFP). 24 hours after transfection, media was removed and cells were washed with 1x PBS. Western blotting was done following the protocol previously outlined above. 100 µL of RIPA buffer was used for cell lysis. For protein detection on the nitrocellulose membrane a primary antibody solution was made containing 1:1000 concentration of both Monoclonal ANTI-FLAG® M2 (Sigma) and Living Colors® Full Length GFP Polyclonal (CloneTech) antibodies

in 1x PBS with 0.02% sodium azide. A secondary antibody solution was made containing IRDye® 800CW Donkey anti-Mouse (Li-COR®) secondary antibody and IRDye® 680LT Donkey anti-Rabbit (Li-COR®) secondary antibody each at a concentration of 1:15000 in a 33% Odyssey® Blocking Buffer and 66% 1x PBS solution.

2.3.2 Migration Shift Following CIP Treatment

HeLa cells were plated at 25% density in 10cm Costar cell culture plates and left to grow to 85% confluency. Cells were transfected using GenJet™ transfection reagent (SignaGen Laboratories) for 24 hours. Each 10cm dish was transfected with 2 µg of plasmid DNA (pCDNA Flag-NIPA363His, Flag-NIPA363Arg or empty pCDNA). Following 24 hour transfection, cells were washed with 1xPBS and harvested using 200 µL of RIPA buffer supplemented with 1x cOmplete EDTA-Free protease inhibitor tablets (Roche). 4 µL of calf intestinal phosphatase (CIP; New England BioLabs) was added to 40 µg protein and incubated at 37°C for one hour. Control samples had 1x PhoSTOP phosphatase inhibitor tablets (Roche) added to the RIPA and cOmplete buffer and were also placed at 37°C for one hour. 20 µg of protein were used for SDS-PAGE and western blotting following the previously stated protocol. Primary antibody solution contain 1:1000 dilution of both Monoclonal ANTI-FLAG® M2 (Sigma) and Rabbit Phospho-Antibody to S354 on NIPA (Abcam) in 1x PBS with 0.02% sodium azide. A secondary antibody solution was made containing IRDye® 800CW Donkey anti-Mouse (Li-COR®) secondary antibody and IRDye®

680LT Donkey anti-Rabbit (Li-COR®) secondary antibody each at a concentration of 1:15000 in a 33% Odyssey® Blocking Buffer and 66% 1x PBS solution.

2.3.3 Migration Shift in Native-PAGE Conditions

HeLa cells were plated at 25% density in 10cm Costar cell culture plates and left to grow to 85% confluency. Cells were transfected with 2 µg of DNA (pCDNA Flag-NIPA363His, pCDNA Flag-NIPA363Arg or empty pCDNA) using 2 µL of GenJet™ transfection reagent (SignaGen Laboratories) per 1 µg DNA. After 24hours cells were harvested by washing with 1x PBS, scraping cells and placing them in a new 1.5mL microcentrifuge tube. Cells were pelleted by centrifugation at 4°C on the Eppendorf Centrifuge 54515C set at 14,000rpm for 30 seconds. Pelleted cells were suspended in 100 µL of Tris Buffer (20 mM Tris-HCl, pH 7.5) followed by sonication 3x 5minutes on the Bioruptor™ UCD-200 (Diagenode). Sonicated samples were spun down by centrifugation at 4°C on the Eppendorf Centrifuge 54515C set at 14, 000rpm for 5 minutes. Supernatant was removed and transferred to a new 1.5mL microcentrifuge tube. 25ug of protein was mixed with sample buffer (2×): 0.187 M Tris/HCl (pH 6.8), 30% glycerol and 80 µg/mL Bromophenol Blue. All samples were then loaded into an 8% running (26% Acryl-Bis (29:1), 73% 0.375m Tris HCl pH 8.8, 0.1% APS, 0.1% TeMed), 4% stacking (13.5% Acryl-Bis (29:1), 85.5% 0.375m Tris HCl pH 8.8, 0.1% APS, 0.1% TeMed) native-PAGE gel. The gel electrophoresis was run in SDS-free running buffer at 200v for 3hrs at 4°C. Western blotting was done following the previously stated protocol. For protein detection on the nitrocellulose membrane, a primary antibody solution was made containing 1:1000 concentration

of Monoclonal ANTI-FLAG® M2 (Sigma) in 1x PBS with 0.02% sodium azide. A secondary antibody solution was made containing IRDye® 800CW Donkey anti-Mouse (Li-COR®) at a concentration of 1:15000 in a 33% Odyssey® Blocking Buffer and 66% 1x PBS solution.

2.3.4 Co-Immunoprecipitation of NIPA

HEK 293T cells were plated at 25% density in a 6-well dish and grown until they reached 80% confluency. The cells were transfected with a total of 2 µg DNA. The following co-transfections were prepared: a) 1.2 µg empty pEYFP + 0.8 µg empty pCDNA, b) 1.2 µg pEYFP FP-NIPA363His + 0.8 µg empty pCDNA, c) 1.2 µg pEYFP FP-NIPA363Arg + 0.8 µg empty pCDNA, d) 1.2 µg empty pEYFP + 0.8 µg pCDNA Flag-NIPA363His, e) 1.2 µg pEYFP FP-NIPA363Arg + 0.8 µg pCDNA Flag-NIPA363His and f) 1.2 µg pEYFP FP-NIPA363His + 0.8 µg pCDNA Flag-NIPA363His. Cells were transfected using 2 µL of GenJet™ transfection reagent (SignaGen Laboratories) per 1 µg DNA. After 24hours, cells were washed and sprayed off with 1x PBS, then transferred to a 1.5mL microcentrifuge tube. Cells were pelleted by centrifugation at 4°C on the Eppendorf Centrifuge 54515C set at 14, 000rpm for 30 seconds. 350 µL of IP Buffer (0.05% Triton X-100 and 1x PBS, supplemented with protease and phosphatase inhibitors) was used to resuspend cells. Cell suspensions were sonicated for 3x 5minutes on the Bioruptor™ UCD-200 (Diagenode). Sonicated samples were spun down by centrifugation at 4°C on the Eppendorf Centrifuge 54515C set at 14, 000rpm for 5 minutes. Supernatant was removed and transferred to a new 1.5mL microcentrifuge tube.

10 µL packed-bead volume of Anti-FLAG® M2 Magnetic Beads, were washed twice by pelleting the beads using a magnetic rack and resuspending in 400 µL of 1x TBS buffer (50 mM Tris-HCl, pH 7.6, 150 mM NaCl). After final removal of 1x TBS buffer, 50 µL of cell lysates were added to the pelleted beads along with 200 µL of supplemented IP buffer to increase total volume for better bead dispersal. Samples were incubated for 2 hours at 4°C with gentle rotation. Beads were collected using a magnetic separator and washed twice with 400 µL of 1x TBS buffer. 30ul of 5x SDS protein loading buffer was added to pelleted beads and boiled at 96°C for 3mins. Beads were pelleted on a magnetic separator and supernatant was transferred to a new 1.5mL microcentrifuge tube. 15ul of each sample was used for SDS-PAGE and western blotting following the previously stated protocol. For protein detection on the nitrocellulose membrane a primary antibody solution was made containing 1:1000 concentration of both Monoclonal ANTI-FLAG® M2 (Sigma) and Living Colors® Full Length GFP Polyclonal (CloneTech) antibodies in 1x PBS with 0.02% sodium azide. A secondary antibody solution was made containing IRDye® 800CW Donkey anti-Mouse (Li-COR®) secondary antibody and IRDye® 680LT Donkey anti-Rabbit (Li-COR®) secondary antibody each at a concentration of 1:15000 in a 33% Odyssey® Blocking Buffer and 66% 1x PBS solution.

2.3.5 Evaluation of NIPA S354 Phosphorylation

HeLa cells were seeded into 6-well Costar cell culture plates at 25% density. They were grown to 80% confluency in order to transiently transfect them. Cells were transfected using GenJet™ transfection reagent (SignaGen Laboratories). Each well was transfected with a total of 1 µg or 2 µg plasmid DNA (pCDNA Flag-NIPA363His, pCDNA Flag-NIPA363Arg or empty pCDNA). 24 hours after transfection, media was removed and cells were washed with 1x PBS. Western blotting was done following the protocol previously outlined above. 50 µL of supplemented RIPA buffer was used for cell lysis. For protein detection on the nitrocellulose membrane a primary antibody solution was made containing 1:1000 concentration of both Monoclonal ANTI-FLAG® M2 antibody (Sigma) and Rabbit Phospho-Antibody to S354 on NIPA (Abcam) in 1x PBS with 0.02% sodium azide. A secondary antibody solution was made containing IRDye® 800CW Donkey anti-Mouse (Li-COR®) secondary antibody and IRDye® 680LT Donkey anti-Rabbit (Li-COR®) secondary antibody each at a concentration of 1:15000 in a 33% Odyssey® Blocking Buffer and 66% 1x PBS solution.

After signal detection, images were analyzed for antibody intensity. This was done using Odyssey® Analysis Software V3.0.21 (LI-COR®). Background values were subtracted from raw values. A ratio of total Phospho-S354 to Flag signal was calculated for each well and an average of six lanes (three biological replicates for each condition loaded into two wells) was taken.

2.4 Cell Proliferation Assay

HeLa cells were grown to 80% confluency then trypsinized. The cells were counted using TC10 Counting Slides (BioRad) and TC10 Automated Cell Counter (BioRad). They were then seeded at 30,000 cells/ well in a Costar 24-well culture plate. Cells were left to grow overnight at 37°C at 5% CO_2. Using GenJet™ transfection reagent (SignaGen Laboratories), cells were transfected with 0.5ug of DNA per well (pCDNA Flag-NIPA363Arg, pCDNA Flag-NIPA363His, empty pCDNA). After 24hours media was removed and replaced with 250ul of media supplemented with Premix WST-1 Cell Proliferation Reagent (TaKaRa; 100ul WST-1 reagent/1mL media). Cells were incubated for 1 hour at 37°C at 5% CO_2. Two wells without cells had 250ul of WST-1 supplemented media and were used for background absorbance readings. Absorbance was read at 460nm on the Synergy MX BioTek® plate reader.

