MOLECULAR AND BIOLOGICAL MECHANISMS OF THE *ARAP1* TYPE 2 DIABETES LOCUS

Jennifer Renée Kulzer

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum in Genetics and Molecular Biology in the School of Medicine.

Chapel Hill
2014

Approved by:

Karen L. Mohlke

Patrick J. Brennwald

Hans E. Hohmeier

Praveen Sethupathy

ABSTRACT

Jennifer Renée Kulzer: Molecular and Biological Mechanisms
of the *ARAP1* Type 2 Diabetes Locus
(Under the direction of Karen L. Mohlke)

Genome-wide association studies (GWASs) have identified more than 70 loci associated with type 2 diabetes (T2D), but for most, the underlying causal variants, associated genes, and functional mechanisms remain unknown. At a T2D- and fasting-proinsulin-associated locus on 11q13.4, we have identified a functional regulatory DNA variant, a candidate target gene, and plausible underlying molecular and biological mechanisms. We confirmed the existence of a single major association signal between fasting proinsulin and noncoding variants ($p = 7.4 \times 10^{-50}$). Measurement of allele-specific mRNA levels in human pancreatic islet samples heterozygous for rs11603334 showed that the T2D-risk increasing and proinsulin-decreasing allele (C) is associated with increased *ARAP1* expression ($p < 0.02$). By performing transcriptional reporter assays in rodent pancreatic beta cell lines, we determined that the C allele of rs11603334, located at the *ARAP1* P1 promoter, exhibits 2-fold higher transcriptional activity than T allele ($p < 0.0001$). Electrophoretic mobility shift assays demonstrated decreased binding of pancreatic beta cell transcriptional regulator PAX6 to the rs11603334 C allele. Collectively, these data suggest that the T2D-risk allele of rs11603334 could abrogate binding of a complex containing PAX6 and thus lead to increased P1 promoter activity and *ARAP1* expression in human pancreatic islets. We went on to determine that the ARAP1 protein isoform corresponding to *ARAP1* mRNA transcripts transcribed from P1 is a major isoform present in

human islets, and that adenoviral overexpression of this ARAP1 isoform in intact human islets led to decreased levels of glucose-stimulated proinsulin secretion ($p = 0.02$). Using G-LISA GTPase activity assays, we also determined that, through the activity of its ARF-GAP domain, increased ARAP1 expression in human islets decreases the levels of ARF6-GTP ($p = 0.02$), a known regulator of glucose-stimulated insulin secretion in the beta cell. We also showed that ARAP1 and ARF6 partially colocalize in the cytoplasm of dispersed human beta cells. Altogether, this body of evidence demonstrates a connected set of molecular and biological mechanisms that may explain GWAS association of the *ARAP1* locus with proinsulin and T2D.

ACKNOWLEDGEMENTS

There are so many people I would like to thank for supporting me over the course of my Ph.D. First, I would like to thank Karen Mohlke. She is an inspiring scientist and a brilliant writer who has contributed greatly to my scientific development and to the enhancement of my own writing; I am fortunate to have been one of her students. I would also like to thank my committee, Patrick Brennwald, Hans Hohmeier, Praveen Sethupathy, and previous member Mara Duncan, for their insightful contributions, key advice, and wonderful mentorship.

My great appreciation also goes to all of the enthusiastic, hard-working members of the Mohlke Laboratory, past and present. Special thanks to Tamara Roman, Marie Fogarty, Rani Vadlamudi, Maren Cannon, and Jim Davis for all of the theoretical, practical, and technical advice shared during Monday afternoon small group meetings; to Martin Buchkovich, Damien Croteau-Chonka, and Ying Wu for hours of computational and statistical consultation; and to everyone for their kind friendship and buoyant support.

I would like to extend my deep gratitude to all of our many talented scientific collaborators and colleagues around the world who have allowed me the privilege of working alongside them, especially: Rama Arumugam, Michael Boehnke, Francis Collins, Michael Erdos, Christian Fuchsberger, Jeroen Huyghe, Anne Jackson, Johanna Kuusisto, Markku Laakso, Ethan Lange, Leslie Lange, Yun Li, Mario Morken, Peter Schwarz, Michael Stitzel, Richard Watanabe, and everyone collaborating with FUSION, METSIM, and GoT2D.

I would also like to sincerely thank the National Institutes of Health for funding our research, the Integrated Islet Distribution Program and the National Disease Research Interchange for their valuable services, and most especially the human organ donors, without whom this research would not have been possible.

For their invaluable friendship, unmatched scientific enthusiasm, and warm support of both my scientific and personal development, I am forever thankful to Kathy Wojcik and Dallas Donohoe. For their encouraging mentorship early in my scientific career, I will also always be thankful to David Winkler, Megan Spence, and Adrian Hanley.

For their unwavering support and love throughout this process, for extending me their courage, advice, and understanding, and for frequently adding important perspective, I am forever indebted to my very dear family, Gary, Laura, Sheila, Cheryl, Kimberly, Richard, Hazel, Ruth, Bill, Dan, Lori, Chrissy, Travis, Gabi, Danny, Jessica, Danny IV, Ellis, Beth, Eric, Steve, Jhaki, Kathy, all of the Kulzers, Downings, and Tynans, and to all of our many incredible friends across the U.S., who I am proud to know, love, and consider additional members of my family.

More than anyone, I would like to thank my husband and best friend, Matthew Tynan. Thank you, Matt, for supporting my passion and creativity, for many hours of helpful discussion and review, for asking great questions and adding insight, for your enthusiastic company at lab on many late nights and weekends, for helping to keep my light shining (both figuratively and literally), and for all of the hugs.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AEI	allelic expression imbalance
ANOVA	analysis of variance
ARAP1	ARF-GAP with RHO-GAP domain, ankyrin repeat and PH domain 1
ARF	ADP-ribosylation factor
BMI	body mass index
cDNA	complementary DNA
CENTD2	centaurin delta 2; another gene symbol for ARAP1
CEU	Caucasian ancestry
ChIP	chromatin immunoprecipitation
Chr	chromosome
CMRL	Connaught Medical Research Laboratories medium
CMV	cytomegalovirus promoter
DMEM	Dulbecco's Modified Eagle Medium
EGF	epidermal growth factor
EGFP	enhanced green fluorescent protein
ELISA	enzyme-linked immunosorbent assay
EMSA	electrophoretic mobility shift assay
ENCODE	Encyclopedia of DNA Elements
eQTL	expression quantitative trait locus
EUR	European ancestry
FAIRE	formaldehyde-assisted isolation of regulatory elements

FBS	fetal bovine serum
FUSION	Finland-United States Investigation of NIDDM Genetics
GAP	GTPase activating protein
gDNA	genomic DNA
GDP	guanosine diphosphate
GLUT4	glucose transporter 4
GoT2D	Genetics of Type 2 Diabetes
GTP	guanosine triphosphate
GWAS	genome-wide association study
H3K4me1	histone H3 lysine 4 monomethylation
H3K4me3	histone H3 lysine 4 trimethylation
H3K9ac	histone H3 lysine 9 acetylation
HRP	horseradish peroxidase
HWE	Hardy-Weinburg equilibrium
ICR/IIDP	Islet Cell Resource/Integrated Islet Distribution Program
LD	linkage disequilibrium
MAC	minor allele count
MAF	minor allele frequency
MALDI-TOF	matrix-assisted laser desorption/ionization-time of flight
METSIM	Metabolic Syndrome in Men
mRNA	messenger RNA
NDRI	National Disease Research Interchange
NS	non-significant

OGTT	oral glucose tolerance test
P1	*ARAP1* promoter at 72.11 Mb (hg18) on chromosome 11
P2	*ARAP1* promoter at 72.14 Mb (hg18) on chromosome 11
PA	phosphatidic acid
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PFA	paraformaldehyde
PFU	plaque-forming units
PH	pleckstrin homology domain
PI	phosphoinositides
PVDF	polyvinylidene difluoride
RA	Ras-associating domain
RefSeq	National Center for Biotechnology Information Reference Sequence
RHO	Ras homolog
RPMI	Roswell Park Memorial Institute medium
SAM	sterile alpha motif domain
SD	standard deviation
SNP	single nucleotide polymorphism
T2D	type 2 diabetes
TBS	tris-buffered saline
TSS	transcription start site
UCSC	University of California Santa Cruz
UTR	untranslated region

CHAPTER 1: INTRODUCTION

Genome-wide association studies (GWAS) have identified over 70 genomic locations, or loci, associated with type 2 diabetes (T2D), but for most of these loci, the underlying functional mechanisms remain unknown. Naturally occurring human genetic variants at a T2D GWAS locus on chromosome 11 containing *ARAP1*, *PDE2A*, *STARD10*, *ATG16L2*, and *FCHSD2* is associated with T2D risk ($p = 1.4 \times 10^{-22}$) and decreased proinsulin levels ($p = 3.2 \times 10^{-102}$).[1,2] At this '*ARAP1*' locus, robust statistical association of twenty frequently co-inherited single nucleotide polymorphism (SNP) variants with T2D, proinsulin, and several additional glycemic traits have been reported in independent populations,[1–4] providing high confidence that there exist unknown underlying biological explanations worthy of exploration. The aims of my research were to identify the functional variant(s) and gene(s) at the *ARAP1* locus and determine how the gene(s) play a role in proinsulin secretion, the most strongly associated trait, from pancreatic islets. Determining the molecular and biological mechanisms underlying high-confidence GWAS loci like *ARAP1* will uncover new biological pathways involved in diabetes, improve our understanding of the disease, and may lead to the discovery of novel therapeutic targets.

Overview of Type 2 Diabetes (T2D) and Insulin

Diabetes is a severe and expensive disease, affecting 11.3% of Americans age 20 years or older.[5] It is the seventh leading cause of death in the U.S., costing $116 billion each year in direct

medical expenses. The overwhelming majority (90-95%) of those with diabetes have T2D.