To analyzed, absorbance of the background control wells were averaged and subtracted from each raw absorbance. Six wells of each condition were averaged for each of the five experiments. Statistical analysis was conducted on averages of five experiments.

2.5 Allelic Expression Analysis in Human Whole Blood

2.5.1 Differential Allelic Expression

RNA was previously isolated from whole blood samples of Ottawa Heart Study Control subjects (cohort population characteristics found in **Table 3**). Both RNA isolation and cDNA synthesis was done by Adrianna Douvris. A total of 120 samples were used for expression analysis.

For each cDNA sample, 4.5 µL of cDNA, 5 µL of TaqMan MasterMix (Thermo Fisher) and 0.5 µL of rs11556924 TaqMan specific probe (Thermo Fisher) were mixed and loaded in a well of a LightCycler® 480 Multiwell Plate 96, White (Roche). A standard curve was made by diluting homozygote DNA samples to 5ng/ µL. The following ratios of CC (R variant) to TT (H variant) patient DNA were used to create the probe efficiency standard curve: 1:9, 2:8, 3:7, 4:6, 5:5, 6:4, 7:3, 8:2, 9:1. A total of 4.5 µL of DNA was used in each reaction.

All samples were run on a LightCycler® 480II instrument (Roche) under the following conditions: 1) heat to 95°C for 10mins, 2) 40 cycles of (i) 95°C for 15sec, (ii) 60°C for 1 min, 3) final cool down. LightCycler® 480 software (Roche) was used to analyze the readings and make the genotype calling. Raw data was exported into excel for further analysis. First a standard curve was made using the values of the standard curve samples. Values were $\log^{10}$ transformed in order to make a more linear relationship. Thus all sample values were $\log^{10}$ transformed, followed by transformation using the standard curve equation. Only heterozygote patient samples were used to assess the expression level of each variant.

2.5.2 Total NIPA RNA Expression

The same patient cDNA made from extracted whole blood RNA was used to measure total NIPA RNA levels. Primers for NIPA (see Appendix A) and peptidylprolyl isomerase A (PPIA -housekeeping gene; Appendix A) were used to determine total amount of NIPA. Each 20 µL reaction was set up in a LightCycler® 480 Multiwell Plate 96, White (Roche) and contained 2.5 µL cDNA (1:5 dilution), 1.5 µL H_2O, 0.5 µM forward primer (0.5 µM), 0.5 µL reverse primer (0.5 µM) and 5 µL of SYBR Green I Master (Roche). Quantitative PCR was performed using the LightCycler® 480II instrument (Roche) under the following conditions: 1) pre-incubation at 95°C for 10mins, 2) 45 cycles of amplification and quantification: (i) 95°C for 10secs, (ii) annealing at 60°C (touchdown program; annealing temperature decreases by 0.5°C/cycle) for 30secs, (iii) extension at 72°C for 20secs and (iv) signal acquisition at 82°C for 5secs; 3) Melting of all products occurred in the following manner: 95°C for 10secs, 40°C for 1min, 60°C for 10secs, and finally melting was achieved by gradually increasing temperature at a rate of 0.11°C/sec and signals were acquired at 5/°C.

Using LightCycler® 480 software (Roche), PCR products were initially verified by melting curves. Relative copy numbers of NIPA transcripts were normalized to the expression levels of the abundantly expressed house-keeping gene PPIA.

2.6 Yeast Expression System

2.6.1 Histidine tagging NIPA

To purify protein using a column, a His-tag was added to the NIPA363Arg and NIPA363His. Through the use of specific primers (Appendix A) MHHHHHHDP (inframe His-tag) was added onto the N-terminus of NIPA. Both variants were His-tagged and amplified using the Roche Fast-Start Taq DNA polymerase dNTP pack. Each 25µl reaction contain the following, 1x PCR Buffer with 2mM $MgCl_2$, 0.2mM of each dNTP, 1µl (~20 µM) each forward and reverse primer (Appendix A), and 1U Taq DNA polymerase and 0.5µl pCDNA containing NIPA (one reaction for NIPA363Arg and one for NIPA363His). Reactions were run on the Applied Biosystems (ABI) GeneAmp® PCR System 9700 using the following cycling conditions: an initial 30 sec denaturation at 98°C, 5 cycles of 98°C for 10sec, annealing at 60°C for 20sec, extension at 72°C for 2min 30sec; followed by 40 cycles of 98°C for 10sec, annealing at 72°C for 20sec, extension at 72°C for 2min 30sec and a final extension for 2min at 72°C at the end of the 40 cycles. 2ul of PCR reaction was used to check for proper product size (~500bp) on a 1% agarose gel. DNA sequencing was conducted to assess frame and to ensure no mutations were generated.

Following PCR, the products were digested to prepare them for plasmid insertion. A 30ul reaction was made for both the His-NIPA variants with the following protocol: 25 µl PCR product (~350ng DNA), 3 µl 10x CutSmart enzyme buffer (New England Biolabs), 1 µl BamH1 digestive enzyme (New England Biolabs) and 1 µl of Xho1 (New England Biolabs). Digestion of the pKLAC2 vector was done in parallel: 1.0ng of pKLAC2 (New England Biolabs), 6 µl 10x CutSmart enzyme buffer (New

England Biolabs), 2 µl BamH1 digestive enzyme (New England Biolabs) and 2 µl of Xho1 (New England Biolabs) for a total reaction volume of 60 µl. Digestions were done at 37°C for 2hours. Products were cleaned up using High Pure PCR Product Purification Kit (Roche).

HIS-NIPA363Arg and HIS-NIPA363His digested inserts were then ligated with the digested plasmid for 4 hours at room temperature. Each ligation reaction contained the following: 2 µl 10X T4 DNA Ligase buffer (New England Biolabs), 2 µl digested vector, 8 µl digested PCR insert, 1 µl (~400 units) T4 DNA Ligase (New England Biolabs) and 7ul of deionized water. A ligation reaction with no insert was done to create an empty pKLAC2 (EV) vector for control experiments: 2 µl 10X T4 DNA Ligase buffer (New England Biolabs), 2 µl digested vector, 1 µl (~400 units) T4 DNA Ligase (New England Biolabs) and 15 µl of deionized water.

2.6.2 Cloning NIPA containing plasmid into *K. lactis* yeast.

Ligation reactions were then used to transform into *Kluyveromyces lactis* yeast. First the plasmids were linearized at 37°C for 2 hours using the following reactions: 1) HIS-NIPA363His: 2.15 µl pKLAC2 HIS-NIPA363His ligation reaction, 1 µl SacII (New England Biolabs), 5ul 10x CutSmart Enzyme Buffer (New England Biolabs) and 41.85 µl deionized water. 2) HIS-NIPA363Arg: 8.69 µl pKLAC2 HIS-NIPA363Arg ligation reaction, 1 µl SacII (New England Biolabs), 5ul 10x CutSmart Enzyme Buffer (New England Biolabs) and 35.31 µl deionized water. 3) EV: 2.2 µl pKLAC2 EV ligation reaction, 1 µl SacII (New England Biolabs), 5ul 10x CutSmart

Enzyme Buffer (New England Biolabs) and 41.80 µl deionized water. Following incubation, all linearization reactions were cleaned up using High Pure PCR Product Purification Kit (Roche).

2.6.3 Transformation and growth

First, *K. lactis* GG799 cells were thawed on ice for 10 minutes. 620 µl of NEB Yeast Transformation Reagent was added to cells. 1 µg of linearized pKLAC2 DNA with insert (NIPA363His:7ul, NIPA363Arg: 15ul, EV:15ul) was added to the cells and incubated at 30° for 30mins. Reactions were then heat shocked in a 37°C water bath for 1 hour. Cells were spun at 7000rpm on an Eppendorf Centrifuge 54515C for 2mins to pellet cells. After pouring off the supernatant, 1mL of YPGlu Media was added (10g Bacto™ Yeast Extract, 20g Bacto™ Peptone, 50mL sterile 40% glucose in 950ml deionized water). Pelleting of cells and resuspension in YPGlu media was repeated twice. Samples were then incubated at 30°C for 4 hours with shaking (250-300rpm). Cells were spun at 7000rpm on an Eppendorf Centrifuge 54515C for 2mins to pellet cells. Supernatant was discarded and yeast was resuspended in 1mL 1x PBS. 10ul of solution was plated on YCB agar plates (15ml 1M Tris-HCl buffer solution, 5.85g YCB medium powder [in *K. lactis* protein expression kit; New England Biolabs], 10g Bacto™ Agar, 1x Acetamide stock solution [in *K. lactis* protein expression kit; New England Biolabs]) and left to grow for 4 days at 30°C. Nine single colonies were picked and streaked on a new YCB agar plate and left to grow at 30°C for 2 days.

2.6.4 *K. lactis* colony selection for NIPA expression

Yeast colonies were PCR screened for genome incorporation of NIPA. First a 0.1mm^2 area of each colony patch was scraped and transferred to a 1.5mL microcentrifuge tube. They were then suspended in 100 µl of 200mM lithium acetate 1% SDS solution and incubated at room temperature for 10min. 300 µl of 96% ethanol was added to the solution, vortexed to mix and centrifuged at 12,000rpm on an Eppendorf Centrifuge 54515C for 3mins to pellet DNA. Supernatant was removed and the pellet was washed with 500 µl of 70% ethanol. DNA was resuspended in 100 µl of TE buffer (10mM Tris-HCl, pH 8 and 0.1mM EDTA). To determine single copy integration, Integration Primers 1 and 2 from the *K. lactis* Protein Expression Kit (New England Biolabs; Appendix A) were used in the PCR reactions to generate a 2.4kb product. To identify colonies with multiple copy integration, Integration Primers 2 and 3 from the *K. lactis* Protein Expression Kit (New England Biolabs; Appendix A) were used in the PCR reactions to generate a 2.3kb product. Both PCR reactions, for all colonies, were run under the following PCR parameters: 30 cycles of 94°C for 30sec, 50°C for 30sec and 72°C for 3min, followed by 10min incubation at 72°C. To check for PCR products, 10µl of each reaction was run on a 1% agarose gel.

2.6.5 NIPA protein expression

Following 2 days of growth at 30°C, a 2mm^2 area of each patched colony was scraped and resuspended in 2mL of YPGal media (10g Bacto™ Yeast Extract, 20g Bacto™ Peptone, 50mL sterile 40% galactose in 950ml deionized water) and

incubated at 30°C with shaking (250rpm) for 2 days. The yeast growth was transferred to three 1.5mL microcentrifuge tubes and spun down at 7000rpm on an Eppendorf Centrifuge 54515C for 2mins to pellet cells. The supernants were transferred to new 1.5mL microcentrifuge tubes. The supernatant was passed through a 0.45µm PVDF membrane filter to remove residual cells. The protein was precipitated using 900µl of supernatant and 100µl 100% Trichloroacetic acid solution (Sigma). The solution was mixed and incubated on ice for 10min then centrifuged at 14,000rpm on an Eppendorf Centrifuge 54515C for 2mins to pellet the protein. Supernant was discarded and the protein pellet was washed twice with 100% cold acetone. The pellet was air dried on ice and resuspended in 50µl of 0.1M Tris, pH 8.8. 5X SDS-PAGE loading buffer and β-mercaptoethanol was added to 20µl of protein suspension for a final concentration of 1X and 5% (v/v), respectively.

To assess protein expression, each sample was boiled for 5 minutes at 96°C and loaded onto an 8% polyacrylamide gel (4% stacking). Electrophoresis was conducted for approximately 45 minutes at 200v. The proteins were transferred to a 0.45 µm nitrocellulose membrane (BioRad) using the Trans-Blot® Turbo™ transfer pack and Trans-Blot® Turbo™ transfer system (BioRad) set at 2.5 constant amps, 25 constant volts for 7 minutes. The nitrocellulose membrane was blocked using a 5% milk- 1x PBS solution for 20 minutes at room temperature with gentle shaking. Blocking solution was washed off with 1x PBS. Membrane was incubated overnight at 4°C in a primary antibody solution containing a 1:1000 dilution of NIPA Rabbit Ab (Cell Signaling), 1x PBS and 0.02% sodium azide. The membrane was washed for

4x5mins with 1xPBS and incubated in the dark for 1 hour in a secondary antibody solution containing IRDye® 680LT Donkey anti-Rabbit (Li-COR®) at a concentration of 1:15000 in a 33% Odyssey® Blocking Buffer and 66% 1x PBS. The membrane was again washed 4x5mins with 1x PBS. The signals were detected using Odyssey® Infrared Imaging System (LI-COR®) and analysis was done using Odyssey® Analysis Software V3.0.21 (LI-COR®).

3 Results

3.1 Western blot analysis identifies slight difference in migration between NIPA363Arg and NIPA363His variants under SDS-PAGE conditions.

In an attempt to analyze NIPA and the impact of Arg363His on its function, Western blot analysis was crucial to analyze potential protein differences. Amino acid changes can have impacts on the overall function, stability and conformation of a protein[41]. Using SDS-PAGE, one can study proteins at a molecular level.

In multiple SDS-PAGE experiments determining NIPA expression, the NIPA variants consistently produced a very small shift between NIPA363Arg and NIPA363His variants. To further investigate if the shift was reproducible at different concentrations of variant expression, HeLa cells were co-transfected with FP (as a transfection efficiency control) and either NIPA363His or NIPA363Arg DNA. Protein lysates were run on an 8% SDS-PAGE gel and visualized using FP and NIPA specific antibodies. As seen in **Figure 4**, even at different levels of protein expression, NIPA363His migrated through the SDS gel slightly slower than NIPA363Arg. This slight, but reproducible shift suggests that at the protein level an innate difference exists between the two NIPA variants.

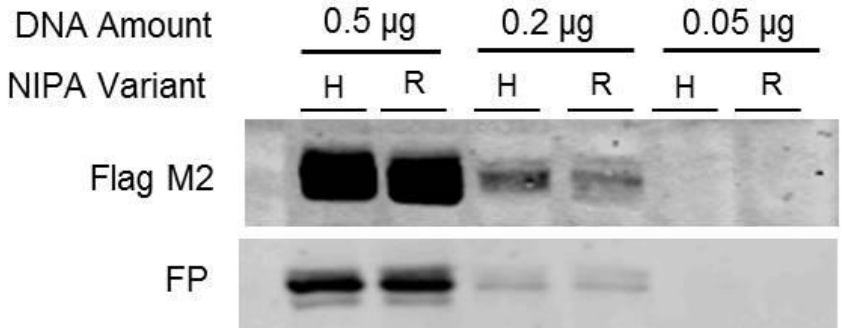

DNA Amount
0.5 µg
0.2 µg
0.05 µg
NIPA Variant
H
R
H
R
H
R
Flag M2
FP

Figure 4. A slight migration shift is seen between NIPA363Arg and NIPA363His variants. HeLa cells were co-transfected with pEYFP fluorescent protein (FP) constructs and pCDNA Flag-NIPA variants in decreasing amounts of total DNA (0.5 µg, 0.2 µg, 0.05 µg).

3.2 Slight shift between NIPA363His and NIPA363Arg remains following phosphatase treatment.

It is well known that phosphorylation of proteins can lead to differences in their ability to migrate through SDS-PAGE gels. NIPA is known to be phosphorylated at various residues in close proximity to amino acid 363 (location of the rs1556924 SNP)[28,38]Any alterations or differences in phosphorylation may affect NIPA's ability to function properly[38]. Phosphorylation at the critical residues (S354, S359 and S395) is vital for proper NIPA function and ultimately a cells timely entrance into mitosis[28,38].

To determine if alternative phosphorylation of NIPA was resulting in alternative migration of the protein through SDS-PAGE gels, samples were treated with CIP (phosphatase). By removing the phosphorylation from the equation the shift could be evaluated based on amino acid sequence in hopes of determining if the slight difference in migration between NIPA363Arg and NIPA363His is amino acid based. **Figure 5** shows that even after phosphatase treatment the slight shift between NIPA363Arg and NIPA363His is still clearly visible. This suggests that the shift is therefore not due to the phosphorylation differences between the variants.

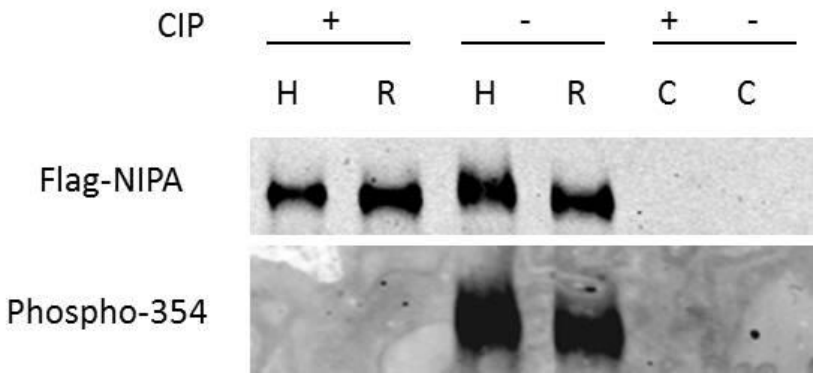

CIP
+
-
+
-
H
R
H
R
C
C
Flag-NIPA
Phospho-354

Figure 5. Removal of phosphate groups using CIP does not affect the shift seen between NIPA363Arg and NIPA363His variants. HeLa cells were transfected with 2ug of pcDNA Flag-NIPA363Arg, Flag-NIPA363His or empty pcDNA vector for a control. Protein harvested from HeLa cells overexpressing NIPA variants was exposed to calf Intestinal phosphatase (CIP) for 1 hour at 37°C (C). Samples were then run on an 8% SDS-PAGE gel.

3.3 Slight shift between NIPA363Arg and NIPA363His is more pronounced when run in non-reducing, non-denaturing PAGE conditions.

Since the shift between NIPA363Arg and NIPA363His variants did not appear to be due to phosphorylation differences, further investigation was done to identify possible explanations. One hypothesis was that there may be a structural and/or conformational difference between the variants as a result of the amino acid substitution. To test this, protein lysate from cells overexpressing each variant were run on an 8% PAGE gel in non-denaturing, non-reducing conditions (native conditions). **Figure 6** shows that the shift initially seen between NIPA363His and NIPA363Arg variants is more pronounced under native conditions. This suggests that a conformational or structural difference exists between the two variants.

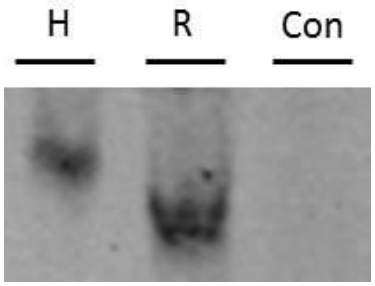

H
R
Con

Figure 6. HeLa cell extracts run on a PAGE-gel under native conditions exaggerates the shift between NIPA363Arg and NIPA363His variants. HeLa cells were transfected with 2ug of pcDNA Flag-NIPA363Arg or Flag-NIPA363His. Protein was run on an 8% native PAGE gel under non-reducing, non-denaturing conditions.

3.4 Successful development of a *K. lactis* yeast expression system for the purification of NIPA363Arg and NIPA363His variants.

The previous experiments suggest that a conformation and/or structural change exists between the two NIPA363Arg and NIPA363His variants. In order to investigate this further, we initiated a collaboration with Dr. Cuture for the eventual crystallization of NIPA. Crystallization requires high volumes of purified protein, using an efficient expression system.