T2D develops when glucose levels in the bloodstream are chronically elevated due to insufficient regulation by insulin. Insulin is a peptide hormone that is secreted from pancreatic beta cells in response to increased levels of glucose and other nutrients present in the circulation after eating. Insulin's precursor protein, proinsulin, is translated in the rough endoplasmic reticulum and transported to the Golgi apparatus for modifications and sorting. In the trans-Golgi, proinsulin is packaged into granules, where it is converted to mature insulin.[6,7] There, enzymes cleave each proinsulin molecule to produce one molecule each of insulin and C-peptide. Upon exposure of the beta cell to glucose, the mature insulin-containing granules fuse with the plasma membrane and exocytose their contents. This is called the regulated release pathway, accounting for greater than 99% of the insulin and proinsulin secreted from beta cells.[8] Insulin signals many bodily tissues, such as liver, muscle, and adipose, to clear the glucose from the blood and use some of it for immediate energy while storing the rest as glycogen and fat for later use. In individuals with T2D, this process is defective due to 1) inefficient insulin signaling in liver, muscle, and adipose, commonly referred to as "insulin resistance," and 2) inadequate insulin secretion by the pancreas.

Insulin resistance can develop slowly and be present for many years before glucose levels rise and diabetes develops. During this time, the pancreas produces more insulin to compensate for insulin resistance in an attempt to maintain glucose homeostasis. For many with insulin resistance, the pancreas does this quite successfully and diabetes never develops. However, for some with insulin resistance, the pancreatic beta cells progressively lose their ability to secrete enough insulin, leading to a rise in blood glucose levels (hyperglycemia). Since chronically elevated blood glucose causes complications, and evidence of sustained hyperglycemia is used clinically to diagnose T2D,

the declining ability of beta cells to secrete adequate amounts of insulin often marks the onset of disease.

Poorly managed diabetes results in chronic hyperglycemia, which is damaging to blood vessels and nerves and can lead to complications in several highly vascularized organs. Diabetes is the primary cause of adult blindness, kidney failure, and leg and foot amputations not resulting from trauma.[5] It is also a major cause of painful peripheral neuropathies and male impotence. Macrovascular complications include 2-4 fold increased risk[5] of heart disease death and stroke. Simply having diabetes, without any prior evidence of coronary heart disease (i.e., myocardial infarction, angina, abnormal electrocardiogram changes indicating ischemia), increases risk of death from a coronary event to the same extent as the risk faced by non-diabetics who have already shown signs of coronary heart disease.[9]

The intricate biological mechanisms causing insulin resistance and impaired insulin secretion are still not well understood, despite several decades of research. However, recent major advances in the field of genetics, including mapping of the human genome and discovering the extent of human genetic variation through deep-sequencing, provide an exceptional opportunity to pose new questions about the nature of genetic contribution to diabetes risk. Research thoroughly examining the biological reasons why some individuals are more likely to develop T2D than others will illuminate novel pathways to explore for potential medical interventions and cures.

Genetics of T2D

T2D is a complex trait: its inheritance is influenced by both the environment and the contributions of many genes, each exerting a small individual effect. Obesity is well known to be a major risk factor for the development of T2D; over 80% of persons with T2D are overweight or

obese.[10] Even so, many who are overweight or obese do not become diabetic, and many others develop T2D even though they are lean. This can be partly explained by the strong genetic component of T2D. Population studies in the U.S.[11] and Finland[12] reported proband concordance in monozygotic twins to be 29-58% and 34%, respectively, and concordance increases with age. In dizygotic twins, proband concordance was 14-17% and 16%, compared to population prevalences of 5.7% and 3.3%, respectively.

T2D is variably penetrant and has high genetic heterogeneity; many individual genetic loci associated with T2D have low penetrance and small effect sizes. Although the disease symptoms manifest as one common disorder, the inheritance pattern does not fit a simple Mendelian model. However, on a molecular level, individual T2D-associated variants may act additively, such that risk for disease rises with each additional risk allele one carries.

To systematically identify genetic loci associated with T2D risk, human geneticists have performed genome-wide association studies.[1,13–15] GWAS is a powerful, unbiased, population-based method that usually uses additive modeling of common variants to discover novel loci statistically associated with a trait. To date, over 70 loci associated with T2D have been identified through GWAS; the first 50 identified were estimated to explain about 10% of the inheritance.[1] For many of the loci, the underlying variant(s) and gene(s) contributing to T2D risk remain unknown.

Multiple Trait Associations at the *ARAP1* Locus

The *ARAP1* locus is robustly associated with T2D ($p = 1.4 \times 10^{-22}$)[1] and has the strongest association with fasting proinsulin of all loci across the genome ($p = 3.2 \times 10^{-102}$).[2] The most strongly associated 'index' SNPs are the same for both traits, implying that a shared set of functional mechanisms may underlie both associations. These index SNPs, rs1552224 and rs11603334, are

located 113 bp apart within the 5' untranslated region (UTR) of the short RefSeq (Reference Sequence) isoform of *ARAP1* (*ARAP1*Δ*SAM*). These SNPs are always inherited together; they are in perfect linkage disequilibrium (LD; r^2=1, European ancestry (EUR) population, 1000 Genomes Phase 1). All other SNPs inherited in a similar pattern ($r^2 > 0.8$) with rs1552224 and rs11603334 are non-coding. The six genes closest to the association signal are *ARAP1*, *STARD10*, *PDE2A*, *ATG16L2*, *FCHSD2*, and mir-139.

The index SNPs are also associated with additional glycemic traits (Table 1), including partially processed ('split') proinsulin, measures of insulin secretion such as the insulinogenic index, and insulin 30 min into an oral glucose tolerance test (OGTT). Evidence of associations with several traits more clearly linked to insulin secretion than to insulin resistance suggests that the *ARAP1* locus may influence T2D risk by altering pancreatic beta cell function.

Table 1. The Index SNPs at the *ARAP1* Locus Are Associated with Multiple Glycemic Traits

Trait	p-value	Effect Size	*n*	Direction of Effect rs11603334		Ref.
				T2D-Risk Allele (C)	Non-Risk Allele (T)	
Type 2 Diabetes	1.4×10^{-22}	1.14 *(OR)*	101,290	-	-	1
Fasting Intact Proinsulin	3.2×10^{-102}	-0.0928 *(β)*	26,599	Decreased	Increased	2
32,33-Split Proinsulin	1.2×10^{-25}	-0.146 *(β)*	6,343	Decreased	Increased	2
Insulinogenic Index	2×10^{-6}	-6.92 %	5,529	Decreased	Increased	3
30 min OGTT Insulin	0.001	-4.19 %	5,529	Decreased	Increased	3
30 min OGTT Glucose	2×10^{-5}	1.97 %	5,529	Increased	Decreased	3
Fasting Glucose	1.20×10^{-4}	0.0193 *(β)*	46,263	Increased	Decreased	1
Fasting Insulin	0.015	-0.0126 *(β)*	38,413	Decreased	Increased	1

Notably, the T2D-risk increasing alleles for the index SNPs are the major alleles (rs11603334, C, minor allele frequency (MAF) = .13, CEU). According to the Hardy-Weinberg formula, 76% of Caucasians (CEU) are predicted to be homozygous for the T2D-risk increasing allele, 23% are T2D-

risk allele carriers, and only 2% are homozygous for the non-risk allele. From the opposite perspective, individuals homozygous for the non-risk allele may also be considered to be "protected" from the increased risk of T2D conferred by rs11603334 C allele carrier status that is inherited in 98% of Caucasians.

Genes at the *ARAP1* Locus

ARAP1, *STARD10*, *PDE2A*, *ATG16L2*, and *FCHSD2* are all expressed in human pancreas, islets, and flow-sorted beta cells.[1,2,16,17] None of these genes can be clearly implicated in T2D based on what has been reported in the literature to date, so observations linking any of these genes to T2D biology will be novel. ARAP1 (ARF-GAP with RHO-GAP domain, ankyrin repeat and PH domain 1) is a GTPase activating protein (GAP) that activates ADP-ribosylation factor (ARF) and Ras homolog (RHO) family GTPases, which regulate Golgi transport, membrane trafficking, and actin cytoskeleton dynamics.[18] STARD10 binds and transfers phospholipids between membranes.[19] PDE2A is a cyclic nucleotide phosphodiesterase which degrades second messengers cGMP and cAMP.[20] ATG16L2 shares sequence homology with ATG16, which is required for autophagy in yeast.[21] FCHSD2 is named after its FCH and SH3 protein domains, and stimulates F-actin assembly in vitro.[22,23] Mir-139 downregulates FOXO1 in mouse neonatal hepatocytes.[24]

Although rs1552224 has been reported previously as an expression quantitative trait locus (eQTL) for *STARD10* expression in blood ($p = 8.6 \times 10^{-7}$),[1] liver ($p = 1.79 \times 10^{-5}$) and omental adipose tissue ($p = 1.21 \times 10^{-6}$),[25] the importance of these associations to T2D is unclear. The top SNP (rs519790, $p = 2.7 \times 10^{-24}$) associated with *STARD10* expression in blood is in low LD ($r^2 = 0.04$) with rs1552224, indicating that the association signals representing T2D risk and *STARD10* expression are probably not the same.

ARAP1 Protein Structure, Function, and Substrates

ARAP1 may be the strongest candidate gene because the index SNPs lie within one of its alternative 5'UTRs, which is used by two out of three RefSeq isoforms. Sequence homology predicts all three RefSeq protein isoforms of ARAP1 (NP_056057, NP_001128662, and NP_001035207) to contain one ARF-GAP domain, one RHO-GAP domain, one Ras-association (RA) domain, and five pleckstrin homology (PH) domains.[26–29] GAP domains catalyze the guanosine triphosphate (GTP) hydrolysis activity of small GTPases. GTP-bound forms of GTPases such as ARF and RHO are considered to be active because they bind to effectors, while guanosine diphosphate (GDP)-bound forms are mostly inactive. ARF and RHO GTPases help regulate a variety of cellular processes, including vesicle trafficking and actin cytoskeleton reorganization. RA domains bind GTPases. PH domains can bind phosphoinositides (PI) in lipid membranes, and they can also interact with GTP-bound small GTPases.[30] Alternatively spliced out exons in the ARF-GAP domain and in the C-terminal PH domain of ARAP1 RefSeq isoform NP_001128662 create differences in these domains.

The longest isoform of ARAP1, NP_001035207, which does not contain the index SNPs, has an additional sterile alpha motif (SAM) domain at its alternative 5' end, as do homologs ARAP2 and ARAP3. SAM domains interact with other proteins to facilitate many biological functions.[31] In contrast to the shorter isoform (ARAP1ΔSAM), the long isoform of ARAP1 does not catalyze ARF6-GTP hydrolysis, suggesting that the SAM domain may function as a negative regulator.[32] ARAP2 additionally differs from the long isoform of ARAP1 and from ARAP3 in that it contains an R>Q substitution at a catalytic residue of the RHO-GAP domain and lacks RHO-GAP activity.[18,33]

The ARAPs are the only known mammalian proteins containing both ARF-GAP and RHO-GAP domains.

Though *ARAP1* messenger RNA (mRNA) is expressed ubiquitously, ARAP1 protein expression and localization may be more variable.[16–18] Endogenous ARAP1 protein localizes near the Golgi in resting HeLa cells[18,34] and relocates to the cytoplasm and at the plasma membrane in spreading HeLa cells and NIH3T3 mouse fibroblasts.[18] A detailed study in Hela cells demonstrates that endogenous ARAP1 is positioned at the *trans* Golgi and with recycling endosomes, multivesicular bodies, and lysosomes.[35] Overexpressed ARAP1 is present throughout the cytoplasm in HeLa[32,35] and NCTC immortalized human keratinocytes.[36] Upon cell stimulation with either epidermal growth factor (EGF)[34] or TNF-related apoptosis inducing ligand[36], endogenous and/or overexpressed ARAP1 relocates to foci along the plasma membrane containing the corresponding receptors. The PH domains also play an important role in ARAP1's cellular localization. Each PH domain has varying binding affinities to phosphorylated PI *in vitro*. Treatment of NRK cells with PI-3 kinase inhibitors shifts ARAP1's cellular distribution towards the Golgi, while PI-4 kinase inhibitors shift ARAP1's cellular distribution towards the endosomal compartment.[35]

ARAP1 plays a role in cytoskeletal reorganization[18] and in the endocytosis and trafficking of certain receptors. Conflicting reports suggest that overexpression[35] and/or knockdown[34] of ARAP1 in HeLa cells leads to increased association of EGF and EGF receptor with early endosomes, increased EGF receptor degradation and reduced activation of the downstream Erk pathway. ARAP1 catalyzes GTP hydrolysis for ARF GTP-binding proteins ARF1, ARF5, and ARF6 *in vitro*[18] and for ARF6 but not ARF1 in BHK cells.[32] *In vitro* and in BHK cells, a R338K substitution in the ARF-GAP domain substantially decreases hydrolysis of GTP-bound ARF1 and ARF6, respectively. The RHO-GAP domain is not required for ARF-GAP activity. However, *in vitro* ARF-GAP activity

is dependent on the presence of phosphorylated PI; PI 3,4,5 triphosphate (PI(3,4,5)P3) is most stimulatory, followed by PI(3,4)P2. ARAP1 also catalyzes *in vitro* GTP hydrolysis for RHO GTPases RHOA and CDC42, while a R753K substitution in ARAP1's RHO-GAP domain decreases CDC42 GTPase activity.[18] Over-expressed ARAP1 catalyzes CDC42 GTPase activity in 293T cells.[18]

ARAP1 has been demonstrated to regulate GTP hydrolysis *in vitro* for at least five GTP-binding proteins: ARF1, ARF5, ARF6, CDC42, and RHOA.[18] The literature suggests plausible biological roles in insulin secretion for four of these. ARF1 might be involved early in the process of secretion as it is known to regulate transport between the endoplasmic reticulum and Golgi by recruiting coat proteins that allow vesicle budding when bound to GTP.[37] ARF5 associates with the Golgi apparatus[38] and may be involved in regulated secretion. ARF5-GDP binds to calcium-dependent activator protein for secretion 1 (CAPS1), a protein involved in secretion of dense core vesicles, in mouse cerebellum lysates.[39] This interaction is required for proper KCl-stimulated release of a secretory protein called chromogranin A from the PC12 rat adrenal medulla cell line.[39] ARF6 localizes to the plasma membrane and is well known to regulate both clathrin-dependent and independent endocytosis as well as membrane recycling, and may play a role in regulated secretion for certain cell types.[37,38] ARF6 is activated upon glucose stimulation in 832/13 cells, and cells expressing a dominant negative ARF6 mutant show reduced glucose-stimulated insulin secretion.[40] MIN6 cells expressing the same ARF6 mutant have reduced glucose- and K+-stimulated insulin secretion.[41] CDC42 is important for glucose-stimulated insulin secretion in MIN6 cells and for the second phase of insulin secretion in mouse islets.[42] Both ARF6 and CDC42 become activated within 3 minutes of glucose stimulation.[40,42] CDC42 also plays a role in insulin-stimulated glucose

transporter 4 (GLUT4) translocation to the plasma membrane of 3T3-L1 mouse adipocytes.[43] RHOA regulates actin polymerization;[44] however, a potential role for it in insulin secretion is unclear.

In this dissertation, I investigate and describe functional molecular and biological mechanisms that may underlie the genome-wide association of the 11q13.4 *ARAP1* locus with proinsulin levels and T2D. In Chapter 2, I identify a plausible functional regulatory variant, rs11603334, and demonstrate its allelic effects on expression of a specific isoform of *ARAP1*. In Chapter 3, I characterize localization and isoform expression of ARAP1 protein in the human pancreatic islet, demonstrate that the ARAP1 isoform implicated in Chapter 2 can regulate glucose-stimulated islet secretion of proinsulin, and show that this ARAP1 isoform regulates ARF6-GTP levels in an ARF-GAP domain-dependent manner. Finally, in Chapter 4, I discuss the implications of my research results and consider the possible role of altered ARAP1 expression in T2D risk.

CHAPTER 2: A COMMON FUNCTIONAL REGULATORY VARIANT AT A TYPE 2 DIABETES LOCUS UPREGULATES *ARAP1* EXPRESSION IN THE PANCREATIC BETA CELL[i,ii]

Introduction

Genome-wide association studies have identified more than 70 loci associated with type 2 diabetes [MIM 125853][1,45–52] and 10 loci associated with fasting proinsulin levels.[2,53] A locus at 11q13.4 near *ARAP1* [MIM 606646], *PDE2A* [MIM 602658], *STARD10*, *ATG16L2*, and *FCHSD2* is strongly associated with T2D (rs1552224, $p = 1.4 \times 10^{-22}$),[1] fasting proinsulin (rs11603334, $p = 3.2 \times 10^{-102}$),[2] and 32,33-split proinsulin (rs11603334, $p = 1.2 \times 10^{-25}$).[2] This locus is also nominally associated with insulinogenic index (rs1552224, $p = 2 \times 10^{-6}$) and both insulin ($p = 0.001$) and glucose ($p = 2 \times 10^{-5}$) levels at 30 minutes during an oral glucose tolerance test.[3] The clustering of multiple phenotypic associations related to proinsulin processing and insulin secretory response during an oral glucose challenge suggests that the affected target gene(s) may play a role in pancreatic beta cell function. Currently, the functional DNA variant(s), the affected gene(s), and the underlying molecular genetic mechanism(s) contributing to these associations are unknown.

---

[i]This chapter previously appeared as an article in the *American Journal of Human Genetics*. The original citation is as follows: **Kulzer JR**, Stitzel ML, Morken MA, Huyghe JR, Fuchsberger C, Kuusisto J, Laakso M, Boehnke M, Collins FS, Mohlke KL. "A common functional regulatory variant at a type 2 diabetes locus upregulates *ARAP1* expression in the pancreatic beta cell." *American Journal of Human Genetics*, 94(2):186-97 (2014).

[ii]Jennifer Kulzer designed, performed, and analyzed the allelic expression imbalance assays, dual-luciferase transcriptional reporter assays, and electromobility shift assays; prioritized regulatory variants and transcription factors; and wrote the manuscript. Coauthors designed, performed, and analyzed the OGTT clinical studies, genotyping, fine-mapping and conditional analyses and allelic expression imbalance assays; and reviewed and edited the manuscript.

SNPs rs1552224 and rs11603334 are in perfect LD with each other ($r^2 = 1$, EUR) and are located within the first exon and 5'UTR of *ARAP1* RefSeq isoforms NM_015242 and NM_001135190, at +305 nt and +418 nt, respectively, from the transcription start site (TSS) at 72.11 Mb (hg18) on chromosome 11. A third *ARAP1* RefSeq isoform, NM_001040118, is expressed from an alternative promoter and TSS located at 72.14 Mb. We designate the promoter at 72.11 Mb as P1 and the promoter at 72.14 Mb as P2.

We hypothesized that functional variant(s) at this locus may be in high LD ($r^2 \geq 0.8$) with rs1552224 and rs11603334. None of the variants in high LD with these SNPs are within gene coding regions where they could alter protein function, suggesting that functional common SNP(s) at this locus may influence gene regulation. Genes within the LD region containing rs1552224 and rs11603334 have reported expression in human pancreas, islets, and flow-sorted beta cells;[1] however, islet expression of *ATG16L2*, *FCHSD2*, and *PDE2A* may not be above background.[16,17,54] None of these genes have been demonstrated to play roles in insulin processing or secretion. ARAP1 is a GTPase activating protein (GAP) that activates ARF and RHO family GTPases, which regulate membrane trafficking and actin cytoskeleton reorganization.[18] STARD10 binds and transfers membrane phospholipids.[19] PDE2A is a cyclic nucleotide phosphodiesterase, degrading second messengers cGMP and cAMP.[20] ATG16L2 shares sequence homology with ATG16, a protein required for autophagy in yeast.[21] FCHSD2 is named for its FCH and SH3 protein domains.[22]

Here, we show data supporting rs11603334 as a functional variant regulating *ARAP1* expression. Dense fine-mapping data and conditional analyses support a single association signal. We demonstrate that the T2D-risk allele (C) of rs11603334 is associated with increased *ARAP1* transcript levels in primary human pancreatic islets, disrupts binding of a protein complex

containing transcriptional regulators, and increases transcriptional activity at the *ARAP1* P1 promoter. These data suggest that increased *ARAP1* P1 promoter activity and *ARAP1* expression may be molecular consequences of the T2D-associated variants in this region.