As seen in **Figure 7**, the first steps of a high volume NIPA expression system were successfully generated. In order to determine if the NIPA expression cassette was successfully inserted into the *K. lactis* genome, PCR of a 300bp segment was generated and run on a gel to identify clones with inserts. It can be seen in **Figure 7a** that 5 successful NIPA363His clones and 6 successful NIPA363Arg clones were generated. In order to check that the NIPA cassettes were inserted into the *K. lactis* genome in the correct orientation, PCR was done to check for orientation and multiple insertions, and then run on an agarose gel. **Figure 7b/c** shows that there are 4 NIPA363His clones and 5 NIPA363Arg clones with correct and multiple insertions. Lastly to check for NIPA expression, protein was precipitated from the growth media and a western blot was generated. **Figure 7d** shows that all clones express the protein; however some clones express the protein at higher levels than others.

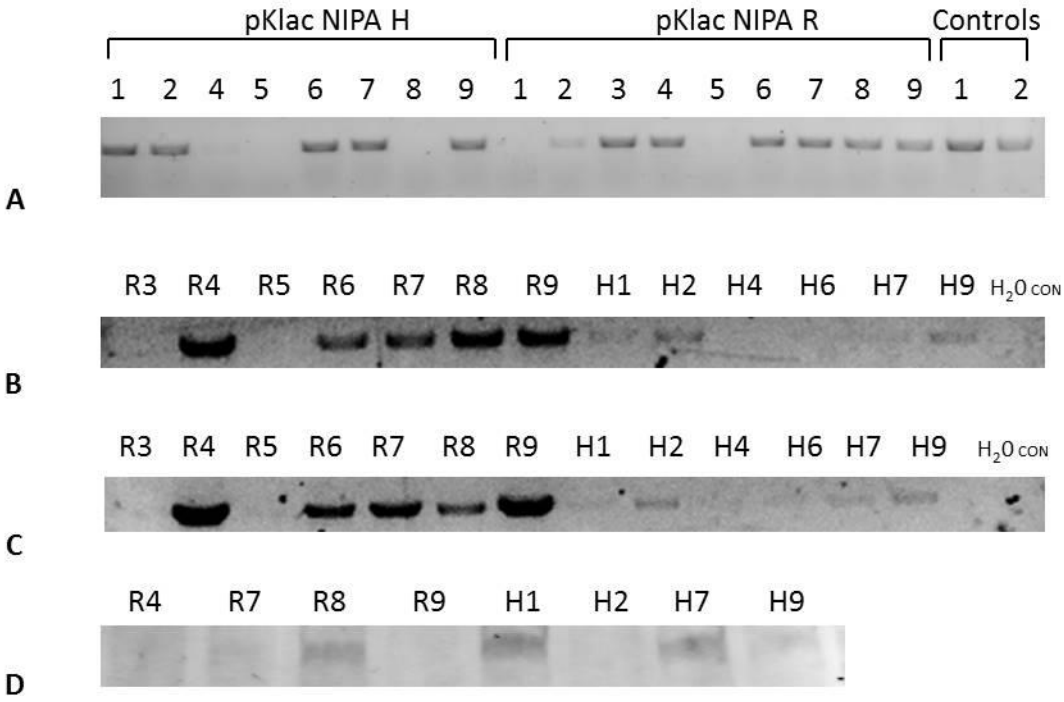

pKlac NIPA H
pKlac NIPA R
Controls
1 2 4 5 6 7 8 9 1 2 3 4 5 6 7 8 9 1 2
A
R3 R4 R5 R6 R7 R8 R9 H1 H2 H4 H6 H7 H9 H2O CON
B
R3 R4 R5 R6 R7 R8 R9 H1 H2 H4 H6 H7 H9 H2O CON
C
R4 R7 R8 R9 H1 H2 H7 H9
D

Figure 7. Development of *K. lactis* yeast colonies overexpressing NIPA R and NIPA H variants. PCR products of a 300bp segment of *ZC3HC1* in independent *K. lactis* yeast colonies (A). PCR products of a region containing the vector and *K. lactis* genome barrier to check for correct plasmid integration (B). PCR products of sequential vector integration regions, to check for multiple copies of integration in the *K. lactis* genome (C). Colonies were grown in 2mLs of growth media for two days. Western blot on 10% SDS-PAGE gel of TCA protein precipitation products from growth media of the various *K. lactis* colonies (D). Antibody for NIPA was used for detection. [Letter denotes the NIPA variant; number denotes colony number]

3.5 Mass overexpression of NIPA results in complex formation with itself.

During the process of crystallization, it is easier to purify stable protein when the protein can create a complex with itself. In addition, the ability of NIPA for homodimerization has never been assessed. In order to determine if NIPA was capable of creating complexes with its self, cells were co-expressed with FP (as a control), FP-NIPA (as one expressed version of NIPA) and/or Flag-NIPA(as the second expressed form of NIPA). After an immunoprecipitation targeted at Flag-NIPA and western blotting for NIPA and FP, it can be seen in **Figure 8** that FP-NIPA and not FP is pulled down when co-expressed with Flag-NIPA. This suggests that NIPA is capable of making complexes with itself.

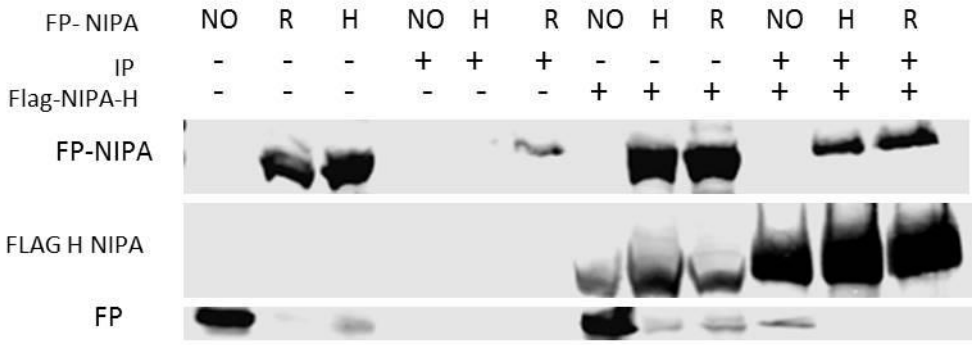

FP- NIPA
NO R H NO H R NO H R NO H R
IP
- - - + + + - - - + + +
Flag-NIPA-H
- - - - - - + + + + + +
FP-NIPA
FLAG H NIPA
FP

Figure 8. Mass overexpression of NIPA results in complex formation with itself. HeLa cells were co-transfected with a total of 5ug of pCDNA comprised of a mix of fluorescently tagged NIPA variants [FP-NIPA363Arg, FP-NIPA363His], fluorescent protein [FP] and Flag-tagged NIPA H variant [Flag-NIPA363His]. Following an immunoprecipitation (IP) with Flag-M2 magnetic beads, untagged FP-NIPA can be seen in the IP samples.

3.6 S354 phosphorylation of NIPA363Arg versus NIPA363His variants.

NIPA activity is highly regulated by the phosphorylation of three critical serine residues [38]. These serine residues lie in close proximity to amino acid 363, which is altered based on the variant present at rs11556924. To determine if this amino acid change alters the phosphorylation of S354, the one responsible for disassociation from SCF^{NIPA}, cells were overexpressed with Fag-NIPA363Arg or Flag-NIPA363His and cell lysates were run on an SDS-PAGE gel followed by western blotting. A ratio of S354 to Flag antibody intensities were calculated for each variant. **Figure 9** shows that although there seems to be a trend towards less phosphorylation of NIPA363Arg, the differences between the phosphorylation of S354 is not statistically significant.

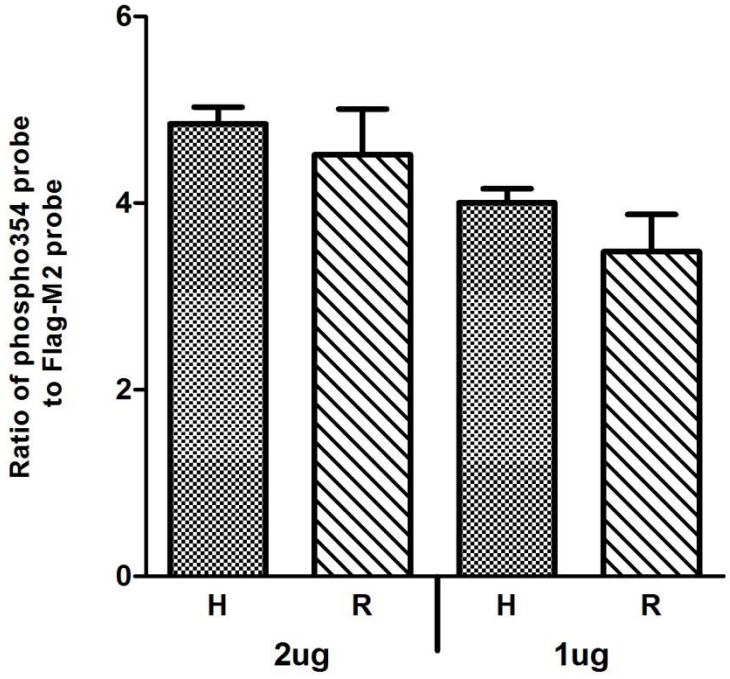

Ratio of phospho354 probe
to Flag-M2 probe
6
4
2
0
H
R
H
R
2ug
1ug

Figure 9. There is no statistically significant difference between phosphorylation of S354 between NIPA363Arg and NIPA363His variants in HeLa cells. HeLa cells were transfected with 2ug or 1ug of pCDNA Flag-NIPA363Arg or Flag-NIPA363His for 24hrs. Protein lysates were run on 8% SDS-PAGE gel, blotted to 0.45um nitrocellulose paper and exposed to Flag-M2 and phospho354 antibodies. Ratios of total NIPA (quantified for total Flag signal detected) and phospho354-NIPA (quantified by total phospho354 signal detected) were compared. Average of three triplicate wells in each experiment; ± Standard Error (SE) (n=3).

3.7 Following overexpression, HeLa cell proliferation is significantly slower in NIPA363His than NIPA363Arg transfected cells.

To better understand the effects of Arg363His variant on cell proliferation, HeLa cells were transfected with pCDNA Flag-NIPA363Arg or Flag-NIPA363His. Following 24hr overexpression, WST-1 reagent was added to the wells and absorbance readings were taken after an hour. Readings were adjusted for background readings and normalized to NIPA363His variants. **Figure 10** shows that there is a slight but significant increased proliferation rate of HeLa cells transfected with NIPA363Arg compared to cells transfected with NIPA363His. However, this figure also shows that neither NIPA363His nor NIPA363Arg expressing cells proliferate at significantly different rate from the control. This suggests that the single amino acid substitution alters overall cell proliferation rate.

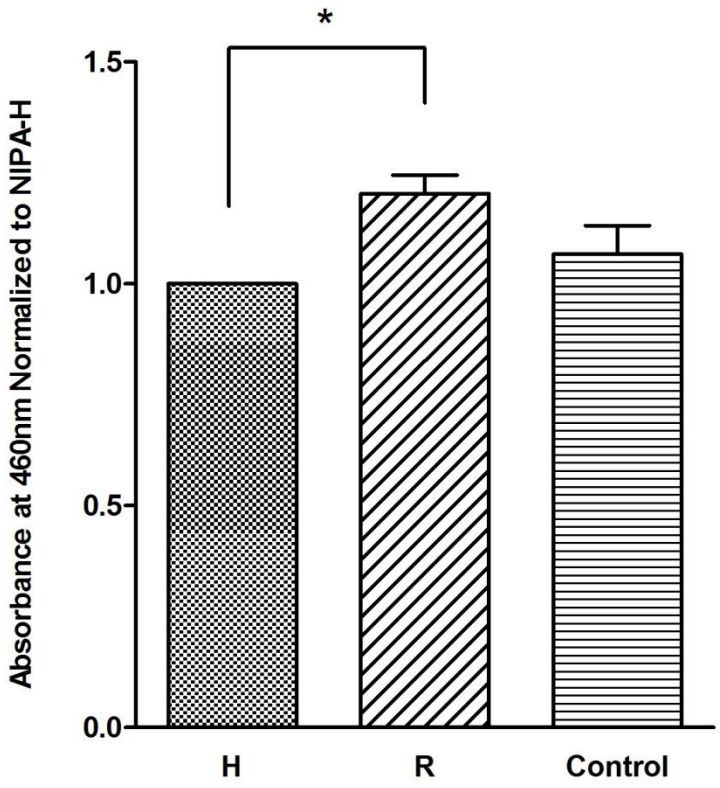

*
1.5
1.0
0.5
0.0
Absorbance at 460nm Normalized to NIPA-H
H
R
Control

Figure 10. HeLa cells overexpressing NIPA363Arg proliferate at a significantly faster rate than HeLa cells overexpressing NIPA363His. HeLa cells were counted and plated in a 24-well plate at a density of 30,000cells/well. 24hrs after seeding, cells were transfected with 0.5 µg pCDNA Flag-NIPA363Arg or Flag-NIPA363His. Following 24hrs of transfection, media was removed and replaced with 100ul of growth media supplemented with 10ul of WST-1 reagent (TaKaRa). Absorbance readings were taken after 1hr. Background absorbance readings were subtracted from raw readings. All adjusted readings were normalized to NIPA363His average absorbance. For each experiment an average of 6 replicate wells were used. A one-way ANOVA was conducted between NIPA363Arg, NIPA363His and Control treatments [$F(2,12)= 5.3512$, $p= 0.0218$]; ± SE (n=5).

3.8 rs11556924 heterozygote control patients have significantly less expression of NIPA363Arg than NIPA363His in whole blood samples.

Previous studies have shown that non-synonymous mutations have the ability to affect expression or stability of proteins particularly in disease related genes[42,43]. This suggests that even the simplest disruption to the conserved amino acid sequence can affect the ability of the protein to function at its highest capacity.

In order to investigate whether the rs11556924 variant was expressed differently, or NIPA RNA levels differed between alleles, an allele specific qPCR assay was undertaken. In order to do this, cDNA libraries that were previously made from the isolation of RNA from human whole blood samples were used (cohort population characteristics found in **Table 3**). A Taqman (Thermo Fisher) assay was used to probe for either the C-allele or T-allele in the heterozygote samples. To account for differences in probe efficiencies between alleles, a standard curve was generated using homozygote CC and TT patient DNA run in parallel with the cDNA samples. The standard curve readings were log transformed to create the standard cure. All raw sample probe readings were $\log^{10}$ transformed and then adjusted using the standard curve. **Figure 11** clearly shows a significant 13% difference between the levels of NIPA363Arg transcripts and NIPA363His transcripts in whole blood samples of rs11556924 heterozygote patients. This suggests that the rs11556924 variant may be affecting the expression or more likely the stability of NIPA RNA.

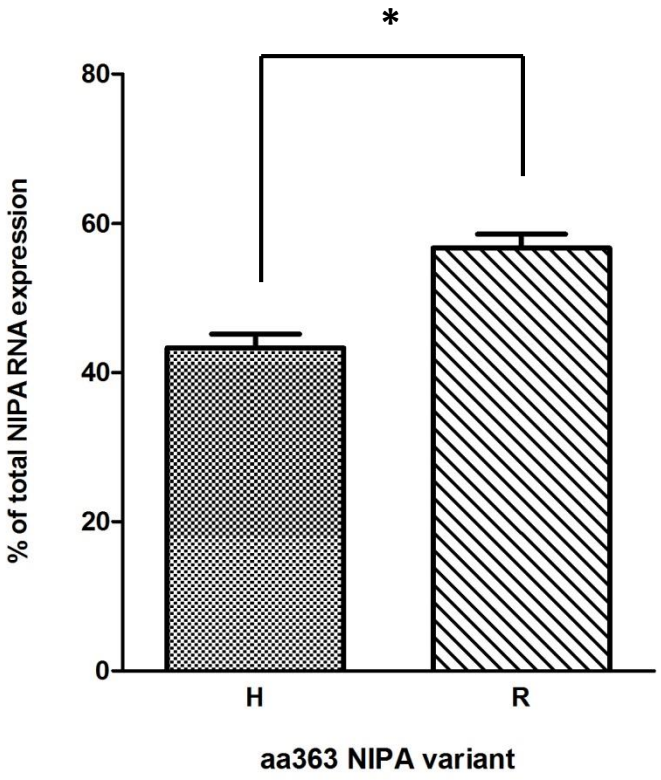

*
80
60
40
20
0
% of total NIPA RNA expression
H
R
aa363 NIPA variant

Figure 11. Allele specific expression analysis of NIPA R and H transcripts in rs11556924 heterozygote human whole blood samples shows a significant decrease in the expression of NIPA363His transcripts compared to NIPA363Arg transcripts. Total RNA was extracted from control patient whole blood samples. A Taqman® (Thermo Fisher) assay for rs11556924 was used in conjunction with LightCycler 480 (Roche) to identify allele specific RNA transcripts in whole blood RNA samples. Probe detection efficiencies were corrected using a standard curve and log10 transformed to fit the curve.[*= $p<0.005$]; ± Standard deviation (SD) (n= 160)

3.9 NIPA RNA expression levels in human whole blood samples showed no significant difference between rs11556924 genotypes.

Differential expression of variants in homozygotes may suggest that there is a genotype-dosage correlation according to rs11556924 genotype. If such a correlation exists then it could be hypothesized that total expression of each NIPA variant is dependent on genotype of rs11556924. By evaluating total RNA expression levels of NIPA, it may provide further insight into the effects of rs11556924 on CAD development.

In order to investigate whether the rs11556924 genotype affected expression of NIPA, total NIPA RNA transcripts were investigated. Complete cDNA libraries of RNA previously isolated from human whole blood samples were used to run quantitative real-time PCR. Using NIPA RNA specific primers (Appendix A), cDNA samples were probed for total amount of NIPA present. Values were compared to the levels of the constitutively high expressing protein PPIA (Peptidylprolyl isomerase A). Ratios were normalized to the average expression ratio of heterozygote patients. As seen in **Figure 12**, there was no significant difference in the levels NIPA in whole blood samples when comparing the rs11556924 genotype. This suggests that the rs11556924 variant does not affect the overall expression levels of NIPA in heterozygote or homozygote (CC or TT) patients.

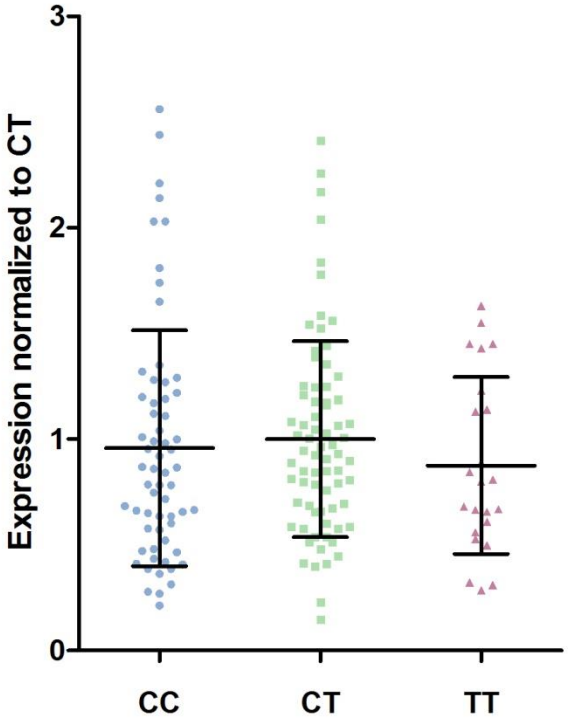

Expression normalized to CT
3
2
1
0
CC
CT
TT

Figure 12. Total NIPA RNA expression in human whole blood samples shows no difference between NIPA363Arg and NIPA363His variants. Whole blood RNA was extracted from 160 human control patients for use in qRT-PCR. Total NIPA RNA levels were measured using qRT-PCR and normalized to the levels of constitutively expressed house-keeping gene *PPIA*. To plot, values were then normalized to CT total-average [CC= NIPA363Arg variant; TT= NIPA363His variant; CT= heterozygote];± SD (n= 160)

4 Discussion

Here I have evaluated the hypothesis that the Arg363His variant (rs11556924) associated with protection against CAD is impeding the phosphorylation of critical serine residues S354 and S359 of NIPA, preventing NIPA's disassociation from SCF^{NIPA} complex and consequently attenuating cell proliferation and leading to delayed atherosclerosis.