Materials and Methods

Study Population and Phenotypes

We included 8,635 Finnish men without diabetes and not taking diabetes medication (mean age 57.2 years (range 45-74 years); mean body mass index (BMI) 26.8 kg/m^2 (range 16.2-51.6 kg/m^2); fasting plasma glucose levels < 7 mmol/l and 2-hour oral glucose tolerance test plasma glucose levels < 11.1 mmol/l) from the population-based Metabolic Syndrome in Men (METSIM) study[55] in the proinsulin association analyses. Fasting plasma-specific proinsulin (Human Proinsulin RIA kit, Linco Research, St. Charles, MO; no cross-reaction with insulin or C-peptide) and fasting insulin (ADVIA Centaur Insulin IRI, 02230141, Siemens Medical Solutions Diagnostics, Tarrytown, NY; minimal cross-reaction with proinsulin or C-peptide) were measured by immunoassay. For the T2D association analyses, 1,389 T2D cases and 5,748 normoglycemic controls were included. The study was approved by the ethics committee of the University of Kuopio and Kuopio University Hospital and informed consent was obtained from all study participants.

Genotyping

Samples were genotyped with the Illumina HumanOmniExpress BeadChip. In sum, 681,789 autosomal SNPs passed quality control (Hardy-Weinberg equilibrium, HWE $p \geq 1 \times 10^{-6}$ in the total sample; call frequency ≥ 0.97). The same samples were previously

genotyped using the Illumina HumanExome BeadChip,[53] which focuses on protein-altering variants selected from more than 12,000 exome and genome sequences.

Imputation and Reference Panel

A two-step imputation strategy was used, wherein samples were pre-phased using ShapeIT version 2[56] before genotypes were imputed with Minimac.[57] To increase imputation quality, 5,474 haplotypes from the 2,737 Central-Northern European samples sequenced within the Genetics of Type 2 Diabetes (GoT2D) study were used as a reference panel (C. Fuchsberger et al., and J. Flannick, C. Fuchsberger, et al., 2012, Am. Soc. Hum. Genet., abstracts).

Fine-Mapping and Conditional Analysis

SNP associations with log-transformed fasting plasma proinsulin levels were tested assuming an additive genetic model, using a linear mixed model with an empirical kinship matrix to account for relatedness as implemented in EMMAX,[58] and with adjustment for age, BMI, and log-transformed fasting plasma insulin. Both raw residuals and rank-based inverse-normal transformed residuals were analyzed to assess robustness of rare-variant associations with outliers. Genotyped variants with minor allele count (MAC) ≥ 5 (MAF ~0.03%); HWE $p \geq 1 \times 10^{-6}$; and imputed variants with imputation quality score Rsq > 0.3 and MAF $\geq 0.03\%$ were included in the analyses. To identify additional independent signals in the region, conditional analyses were performed on previously reported lead SNP rs11603334 or on our fine-mapped lead SNP rs7109575, using allele count (genotyped variants) or allelic dosage (imputed variants) as a covariate in the model. SNP associations with T2D were tested similarly

but with adjustment for age only. Kang et al.[58] describe and motivate the use of a linear mixed model for analysis of a binary outcome.

Allelic Expression Imbalance (AEI)

Human islets from non-diabetic organ donors were obtained from the National Disease Research Interchange (NDRI) and the Islet Cell Resource/Integrated Islet Distribution Program (ICR/IIDP). DNA and RNA were obtained from 87 primary human pancreatic islet samples using the PureGene (DNA) or RNAeasy kits (RNA) (Qiagen, Gaithersburg, MD). Reverse transcription of the RNA was performed using the Superscript III first-strand synthesis system (Life Technologies, Grand Island, NY). Common, transcribed SNPs (minor allele frequency ≥ 0.1) with the highest LD values with previously reported lead SNPs rs11603334 and rs1552224 were selected for testing allelic expression of *ARAP1*, *STARD10*, *PDE2A*, and *FCHSD2* (Table 2). LD plots including the transcribed SNPs were created using the HapMap genome browser (Figure 7).[59] High quality genomic DNA (gDNA) and mRNA were available for five islet samples heterozygous for rs11603334 and rs1552224; one additional high quality heterozygous sample each was available for mRNA only or gDNA only. For each transcribed SNP, the relative proportions of each allele comprising the gDNA and complementary DNA (cDNA) were quantified using Sequenom iPLEX matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry-based genotyping (Sequenom, San Diego, CA). Primers were designed using MassARRAY (Sequenom) (Table 3). The primers were designed within a single exon to avoid amplicon size differences between gDNA and cDNA, except for rs2291289, which is located near an exon/intron boundary. To control for assay variation, the proportions of each SNP allele measured in the cDNA were analyzed relative to measurements taken in the gDNA. Data are

reported as percent of total gDNA or cDNA containing a given transcribed SNP allele. Statistical significance of differences in allelic representation was determined based on LD scenarios as previously described.[60] Briefly, two-sided t tests were used when the transcribed SNP and lead SNPs were in perfect LD (D' = 1.0, r^2 = 1.0), or when the transcribed SNP and lead SNPs were in intermediate LD with low pairwise correlation (D' ≈ 1.0, r^2 < 0.2). For all t tests, F tests were first used to determine equal or unequal variance. Non-parametric Wilcoxon pair-wise tests were used instead of t tests when gDNA or cDNA measurement data were not normally distributed. One-sided F tests were used when the transcribed SNP and lead SNPs were in low LD (D' < 0.2, r^2 < 0.2).

Regulatory Variant Selection

All known variants in high LD ($r^2 \geq 0.8$, EUR) with rs11603334 and rs1552224 were identified using EUR data from Phase 1 of the 1000 Genomes project.[61] Variants were prioritized based on location relative to regions of potential regulatory function, marked by the following: open chromatin in primary human pancreatic islets detected by formaldehyde-assisted isolation of regulatory elements (FAIRE)[62] and DNase hypersensitivity,[63] accessible chromatin in primary human pancreatic islets detected by chromatin immunoprecipitation (ChIP)-seq for active histone H3 lysine methylation or acetylation modifications (H3K4me3, H3K4me1, and H3K9ac) as reported in the Human Epigenome Atlas,[64] and transcription factor binding in Encyclopedia of DNA Elements (ENCODE) tissues or cell lines detected by ChIP-seq.[65]

Cell Culture

MIN6 mouse insulinoma beta cell line[66] was grown in Dulbecco's Modified Eagle's medium (DMEM) culture medium (Sigma, St. Louis, MO) supplemented with 10% fetal bovine serum (FBS), 1 mM sodium pyruvate, and 100 μM 2-mercaptoethanol. INS-1-derived 832/13 rat insulinoma beta cell line[67] (a gift from C.B. Newgard, Duke University, Durham, NC) was grown in Roswell Park Memorial Institute (RPMI) 1640 culture medium (Corning Cellgro, Mannassas, VA) supplemented with 10% FBS, 2 mM L-glutamine, 1 mM sodium pyruvate, 10 mM Hepes, and 0.05 mM 2-mercaptoethanol. Both lines were maintained at 37°C and 5% CO_2. One day prior to transfection, MIN6 cells were seeded at a density of 200,000 cells per well on 24-well plates. 832/13 cells were seeded at a density of 300,000 cells per well on 24-well plates, or 150,000 cells per well on 48-well plates.

Dual-Luciferase Transcriptional Reporter Assay

Genomic regions including variant positions were polymerase chain reaction (PCR)-amplified from human gDNA using 5 PRIME Mastermix (5 PRIME, Inc., Gaithersburg, MD) or Expand High Fidelity PCR System (Roche, Indianapolis, IN) with the primers listed in Table 4. Amplified DNA was subcloned using XhoI, KpnI, and/or NheI restriction enzymes and T4 DNA Ligase (New England Biolabs, Inc., Ipswich, MA) into the multiple cloning site upstream of the firefly luciferase gene in the pGL4.10 promoterless vector (Promega, Madison, WI). Site-directed mutagenesis was performed using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent, Santa Clara, CA) and the primers listed in Table 4. For each construct, two to ten independent clones were selected. The fidelity and genotype of each gDNA construct was verified by sequencing. Equal amounts, between 600 and 800 ng, of each gDNA construct or

empty pGL4.10 promoterless vector were co-transfected with 80 ng of Renilla luciferase vector using Lipofectamine 2000 (Life Technologies, Grand Island, NY) into duplicate wells of MIN6 or 832/13 cells. Transfected cells were incubated at 37°C and 5% CO_2 overnight and the transfection media was replaced with fresh culture media after 24 hours. Forty-eight hours post-transfection, cell lysates were collected and assayed for luciferase activity using the Dual-Luciferase Reporter Assay System (Promega). Firefly luciferase activity was normalized to that of Renilla luciferase to control for differences in transfection efficiency. Normalized data are reported as the fold change (± standard deviation, SD) in relative activity compared to that of the empty pGL4.10 promoterless vector control. Two-sided t tests were used to compare luciferase activity between genotypes or haplotypes. F tests were used to determine equal or unequal variance. Two-way analysis of variance (ANOVA) was used for simultaneous comparisons of haplotypes resulting from two independently tested SNPs.