This thesis was able to demonstrate that there is a statistically significant decrease in cellular proliferation of cells expressing NIPA363His versus cells expressing NIPA363Arg despite no statistically significant difference in S354 phosphorylation. I was also able to identify a conformational or structural difference between the variants through multiple western blots. Finally, a 15% difference of NIPA mRNA levels in heterozygote human whole blood samples was found. These findings support the hypothesis that NIPA363His is affecting the development of atherosclerosis through delayed cell proliferation. The exact molecular mechanisms which are responsible for this difference have yet to be fully described; however this thesis has provided a good foundation for future experiments to fully understand molecular mechanisms behind the findings presented here.

Initially in the lab NIPA was being investigated at a protein level by multiple western blot experiments. From these experiments, it was clear that a shift could be seen between the two variants (**Figure 4**). This shift is not due to phosphorylation which is one of the first modifications that is attributed to slight migration shifts seen

between protein variants on an SDS-PAGE gel (**Figure 5**). Other modifications such as glycosylation or acetylation may be the cause of this shift. However after thorough evaluation using online prediction programs it is not likely that glycosylation is affecting this region of the protein. Prediction software suggests that there exists only a single site on NIPA that is possibly glycosylated and is located close to the C terminus of the protein making it unlikely to be affected by the R363H variant.

If phosphorylation and glycosylation are predictably not the source of the difference seen between the two variants, another possible and more probable explanation is a conformational or structural change. Prediction programs such as DIM-pred[44], suggests that this amino acid substitution may be affecting the organization of this region. With 90% confidence, DIM-Pred suggests that upon the substitution of the arginine for a histidine at amino acid position 363, the region surrounding it will go from being an organized sequence to a disorganized sequence, with the potential to cause inactivity or other functional problems related to this region. This is plausible in the sense that an arginine to histidine substitution is not a minor change. Histidines are known to cause kinks in amino acid sequences, thus going from an arginine to a histidine can cause a major local structural change to the region. This information further suggests that this substitution is affecting NIPA folding structure or overall conformation.

Upon the confirmation of the shift through various follow up experiments and eventually under native conditions (**Figure 6**), it is evident that this shift must be due

to a conformational change of the protein. Singh *et al.* (2014) showed that a single amino acid substitution was fully capable of disrupting protein stability without affecting its function. This example supports the idea that the NIPA variant may be disrupting the stability of the protein rather than the overall function. A disruption to protein stability has the capacity to influence phenotype through incorrect protein formation or a complete lack of functional protein, resulting in inhibited cellular function. In order to better understand either the potential effects of the rs11556924 variant on NIPA stability or formation, it is necessary to crystalize the structure.

NIPA is a newly identified protein and thus has yet to be completely understood. As a result the crystalized structure of NIPA has not been determined. As a result of our findings here, we began the process of NIPA crystallization. Crystallization is a long and tedious process. The first step of the process is to mass produce and purify each NIPA variant. The most logical system to do this in is yeast- a eukaryotic system which will produce NIPA in a more biologically relevant way (eg. phosphorylation, folding, etc.) and is cost effective. **Figure 7** shows the effective expression of NIPA protein in the *K. lactis* yeast system. The next steps of this process will involve high-volume growth of yeast colonies and purification of NIPA using the His-tag with a bead and column protocol. Once the protein is of highest purity, two things will happen. The first is the purified protein will be used for future functional experiments to assess possible binding differences between the NIPA363Arg and NIPA363His variants. Secondly, the samples will be sent to collaborators for the crystallization steps. A crystalized structure of NIPA is not only

beneficial to evaluate the structural/conformational differences between the NIPA363Arg and NIPA363His variants, but also for future investigation into the NIPA protein from all angles.

Although phosphorylation doesn't seem to be the cause of the slight migration shift seen on western gels, as determined in the experiments presented here, phosphorylation of S354 is crucial for NIPA disassociation from the SCF complex and its subsequent association with the APC/C complex[28,38]. If phosphorylation of this residue becomes disrupted, the SCF^{NIPA} complex may remain active thus prolonging G2 phase and overall slowing cell proliferation[28]. **Figure 9** shows that although there is no statistically significant difference in S354 phosphorylation between NIPA363Arg and NIPA363His variants, there seems to be a trend towards less phosphorylation of the NIPA363Arg variant than NIPA363His. The S354 has been identified as the critical serine residue for NIPA regulation and thus a slight difference may result in a change in proliferation rate. This slight difference could be due to the potential conformational change identified in this study. Conformational changes due to a single amino acid substitution can alter critical phosphorylation of proteins[45,46]. The hypothesized structural change in the NIPA363His variant could be exposing the protein surface surrounding S354 more, allowing for ERK2 to phosphorylate it more efficiently, and theoretically leading to an overall increase cell proliferation.

It was initially hypothesized that NIPA363His would be associated with reduced rates of cell proliferation, based on the CARDIoGRAM meta-analysis study[21]

which associated rs11556924C/T with protection against CAD. However, based on the current understanding of NIPA, the observed increase in S354 phosphorylation of NIPA363His should result in cells entering mitosis more rapidly and thus increased cell proliferation. In light of this contradiction, NIPA may be involved in other pathways affecting cell cycle that have yet to be identified. Current understanding of NIPA's role in cellular function is still quite limited. To date only 10 papers have been published on NIPA, leading to our limited understanding of this protein. Future studies on the function of NIPA may elucidate other consequences of S354 phosphorylation and a better understanding of the increased phosphorylation of the NIPA363His variant and its implications on cell proliferation.

With the knowledge that NIPA is involved in cell cycle regulation, determining whether the NIPA p.R363H variant affects overall cell proliferation is a critical step in understanding the association of rs11556924 with CAD risk. **Figure 10** clearly shows that there is a statistically significant difference in cell proliferation between HeLa cells expressing the NIPA363Arg versus the NIPA363His variant. NIPA363Arg expressing cells show significantly more proliferation, further suggesting that rs11556924C/T affects proliferation of cells. This supports the hypothesis that the risk allele at rs11556924 which codes for an arginine at amino acid position 363 (NIPA363Arg), increases cell proliferation resulting in higher risks for CAD development. The mechanism(s) through which this variant is affecting proliferation is currently being investigated. As previously stated it could be affecting 1) phosphorylation of regulatory critical serine residues, 2) the stability of the protein and it's affinity for Skp1

or cyclin B1 binding, or 3) the structure/conformation of the protein. Future studies that evaluate the molecular basis of rs11556924 effects will add to a better understanding of the difference in proliferation rates shown in **Figure 10.**

In recent years, cell proliferation of both VSMC and macrophages has been shown to play an important role in maturation of arterial plaque[10–13]. Future experiments should be focused on expressing NIPA363Arg and NIPHA-H variants in VSMCs and macrophages to determine if the effect observed in HeLa cells is replicated in more CAD relevant cell types. The development of two homozygous NIPA (R and H) cell lines from one patient would be ideal for conducting these types of experiments. Using iPSC (induced pluripotent stem cell) and genome editing protocols, cell lines could be developed from a single patient sample, genetically identical except for the rs11556924 variant.

Cell proliferation and NIPA phosphorylation experiments seem the most logical way to evaluate the impact of the NIPA p.R363H variant. However, investigating other potential properties and molecular functions of NIPA will undeniably shed light on other, maybe less obvious functional effects of the rs11556924 variant. To date, it is understood that F-Box proteins, including NIPA, are typically part of the SCF complex and function in the complex as a single protein[40]; thus the ability of NIPA to form a complex with itself has never been explored. As seen in **Figure 8** NIPA is able to form a complex with itself when a surplus of protein exists; however the formation doesn't seem likely to occur naturally. If this complex was found to occur endogenously the efficiency at which the complex is formed would be high as a result

of high affinity for each other. **Figure 8** shows that although the complex does form, it is not an efficient formation— you would expect almost equal amounts of FP-NIPA and Flag-NIPA to be pulled down during immunoprecipitation. It should also be noted that there is no difference between NIPA363Arg and NIPA363His's ability to form this complex; each FP-NIPA variant seems to be able to form a complex at similar efficiency. Despite the fact no difference was seen between the two variants, this evidence still provides further understanding of the NIPA protein. Identifying NIPA's activity and interactions in the cell adds to the overall understanding of NIPAs properties and its role in cellular processes.

Up until this point, studies of NIPA have focused mainly on the activity and function of NIPA at the protein level and its implications on cell proliferation. Although the effects of the rs11556924C/T is presumed to affect NIPA at the protein level, evaluating all possible implications of this variant, including effects on RNA, could provide a more complete understanding of its role in CAD development. Changes to RNA levels or expression can also have phenotypic consequences. Using allele specific probes, it was determined that there were statistically more NIPA363Arg transcripts than NIPA363His transcripts in rs11556924 heterozygote human whole blood samples (**Figure 11**). It is well known that regulatory regions are imperative for gene expression and SNPs that disrupt these regions can alter levels of mRNA transcripts[47] The difference between transcript levels of NIPA363Arg and NIPA363His is not likely due to effects of expression, as the variant does not lie within known regulatory regions. The difference in NIPA363Arg and NIPA363His transcript

levels in heterozygote patients may instead be due to a difference in stability. There is some evidence that suggests that coding variants can affect mRNA stability and translation[48,49]. It is possible that in addition to protein level effects, the rs11556924 SNP can also affect NIPA mRNA stability and/or translation. Future studies targeting NIPA mRNA expression are needed to determine if this variant is affecting mRNA stability or translation efficiencies.