Electrophoretic Mobility Shift Assay (EMSA) and Transcription Factor Prediction

Four sets of complementary 21-mer oligonucleotides centered on rs11603334 (C/T) or rs1552224 (T/G) were generated, containing either of the two alternate alleles, both with and without biotin end-labeling (Integrated DNA Technologies, Coralville, IA). Each set was annealed to create double-stranded oligos by incubating 50 pmol of each single-stranded oligo together in buffer containing 10 mM Tris, 1 mM EDTA, and 50 mM NaCl at 95°C for 5 minutes, followed by gradual cooling to 4°C. Double-stranded oligo sequences are specified in Table 5. Nuclear protein lysates were extracted from MIN6 and 832/13 cell pellets with NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo Scientific, Waltham, MA) and protein concentrations were determined using Pierce BCA Protein Assay (Thermo Scientific). EMSAs

were carried out with the LightShift Chemiluminescent EMSA Kit (Thermo Scientific) according to the manufacturer's instructions. Each protein-DNA binding reaction contained 1X binding buffer, 1 μg poly(dIdC), 4 μg nuclear protein lysates, and 100 fmol biotin-labeled double-stranded oligo in 20 μl total reaction volume. For samples demonstrating DNA competition, 60-fold excess unlabeled double-stranded oligo was pre-incubated with nuclear protein lysates in the reaction mixture for 20 minutes before the addition of biotin-labeled oligo. For EMSA reactions with supershift, 4 μg antibody was added to the final reaction mixture and the samples were incubated for an additional 30 minutes. JASPAR,[68] TRANSFAC,[69] CONSITE,[70] PWMSCAN,[71] and Tfsitescan[72] were used to computationally predict transcription factors that may differentially bind at rs11603334 by contrasting the predictions and/or scores generated for each SNP allele. Transcription factor ChIP-seq data for the region overlapping rs11603334 were obtained from the ENCODE project.[65] Antibody to PAX6 (PRB-278P) was purchased from Covance (Princeton, NJ). Antibodies to PAX4 (M-45X, sc-98942X) and AP-2β (H-87X, sc-8976X) were purchased from Santa Cruz Biotechnology (Dallas, TX). Antibodies to HIF1α (MA1-516) and HIF1β (MA1-515) were purchased from Thermo Scientific. Additional antibodies to the following factors were also purchased and tested for supershift: YY1 (39071, 39345), Active Motif (Carlsbad, CA); NFIC (ab86570), Abcam (Cambridge, UK); AP-2α (C-18, sc-184X), YY1 (C-20, sc-281X; H-414, sc-1703X), MAX (C-124, sc-765X), p53 (DO-1, sc-126X), OCT3/4 (C-10, sc-5279X), SP1 (E-3, sc-17824X), USF-1 (C-20, sc-229X), EGR-1 (C-19, sc-189X), RXRα (D-20, sc-553X), E4BP4 (V-19, sc-9549X), PDX1 (N-18, sc-14662X), CEBPβ (C-19, sc-150X), CUX1 (C-20, sc-6327X), and p300 (C-20, sc-585X), Santa Cruz Biotechnology. Protein-DNA complexes were separated by electrophoresis on 6% DNA retardation gels (Life Technologies, Grand Island, NY) in 0.5X TBE

(Lonza, Rockland, ME), transferred to nylon membranes (Thermo Scientific), and UV cross-linked before proceeding with chemiluminescent detection.

Results

Fine-Mapping of the 11q13.4 Locus Associated with Fasting Proinsulin and T2D

To fine-map the fasting proinsulin association signal, we performed association analyses in 8,635 Finnish subjects from the METSIM study using SNPs imputed from a reference panel of 2,737 sequenced Central-Northern European individuals and SNPs directly genotyped on an exome array.[53] The lead SNP was rs7109575 ($p = 7.4 \times 10^{-50}$; n = 8,635), a non-coding variant located at the *ARAP1* P2 TSS (Figure 1, Table 6) and in high LD ($r^2 = 0.94$) with previously reported lead proinsulin-associated SNP rs11603334[2] ($p = 2.2 \times 10^{-48}$; n = 8,630). Conditional analyses using rs11603334 or rs7109575 as covariates largely attenuated the proinsulin association signal (all SNPs, $p > 1.0 \times 10^{-5}$; Figure 1, Figure 8, Table 6, and data not shown). One rare missense variant genotyped in a 1 Mb region on the exome array was associated ($p < 0.05$) with proinsulin after conditional analysis (rs202137453, MAF = 0.003, p = 0.015, $p_{conditional\text{-}rs11603334} = 8.5 \times 10^{-5}$, Table 6). Association with T2D using 1,389 case and 5,748 control subjects (rs7109575, $p = 2.2 \times 10^{-3}$; rs11603334, $p = 3.6 \times 10^{-3}$) was also attenuated after conditional analysis ($p > 0.05$; Table 6, Figure 9, and data not shown). Taken together, fine-mapping and conditional analyses provide evidence supporting a functional regulatory role for common variant(s) in high LD with rs7109575 and rs11603334.

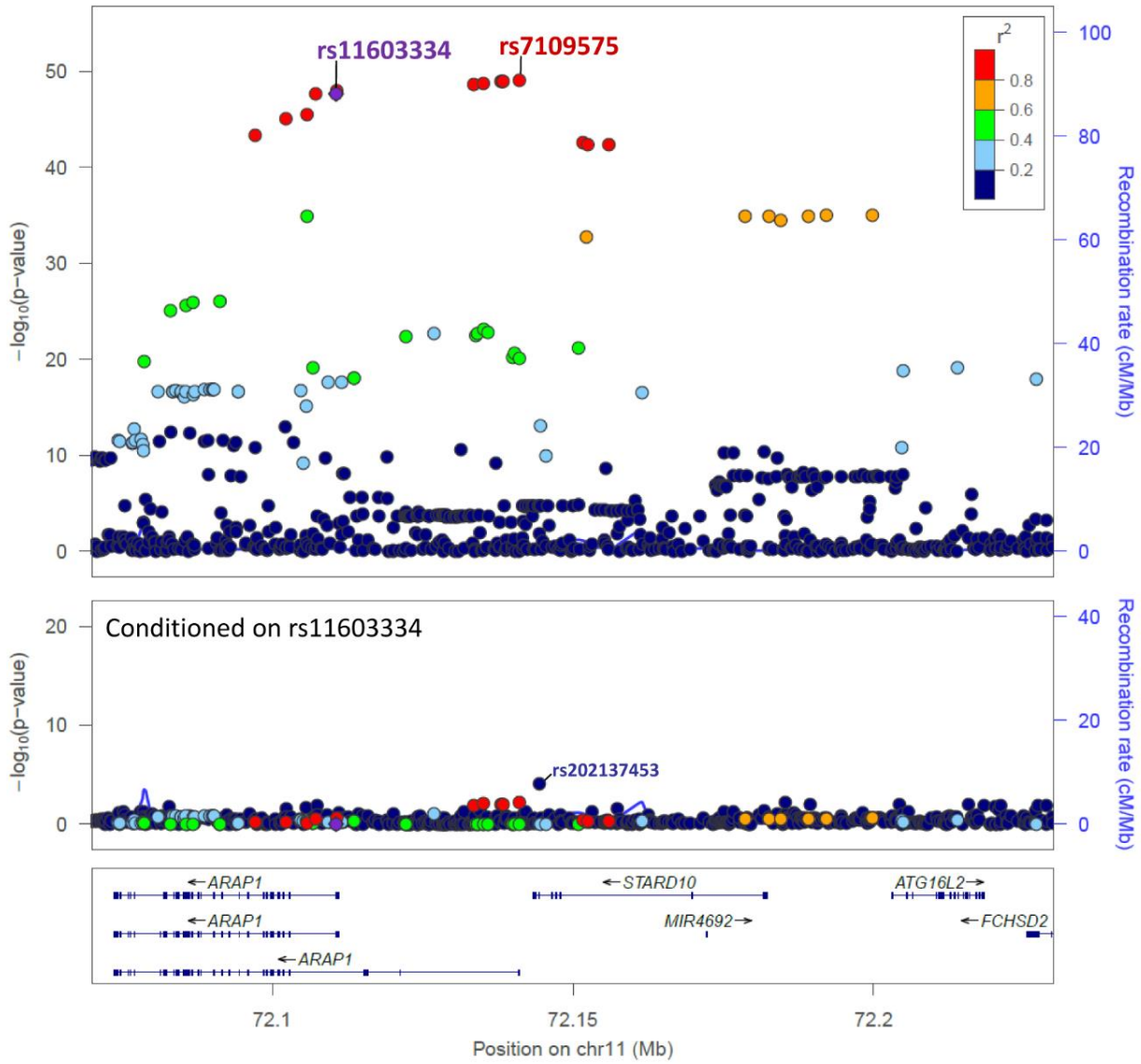

Figure 1. Variants Associated with Proinsulin in the METSIM Study
LocusZoom[73] plots of fasting proinsulin association adjusted for age, BMI, and fasting insulin. The robust association with proinsulin levels (top panel: rs7109575, $p = 7.4 \times 10^{-50}$) is attenuated upon inclusion of rs11603334 as a covariate (bottom panel, only showing y axis to 10^{-20}). Each SNP is colored according to its LD (r^2) with rs11603334. Genomic coordinates refer to hg18 (UCSC Genome Browser).