If a difference in expression existed between variants, one would assume that the levels of expression would correspond to a similar pattern of expression in homozygote samples; however **Figure 12** does not correlate to this hypothesis. Expression of the variants across all genotypes seems to be similar with no statistically significant differences. One possible explanation for this could be the small sample size. There is a huge variation in the amount of transcript levels within each population of samples, as indicated by the large error bars. Future studies with a larger sample size would diminish variation and could potentially elucidate differences in RNA expression with higher confidence.

5 Conclusion

This study is the first to attempt to understand the molecular basis of the association of rs11556924 with CAD risk. Here I have established that a conformational or structural difference likely exists between the two NIPA variants (NIPA363Arg and NIPA363His). In addition, there is a slight difference in phosphorylation of a critical serine residue in close proximity to the SNP with the potential to affect cell proliferation through an unknown mechanism. Cell proliferation, the main process NIPA is involved in, is apparently more rapid in cells expressing the risk allele (NIPA363Arg) than in cells expressing the alternative allele (NIPA363His). With the growing body of evidence that supports the idea that cell proliferation plays a key role in atherosclerosis development, the difference in cell proliferation rates supports the hypothesis that rs11556924C/T is affecting CAD risk through differences in cell cycle regulation. Lastly, I have demonstrated that in heterozygote patients, there are higher levels of NIPA363Arg mRNA transcripts than NIPA363His transcripts. This suggests that in addition to protein effects, rs11556924 may affect mRNA stability or translation. However the difference in transcript levels in heterozygotes does not correlate to total NIPA RNA transcript levels across all genotypes. An increase in sample size may reduce the large expression variation seen between samples, and either confirm the difference or identify a new trend of NIPA mRNA levels between rs11556924 genotypes.

Future experiments should focus on crystalizing the structure of NIPA and the p.R363H variant to further investigate potential conformational or structural differences. RNA and protein stability experiments are needed to evaluate potential

explanations for the results presented in this study. Cell proliferation experiments that express NIPA variants at endogenous levels in macrophages and VSMC will be crucial experiments to understand the direct link between rs11556924C/T and CAD development.

In conclusion, the effects of the nonsynonymous SNP rs11556924C/T in ZC3HC1 encoding NIPA, have yet to be fully elucidated. As this is the first study to identify potential molecular mechanisms through which the rs11556924 SNP is affecting CAD development, there are many follow up experiments that are needed to more completely understand the implications of this SNP. This study has not only provided the basis for future experiments on NIPA and the effects of the rs11556924 variant but it has also added to ever-growing body of knowledge of CAD development. GWAS studies are crucial for identifying potential genetic contributors to CAD, however functional analyses, such as the present study, are imperative for understanding the link between SNPs and CAD development.

6 References

1. World Health Oraganization. The 10 leading causes of death in the world, 2000 and 2012. (2014). at <http://www.who.int/mediacentre/factsheets/fs310/en/>

2. Libby, P. & Theroux, P. Pathophysiology of coronary artery disease. *Circulation* **111,** 3481–3488 (2005).

3. Siemelink, M. A. *et al.* Common variants associated with blood lipid levels do not affect carotid plaque composition. *Atherosclerosis* **242,** 351–6 (2015).

4. Tabas, I., Garcia-Cardena, G. & Owens, G. K. The cell biology of disease: Recent insights into the cellular biology of atherosclerosis. *J. Cell Biol.* **209,** 13–22 (2015).

5. Weissberg, P. L. Atherogenesis: current understanding of the causes of atheroma. *Heart* **83,** 247–252 (2000).

6. Rader, D. J. & Daugherty, A. Translating molecular discoveries into new therapies for atherosclerosis. *Nature* **451,** 904–913 (2008).

7. Ross, R. Inflammation or Atherogenesis. *N. Engl. J. Med.* **340,** 115–126 (1999).

8. Weber, C. & Noels, H. Atherosclerosis: current pathogenesis and therapeutic options. *Nat. Med.* **17,** 1410–1422 (2011).

9. Libby, P., Ridker, P. M. & Hansson, G. K. Progress and challenges in translating the biology of atherosclerosis. *Nature* **473,** 317–325 (2011).

10. Bruemmer, D. & Law, R. E. Thiazolidinedione regulation of smooth muscle cell proliferation. *Am. J. Med.* **115 Suppl ,** 87S–92S (2003).

11. Jenkins, S. J. *et al.* Local macrophage proliferation, rather than recruitment from the blood, is a signature of TH2 inflammation. *Science* **332,** 1284–8 (2011).

12. Robbins, C. S. *et al.* Local proliferation dominates lesional macrophage accumulation in atherosclerosis. *Nat. Med.* **19,** 1166–72 (2013).

13. Tang, J. *et al.* Inhibiting macrophage proliferation suppresses atherosclerotic plaque inflammation. *Sci. Adv.* **1,** e1400223–e1400223 (2015).

14. Deloukas, P. *et al.* Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat. Genet.* **45,** 25–33 (2013).

15. Manolio, T. a *et al.* Finding the missing heritability of complex diseases. *Nature* **461,** 747–753 (2009).

16. Hardy, J. & Singleton, A. Genomewide association studies and human disease. *N. Engl. J. Med.* **360,** 1759–1768 (2009).

17. Hirschhorn, J. N. & Daly, M. J. Genome-wide association studies for common diseases and complex traits. *Nat. Rev. Genet.* **6,** 95–108 (2005).

18. Johansen, C. T., Kathiresan, S. & Hegele, R. a. Genetic determinants of plasma triglycerides. *J. Lipid Res.* **52,** 189–206 (2011).

19. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* **11,** 499–511 (2010).

20. Lewontin, R. C. The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. *Genetics* **49,** 49–67 (1964).

21. Schunkert, H. *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat. Genet.* **43,** 333–338 (2011).

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

23. López-Mejías, R. *et al.* The ZC3HC1 rs11556924 polymorphism is associated with increased carotid intima-media thickness in patients with rheumatoid arthritis. *Arthritis Res. Ther.* **15,** R152 (2013).

24. Kunnas, T. & Nikkari, S. T. Association of Zinc Finger, C3HC-Type Containing 1 (ZC3HC1) rs11556924 Genetic Variant With Hypertension in a Finnish Population, the TAMRISK Study. *Medicine (Baltimore).* **94,** e1221 (2015).

25. Ouyang, T. *et al.* Identification and characterization of a nuclear interacting partner of anaplastic lymphoma kinase (NIPA). *J. Biol. Chem.* **278,** 30028–30036 (2003).

26. Bassermann, F. *et al.* NIPA defines an SCF-type mammalian E3 ligase that regulates mitotic entry. *Cell* **122,** 45–57 (2005).

27. Bassermann, F., Peschel, C. & Duyster, J. Mitotic entry: a matter of oscillating destruction. *Cell Cycle* **4,** 1515–1517 (2005).

28. Illert, A. L. *et al.* Extracellular signal-regulated kinase 2 (ERK2) mediates phosphorylation and inactivation of nuclear interaction partner of anaplastic lymphoma kinase (NIPA) at G2/M. *J. Biol. Chem.* **287,** 37997–38005 (2012).

29. Li, J., Meyer, A. N. & Donoghue, D. J. Nuclear localization of cyclin B1 mediates its biological activity and is regulated by phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 502–7 (1997).

30. Pines, J. & Hunter, T. Human cyclins A and B1 are differentially located in the cell and undergo cell cycle-dependent nuclear transport. *J. Cell Biol.* **115,** 1–17 (1991).

31. Yang, J. *et al.* Control of cyclin B1 localization through regulated binding of the nuclear export factor CRM1. *Genes Dev.* **12,** 2131–43 (1998).

32. Li, J., Meyer, A. N. & Donoghue, D. J. Nuclear localization of cyclin B1 mediates its biological activity and is regulated by phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 502–7 (1997).

33. Feldman, R. M. R., Correll, C. C., Kaplan, K. B. & Deshaies, R. J. A Complex of Cdc4p, Skp1p, and Cdc53p/Cullin Catalyzes Ubiquitination of the Phosphorylated CDK Inhibitor Sic1p. *Cell* **91,** 221–230 (1997).

34. Skowyra, D., Craig, K. L., Tyers, M., Elledge, S. J. & Harper, J. W. F-Box Proteins Are Receptors that Recruit Phosphorylated Substrates to the SCF Ubiquitin-Ligase Complex. *Cell* **91,** 209–219 (1997).

35. Cardozo, T. & Pagano, M. The SCF ubiquitin ligase: insights into a molecular machine. *Nat. Rev. Mol. Cell Biol.* **5,** 739–751 (2004).

36. Uro-Coste, E. *et al.* The cell cycle gene SKP1 is regulated by light in postnatal rat brain. *Mol. Brain Res.* **56,** 192–199 (1998).

37. Deshaies, R. J. SCF and Cullin/Ring H2-based ubiquitin ligases. *Annu. Rev. Cell Dev. Biol.* **15,** 435–67 (1999).

38. Bassermann, F. *et al.* Multisite phosphorylation of nuclear interaction partner of ALK (NIPA) at G2/M involves cyclin B1/Cdk1. *J. Biol. Chem.* **282,** 15965–15972 (2007).

39. Von Klitzing, C. *et al.* APC/C Cdh1-mediated degradation of the F-box protein NIPA is regulated by its association with Skp1. *PLoS One* **6,** 1–8 (2011).

40. Skaar, J. R., Pagan, J. K. & Pagano, M. Mechanisms and function of substrate recruitment by F-box proteins. *Nat. Rev. Mol. Cell Biol.* **14,** 369–81 (2013).