The Proinsulin-Decreasing and T2D-Risk Alleles are Associated with Increased *ARAP1* mRNA Levels in Primary Human Pancreatic Islets

To determine if expression of *ARAP1*, *STARD10*, *PDE2A*, or *FCHSD2* is associated with the T2D- and fasting proinsulin-associated lead SNPs, we evaluated allelic expression imbalance (AEI) in primary human pancreatic islets. AEI maximizes sensitivity and power to detect modest *cis*-acting effects on mRNA expression through quantification of allele-specific transcript levels within heterozygous samples, thereby controlling for differential *trans*-acting and environmental exposures between individuals[74] as well as differential sample purity and viability between islet preparations.[75,76] The allelic composition of the total transcript levels for each gene was quantitatively determined with Sequenom MALDI-TOF mass spectrometry, a sensitive approach that conserves limited biological sample, and statistically analyzed relative to the allelic composition of the gDNA (Methods). Located in the first exon and 5'UTR of *ARAP1* isoforms transcribed from P1, previously reported lead SNPs rs11603334 and rs1552224 demonstrated AEI (Figure 2). In six samples heterozygous for rs11603334 and rs1552224, we observed *ARAP1* expression differences of 7 to 12% between alleles. For both SNPs, the T2D-risk alleles were associated with higher *ARAP1* transcript levels (rs11603334, $p = 0.0002$; rs1552224, $p = 0.015$). Due to lower LD between the lead SNPs and a third *ARAP1* transcribed SNP located in a downstream exon, rs2291289 ($r^2 = 0.12$, D' = 1.0, rs11603334, EUR), only three samples were heterozygous for both rs2291289 and rs11603334 (Table 2). The gDNA-to-cDNA comparison showed no AEI, while a comparison among cDNA samples showed the non-risk allele associated with modestly increased transcript levels (Figure 10A). We also attempted to evaluate AEI for nearby genes containing transcribed SNPs in lower LD with rs11603334, using previously described statistical methods.[60] We did not observe significant evidence of AEI for *STARD10, PDE2A*, or *FCHSD2* (Figure 10B-D and data not shown); zero to four samples were heterozygous for each transcribed

SNP and rs11603334 (Table 2). Overall, these results suggest that one or more functional variants in high LD with rs11603334 and rs1552224 may increase the transcriptional activity or message stability of *ARAP1* isoforms containing these SNPs.

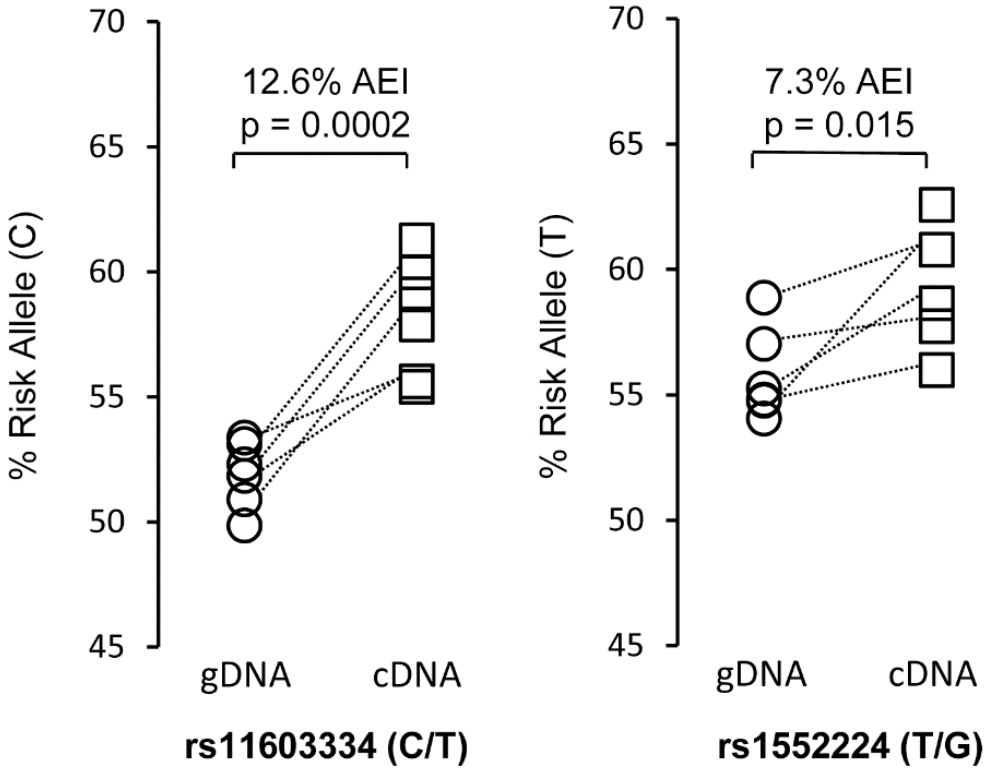

Figure 2. The T2D-Risk Alleles of rs11603334 and rs1552224 Are Associated with Increased *ARAP1* Expression

Allelic expression analysis in human pancreatic islet gDNA and cDNA from six donors each, including five matched pairs (matched pairs only: rs11603334, p = 0.0008; rs1552224, p = 0.06). The proportion of total gDNA and cDNA containing the T2D-risk alleles of rs11603334 (left) and rs1552224 (right) from heterozygous samples are quantified on the y axis. The cDNA containing the T2D-risk allele was expressed 12.6% and 7.3% more than the cDNA containing the nonrisk allele in the rs11603334 and rs1552224 assays, respectively. The p-values were calculated with a two-sided t test.

Table 2. Transcribed SNPs Used to Evaluate Allelic Expression Imbalance for *ARAP1, STARD10, PDE2A*, and *FCHSD2*

Gene	SNP	Annotation	Alleles Major/ Minor	MAF	LD with rs11603334		# Informative Heterozygotes[b]
					D'	r^2	
ARAP1	rs11603334	5'UTR	C/T	.16	1	1	6
ARAP1	rs1552224	5'UTR	T/G	.16	1	1	6
ARAP1	rs2291289	Synonymous	A/G	.41	.94	.12	3
STARD10	rs2291290	3'UTR	A/G	.11	1	.025	2
STARD10	rs519790	5'UTR	C/G	.28	.97	.071	2
PDE2A	rs392565	Synonymous	C/T	.30	.011	.0053	4
FCHSD2	rs76762469	3'UTR	G/T	.10	.18[a]	.001[a]	0

MAF, minor allele frequency; LD, linkage disequilibrium; UTR, untranslated region. MAF, D', and r^2 data are estimated from EUR, 1000 Genomes Phase 1, except [a]CEU, 1000 Genomes Pilot. [b]Of 87 total human pancreatic islet samples.

Candidate Regulatory Variants at the *ARAP1* Promoters

The originally identified lead SNPs rs11603334 and rs1552224 at the P1 promoter overlap ENCODE transcription factor binding sites within a region of open and accessible chromatin in human pancreatic islets that is marked by DNase hypersensitivity and active histone H3 lysine modifications H3K4me3 and H3K9ac (Figure 3 and Table 7). The lead fine-mapped SNP in this study, rs7109575 ($r^2 = 0.94$ with rs11603334, EUR), is located within an islet DNase hypersensitivity region and clustered transcription factor binding at the P2 promoter (Figure 11 and Table 7). Among the 17 additional variants in high LD with rs11603334 ($r^2 \geq 0.8$, EUR; Table 7), rs77464186 ($r^2 = 0.95$ with rs11603334, EUR) is the next most promising functional candidate due to its location within a region of islet DNase hypersensitivity and clustered transcription factor binding in the first intron of *ARAP1* isoforms transcribed from the P2 promoter (Figure 11 and Table 7).

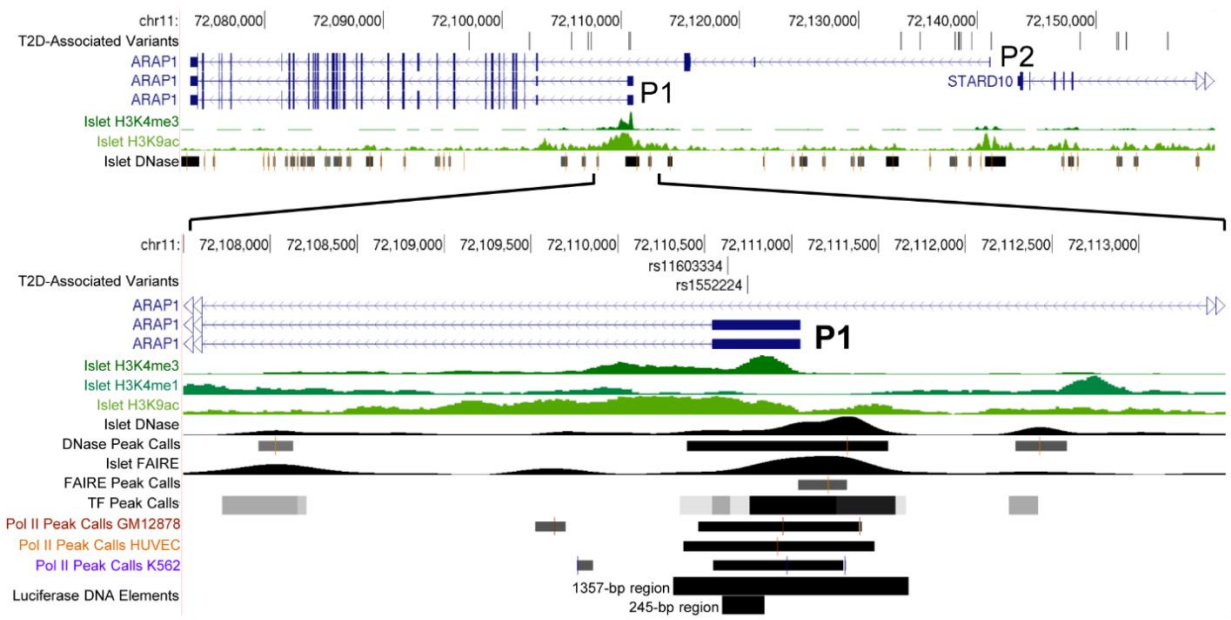

Figure 3. SNPs rs11603334 and rs1552224 Are Located in a Region with Evidence of Regulatory Potential at the *ARAP1* P1 Promoter at 72.11 Mb
University of California Santa Cruz (UCSC) Genome Browser (hg18) diagram showing rs11603334 and rs1552224 situated at the *ARAP1* P1 promoter and 18 additional variants in high LD ($r^2 \geq 0.8$). *ARAP1* is transcribed from right to left. SNPs rs11603334 and rs1552224 overlap regions of accessible chromatin in human pancreatic islets detected by H3K4me3 and H3K9ac ChIP-seq, are situated within or proximal to regions of open chromatin in human pancreatic islets detected by DNase hypersensitivity and FAIRE, and overlap binding sites for Pol2 and other transcription factors detected by ChIP-seq in several ENCODE cell lines (see also Table 7). The DNA sequences cloned for evaluating differential allelic transcriptional activity in the dual-luciferase reporter assays are indicated.