41. Wang, Z. & Moult, J. SNPs , Protein Structure , and Disease. **270,** 263–270 (2001).

42. Singh, S. M., Bandi, S., Shah, D. D., Armstrong, G. & Mallela, K. M. G. Missense Mutation Lys18Asn in Dystrophin that Triggers X-Linked Dilated Cardiomyopathy Decreases Protein Stability, Increases Protein Unfolding, and Perturbs Protein Structure, but Does Not Affect Protein Function. *PLoS One* **9,** e110439 (2014).

43. Yue, P., Li, Z. & Moult, J. Loss of protein structure stability as a major causative factor in monogenic disease. *J. Mol. Biol.* **353,** 459–473 (2005).

44. Anoosha, P., Sakthivel, R. & Gromiha, M. M. Prediction of protein disorder upon amino acid substitutions. *Anal. Biochem.* (2015). doi:10.1016/j.ab.2015.08.028

45. Vaughn, L. S. *et al.* Altered Activation of Protein Kinase PKR and Enhanced Apoptosis in Dystonia Cells Carrying a Mutation in PKR Activator Protein PACT. *J. Biol. Chem.* **290,** 22543–57 (2015).

46. Yuan, C.-C. *et al.* Constitutive phosphorylation of cardiac myosin regulatory light chain prevents development of hypertrophic cardiomyopathy in mice. *Proc. Natl. Acad. Sci. U. S. A.* **112,** E4138–46 (2015).

47. Erbilgin, A. *et al.* Identification of CAD candidate genes in GWAS loci and their expression in vascular cells. *J. Lipid Res.* **54,** 1894–905 (2013).

48. Duan, J. *et al.* Synonymous mutations in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor. *Hum. Mol. Genet.* **12,** 205–16 (2003).

49. Capon, F. A synonymous SNP of the corneodesmosin gene leads to increased mRNA stability and demonstrates association with psoriasis across diverse ethnic groups. *Hum. Mol. Genet.* **13,** 2361–2368 (2004).

7 Appendix A

Table 1. Primers used for qPCR and yeast expression experiments.

Purpose	Direction	Primer	Sequence (5' to 3')
PPIA qPCR	FWD	PPIA F	ACCGTGTTCTTCGACATTGC
	RVS	PPIA R	TTCTGTGAAAGCAGGAACCC
NIPA qPCR	FWD	F3-NIPA-RT	TTTGAAAAATTAAGGCTGAGTTC
	RVS	R3-NIPA-RT	AATCCCACACTCCTTTAACAGAAC
His Tagging NIPA for *K. lactis*	FWD	HISNIPAS1	CACCATCACATGGCGGCGCCCTG
	RVS	HISNIPARC1	ATGATGATGCATGGTGGCGGCCGCC
	FWD	HISNIPAS2	AAACTCGAGAAAAGAATGCATCATCATCACCATCACATG
	RVS	HISNIPARC2	TTTGGATCCTCAGCATGAGCACAGAGATTCCC
K. lactis yeast expression	-	Integration Primer 1	ACACACGTAAACGCGCTCGGT
	-	Integration Primer 2	ATCATCCTTGTCAGCGAAAGC
	-	Integration Primer 3	ACCTGAAGATAGAGCTTCTAA

Table 2. Plasmids used for mammalian expression.

Plasmid	Plasmid Type	Name	Description
pCDNA (ThermoFisher)	Mammalian expression	Flag-NIPA363Arg	CMV promoter drives expression of a Flag-tagged ZC3HC1 with C allele at rs11556924
		Flag-NIPA363His	CMV promoter drives expression of a Flag-tagged ZC3HC1 with T allele at rs11556924
		Empty	CMV promoter with no gene insert
pEYFP-C1 (Clonetech)	Mammalian expression	FP-NIPA363Arg	CMV promoter drives expression of a Florence-tagged ZC3HC1 with C allele at rs11556924
		FP-NIPA363His	CMV promoter drives expression of a Florence-tagged ZC3HC1 with T allele at rs11556924
		Empty	CMV promoter with no gene insert

Table 3. Population characteristics of human whole blood RNA cohort.

Characteristics	Genotype		
	CC (n=61)	**CT (n=75)**	**TT (n=23)**
M	31	35	15
F	30	40	8
Age	75.7 ± 5.2	75.6 ± 5.1	73.5 ± 3.6
BMI	26.45 ± 3.0	25.98 ± 2.7	26.14 ± 3.0
Hypertension	50.82%	38.66%	53.33%

8 Appendix B

Tara Linseman

EDUCATION

Masters of Science **2013-2015**
Biochemistry- Specialization in Human and Molecular Genetics
University of Ottawa

Bachelor of Science **2008-2013**
Honors Specialization in Genetics with Internship
The University of Western Ontario

RESEARCH AND INSTRUCTING EXPERIENCE

Masters Research **2013-2015**
Dr. Ruth McPherson, University of Ottawa Heart Institute
University of Ottawa, Ottawa, ON, Canada

- Functional analysis of a coding variant associated with protection against Coronary Artery Disease (CAD).
- Using biochemical and cell proliferative assays to elucidate subtle changes between genetic variants and cell growth.
- Used western blots, qPCR and sequencing analysis to determine the effects of the coding variant on cell proliferation.
- Extracted DNA and RNA from human whole blood samples.

Teaching Assistant **2015**
Luc Poitras, Biochemistry Department
University of Ottawa, Ottawa, ON, Canada

- Oversaw the experimental procedures carried out by 2nd year biochemistry students, resulting in their successful and timely completion of the lab.
- Corrected over 160 lab reports in a timely manner; consulted with students who had problems both writing the report and understanding where they went wrong once marked.

4Th Year Research Thesis Student **2012-2013**
Dr. Robert Hegele, Robarts Research Institute
The University of Western Ontario, London, ON, Canada

- Identified novel common and rare genetic variants associated with CAD in 1600 Canadian aboriginals using Illumina HumanExome BeadChip and linear regression analysis (PLINK/Seq, coding in R).
- Helped conduct genetic testing on clinical patients with abnormal lipid levels; duties included: whole blood DNA extraction, sequencing preparation and mutation analysis.

Packaging Quality Lab Technician (Science Internship) **2011-2012**
Kirk Massey, Labatt Breweries of Canada
London, ON, Canada

- Conducted CO_2 , pH , taste, colour and oxygen quality control tests on packaged beer products to ensure the production of beer was running smoothly.

- While following laboratory SOPs and maintenance procedures, conducted tests, data collection and entry into LIMS in an orderly and timely manner.

Cell Biology Laboratory Assistant **2009-2011**
Dr. Sashko Damjanovski,
The University of Western Ontario, London, ON, Canada

- Ensured proper care of over 40 frogs, abiding by animal rights and care guidelines.
- Maintained laboratory cleanliness through proper glassware washing, autoclaving tips and biological waste all while following proper safety guidelines.
- Prepared solutions and ran various agarose gels for graduate student projects.

KEY SKILLS

- Scientific literature searches and reviews
- Preparing scientific papers, posters and grant proposals.
- Software and computational tools: MS Office (Word, PowerPoint, Excel), GraphPad Prism, Blast, primer design, UCSC genome browser, Illumina Genome Studios.
- Extensive PCR, qPCR, Western Blotting and mammalian cell culture experience.

VOLUNTEER EXPERIENCE

Social Co-Ordinator **2014-2015**
University of Ottawa Heart Institute Trainee Committee

- Attend monthly executive meetings to discuss trainee concerns and ideas.
- Organize, promote and book social events for all trainees at the Heart Institute, including joint events with the rest of the University of Ottawa Faculty of Medicine

Let`s Talk Science Volunteer **2013-2015**
University of Ottawa

- Present age-appropriate science demonstrations to elementary and high school students.
- Explain difficult scientific concepts to students in a way which is comprehendible.
- Present various science-based demonstrations at community out-reach programs.

Mentor Speaker **2011**
St. Joseph's High School, Barrie, ON

- Spoke to Gr.12 science students about university life, what to expect from courses and how to prepare to be successful in university.
- Provided an up-to-date overview of interesting genetic concepts to high school biology students, expanding their knowledge of genetic concepts beyond the out of date curriculum.

Publications and Papers

Linseman TA. Functional analysis of a coding variant in *ZC3HC1* at 7q32.3 associated with protection against Coronary Artery Disease (CAD). Department of Biochemistry, Microbiology and Immunology, University of Ottawa. MSc. Thesis (2015)

Turner A., Nikpay M, Silva A, Lau P, Martinuk A, **Linseman TA**, Soubeyrand S, McPherson R. Functional interaction between *COL4A1/COL4A2* and *SMAD3* risk loci for coronary artery disease. *Atherosclerosis.* **242(2),** 543-52 (2015)

Linseman TA. Identifying genetic variants associated with cardiovascular disease susceptibility in Canadian aboriginals. Department of Biology, The University of Western Ontario. Undergraduate Thesis (2013)

Presentations and Abstracts

Linseman TA. Functional analysis of a coding variant in *ZC3HC1* at 7q32.3 associated with protection against Coronary Artery Disease (CAD). European Society of Human Genetics Conference. Poster Presentation. Glasgow, Scotland. June 2015.

Linseman TA, Soubeyrand S., Lau P., Martinuk A., Turner A. and McPherson R. Functional analysis of a coding variant in *ZC3HC1* at 7q32.3 associated with protection against Coronary Artery Disease (CAD). Ottawa Heart Institute Annual Research Day. Poster Presentation. Ottawa, Ontario. May 2015.

Linseman TA, Soubeyrand S., Lau P., Martinuk A., Turner A. and McPherson R. Functional analysis of a coding variant in *ZC3HC1* at 7q32.3 associated with protection against Coronary Artery Disease (CAD). Ottawa Heart Institute Work in Progress. Oral Presentation. Ottawa, Ontario. November 2014.