The rs11603334 T2D-Risk and Proinsulin-Decreasing Allele Increases Transcriptional Reporter Activity in Rodent Pancreatic Beta Cell Lines

To interrogate allelic differences in transcriptional activity of rs11603334, rs1552224, rs7109575, and rs77464186, we cloned DNA sequences containing the T2D-risk (proinsulin-decreasing) or non-risk (proinsulin-increasing) alleles into a promoterless luciferase vector and measured luciferase activity in the MIN6 mouse and 832/13 rat beta cell lines. Due to the perfect LD and short distance between rs1552224 and rs11603334, we first tested them together as a haplotype in the context of the *ARAP1* P1 promoter. A 1,357-bp region of DNA, defined by DNase

hypersensitivity in human pancreatic islets and Pol2 binding in several other cell types (Figure 3), exhibited strong transcriptional activity compared to empty vector control in both MIN6 (Figure 4A) and 832/13 (Figure 4B). The rs1552224 and rs11603334 T2D-risk haplotype (TC) demonstrated a two-fold increase in transcriptional activity compared to the non-risk haplotype (GT) in both MIN6 ($p < 0.0001$) and 832/13 ($p = 0.015$), consistent with our finding that the T2D-risk alleles are associated with increased *ARAP1* expression. The 1,357-bp region also contained two short sequences of variable CT repeat length ($CT_{6\text{-}10}$); however, the numbers of repeats at the two sites did not influence transcriptional activity (Figure 12A-B). A narrower, 245-bp DNA region (Figure 3) containing only variants rs1552224 and rs11603334 also exhibited strong transcriptional activity compared to empty vector control (Figure 4C). As observed with the 1,357-bp region, the two-SNP T2D-risk haplotype (TC) demonstrated a two-fold increase in transcriptional activity compared to the non-risk haplotype (GT) ($p = 0.012$). To determine which of the SNPs is responsible for the haplotype effects on transcriptional activity at the P1 promoter, we used site-directed mutagenesis to synthetically create the alternate 'missing' haplotypes. The T2D-risk C allele of rs11603334 replicated the previously observed two-fold increase in transcriptional activity compared to the non-risk T allele ($p < 0.0001$), while the T2D-risk and non-risk alleles of rs1552224 (T/G) showed no significant difference (Figure 4D). We did not observe evidence of allelic differences in transcriptional activity for either rs7109575 (Figure 4E) or rs77464186 (Figure 12C) near the P2 promoter. However, the 287-bp region of DNA surrounding rs7109575 increased transcriptional activity nine-fold compared to the empty control vector ($p < 0.0001$), consistent with promoter activity.

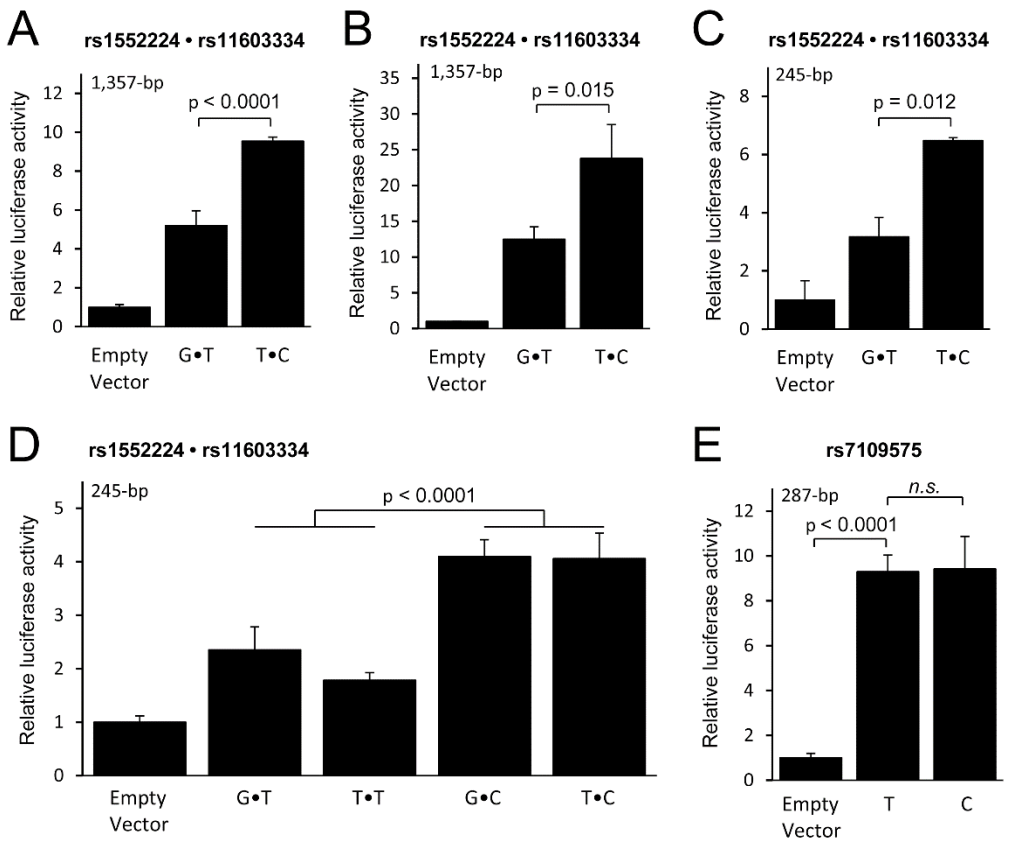

Figure 4. The T2D-Risk Allele, C, of rs11603334 Increases Transcriptional Activity of the *ARAP1* P1 Promoter

Transcriptional activity was evaluated with dual-luciferase reporter assays in the MIN6 mouse and 832/13 rat insulinoma β-cell lines 48 hr after transfection with recombinant vector containing selected regions of the human *ARAP1* promoters cloned upstream of the luciferase gene. (A) rs1552224 and rs11603334 evaluated as a haplotype in a 1,357 bp region at the *ARAP1* P1 promoter in MIN6. (B) rs1552224 and rs11603334 evaluated as a haplotype in the 1,357 bp region in 832/13. (C) rs1552224 and rs11603334 evaluated as a haplotype in a narrower 245 bp region at the *ARAP1* P1 promoter in 832/13. (D) Site-directed mutagenesis separated the effects of rs1552224 and rs11603334 in the 245 bp region in 832/13. (E) rs7109575 evaluated in a 287 bp region at the *ARAP1* P2 promoter in 832/13. The p-values were calculated with a two-sided t test (A–C and E) or two-way ANOVA (D). Data represent the mean of two to ten clones per allele ± SD. A significance threshold of $p < 0.05$ was used. NS, not significant.

Decreased *in vitro* Binding of PAX6 and PAX4 to the rs11603334 T2D-Risk and Proinsulin-Decreasing Allele (C)

We performed EMSAs using nuclear protein lysates extracted from MIN6 and 832/13 to evaluate differences in transcription factor binding to alleles of rs11603334 and rs1552224. Multiple protein-DNA complexes were observed in the EMSAs, consistent with expectations for DNA within a promoter region (Figure 5 and Figure 13). At least one protein complex was observed to bind specifically to the non-risk (T) allele of rs11603334 in MIN6 (Figure 5A, arrow). DNA competition assays using excess unlabeled oligo containing the T allele reproducibly disrupted the protein-DNA complex more efficiently than excess unlabeled oligo containing the C allele, further supporting allelic differences in protein binding (Figure 5A, lanes 8 vs. 9; Figure 13A, lanes 7 vs. 8). Preferential protein binding to the T allele was also detected in 832/13 for a complex with similar mobility (Figure 13A, arrow). To identify the differentially bound protein(s), we performed EMSAs with supershift using antibodies targeting transcription factors selected from those with computationally predicted or empirically determined (ChIP-seq) binding at rs11603334 (Materials and Methods). Incubation with antibody targeting Pax6 partially disrupted the T allele-specific protein-DNA complex (Figure 5A, lane 10; Figure 5B, lane 8) and incubation with antibodies targeting both Pax6 and Pax4 fully disrupted the complex (Figure 5B, lane 10). Among the other transcription factors tested with supershift assay, only AP-2β, Hif1α, and Hif1β were observed to bind to the 21-bp DNA sequence containing rs11603334; these proteins bound to both the C and T alleles with approximately equal affinity (Figure 13B-C, arrows). We did not observe clear evidence of allele-specific protein binding to rs1552224 (Figure 13D).

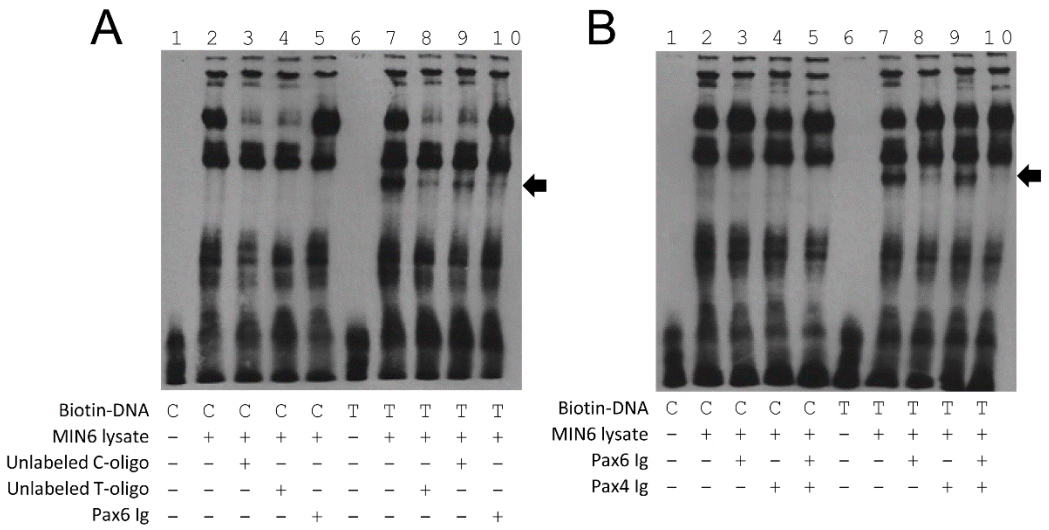

Figure 5. The T2D-Risk Allele, C, of rs11603334 Disrupts Binding of Transcriptional Regulators PAX6 and PAX4

Differential allelic protein-DNA binding was evaluated in vitro with EMSAs. Biotin-labeled 21 bp double-stranded DNA oligonucleotide centered on the reference (C) and alternate (T) alleles of rs11603334 was incubated with MIN6 nuclear lysate, and resulting protein-DNA complexes were separated by electrophoresis (A and B: lanes 2 and 7, respectively). For the competition assays, EMSA reactions were incubated with 60-fold excess unlabeled DNA oligonucleotide containing either the T or C allele (A: lanes 8 and 9, respectively). For the supershift assays, EMSA reactions were incubated with antibody targeting PAX6 (A: lane 10; B: lane 8), PAX4 (B: lane 9), or both PAX6 and PAX4 (B: lane 10).

Discussion

For most T2D-associated loci identified through GWAS, the functional variants and affected genes are not known. In this study, we identify a plausible functional regulatory variant, rs11603334, and demonstrate its allelic effects on expression of *ARAP1*, a gene at the 11q13.4 locus. We provide evidence suggesting that the C allele of rs11603334 disrupts binding of transcriptional regulators and increases transcriptional activity at the *ARAP1* P1 promoter, leading to higher *ARAP1* expression in pancreatic islets. Increased *ARAP1* expression may be one molecular genetic contributor to T2D susceptibility.

We conducted association analyses using the densest set of variants to date, reporting rs7109575 as the lead fine-mapped SNP, and confirming the primary robust association with fasting proinsulin for previously reported lead SNPs rs11603334 and rs1552224. Although not the most strongly associated SNP in this study, the association p-value for our proposed functional SNP rs11603334 ($p = 2.2 \times 10^{-48}$) was similar to that of rs7109575 ($p = 7.4 \times 10^{-50}$); sampling variability, missing genotypes, or imputation uncertainty may influence the relative p-values. Conditioning on either rs11603334 or rs7109575 attenuated the proinsulin association at this locus, consistent with a single signal at this locus and supporting the use of the r^2 measure of LD to identify candidate functional SNPs. Rare variants may also contribute to variation in proinsulin levels among carriers. The variant that became more significant after adjusting for rs11603334 (rs202137453, *STARD10* Pro196Ser) may be a true signal or a false positive among 3,818 variants tested. Further fine-mapping and conditional analyses using a larger T2D case-control population will be needed to determine if other variants in the region also contribute to T2D risk.

We demonstrate for the first time that the T2D-risk alleles of rs11603334 and rs1552224 are associated with increased *ARAP1* mRNA expression in primary human pancreatic islets. This finding was reproducibly evident with a small number of samples because AEI compares quantification of allele-specific mRNA levels within heterozygous individuals, thereby minimizing confounding effects such as inter-individual environmental exposures that can mask genotypic effects in eQTL studies. We had the best power to detect association of rs11603334 and rs1552224 with *ARAP1* mRNA expression because the SNPs are located within an exon; this exon is specific to *ARAP1* isoforms transcribed from the P1 promoter. Previously, rs519790 was reported to be a strong eQTL for *STARD10* in blood ($p = 2.7 \times 10^{-24}$);[1] however, this SNP is in low LD with rs1552224

($r^2 = 0.071$, EUR), which shows only a modest, residual eQTL association with *STARD10* (blood, $p = 8.6 \times 10^{-7}$; liver, $p = 1.79 \times 10^{-5}$; omental adipose tissue, $p = 1.21 \times 10^{-6}$).[25] A nominal association between rs1552224 and *ARAP1* (*CENTD2*) expression has been previously reported in T-cells ($p = 0.019$) in a direction consistent with our results.[77,78]

In agreement with our finding that the T2D-risk alleles of rs11603334 and rs1552224 are associated with increased *ARAP1* expression, our experiments using the transcriptional reporter assay show that the T2D-risk allele of rs11603334 upregulates transcriptional activity at the *ARAP1* P1 promoter. While transcriptional activity at a gene promoter may be expected, allelic differences in promoter activity cannot be predicted easily; functional assays are necessary to demonstrate allelic effects. Among three tested variants situated at the *ARAP1* P1 and P2 promoters in this study, only rs11603334 demonstrated allelic effects on promoter activity. Taken together, the results from the AEI and transcriptional reporter assays implicate increased expression of *ARAP1* isoforms containing rs11603334 and transcribed from the P1 promoter (NM_015242, NM_001135190) in T2D susceptibility.

We show that the beta cell transcriptional regulator Pax6 preferentially binds to the mRNA-lowering (T) allele of rs11603334 in a protein-DNA complex that may also contain Pax4, suggesting a repressive role for one or both factors in transcription at the *ARAP1* P1 promoter (Figure 6). Pax6 and Pax4 have similar binding sequences, although they bind DNA with different relative affinities.[79–81] In the endocrine pancreas, Pax6 is often reported to function as a transcriptional activator, while Pax4 often functions as a repressor. Pax4 has been shown to inhibit the transcriptional activity of co-expressed Pax6 in the HIT-T15 hamster beta cell line and in 293T cells, possibly through competition for DNA binding sites.[80] Pax4 also represses basal transcriptional activity of the glucagon promoter in αTC1.9 and InR1G9 cells and

strongly represses Pax6-activated transcription of the both the glucagon[81] and insulin promoters.[82] However, repressive activity for human and mouse PAX6 has also been demonstrated for specific regulatory elements in other cell types.[83–85] The observed inhibition of transcriptional activity for *ARAP1* transcripts containing the rs11603334-T allele may be mediated by PAX4, PAX6, or both. Human coding variants in PAX4 or PAX6 have been previously implicated in impaired glucose tolerance,[86,87] early-onset diabetes,[86,88] ketosis-prone diabetes,[89] and proinsulin processing.[90] Recently, genome-wide association signals for T2D have also been identified near *PAX4* in East Asian[45] and Chinese populations.[91]

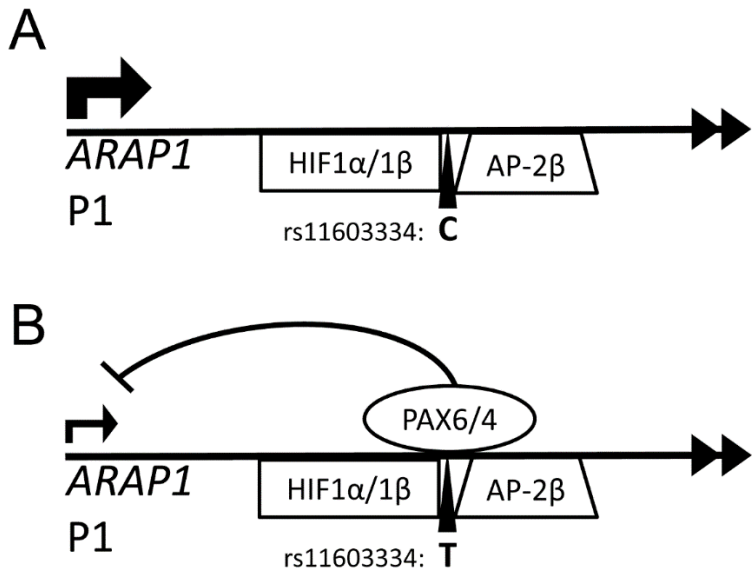

Figure 6. Proposed Model for PAX6 and PAX4 Repression of Transcriptional Activity at *ARAP1*
PAX6 and PAX4 at rs11603334 in a complex with stronger binding to the *ARAP1* mRNA-decreasing T allele (B, T2D nonrisk) than to the mRNA-increasing C allele (A, T2D risk) suggests a role for one or both of these factors in repressing transcription at the *ARAP1* P1 promoter. Transcription factors HIF1α, HIF1β, and AP-2β bind both alleles.

Our results do not preclude the existence of additional functional genes, variants, and/or mechanisms at this locus. Power to detect association of the lead SNPs with expression levels of *STARD10*, *PDE2A*, and *FCHSD2* was limited due to small numbers of SNPs and samples informative for AEI. Although 87 total islet samples were available, only six were heterozygous for rs11603334 and zero to four samples were informative for the *STARD10*, *PDE2A*, and *FCHSD2* assays. We cannot exclude possible associations between rs11603334 and expression levels of these genes. An assay to test AEI for all isoforms of *ARAP1,* using rs2291289, similarly was of limited utility due to the small number of heterozygous samples (n = 3) for this SNP and due to wide assay variance. Moreover, *ATG16L2* could not be tested using this approach because it does not contain any common transcribed SNPs. Future analysis of a larger pancreatic islet expression dataset will be required to validate our findings for *ARAP1* and assess allelic differences in gene expression for all other genes.

We chose to investigate the mechanisms underlying the set of variants most strongly associated with proinsulin and T2D following dense fine-mapping analyses. We prioritized among highly significant non-coding variants based on their locations within regions of open chromatin in pancreatic islets and successfully identified a variant with allelic differences in transcriptional activity. Variants located outside of these open chromatin regions may also be found to have regulatory function. In addition, functional non-coding variants at this locus may regulate gene expression through allelic effects on splicing, message stability, or translational control. Based on sequence, rs1552224 may create a splice site, and rs11603334-T creates a methionine codon upstream of the RefSeq start codon. If used, this codon could potentially lead to attempted translation of a 90-nt upstream open reading frame, which could interfere with efficient translation of *ARAP1*.[92]

Specific roles for ARAP1 in the process of proinsulin conversion and/or insulin secretion in the beta cell have not yet been reported. ARAP1 is located near both the Golgi and the plasma membrane in other cell types[18,34–36] and has been demonstrated to regulate several members of the Arf and Rho small GTPase protein families in vitro and in cell culture.[18,32] Of the reported substrates, both ARF6 and CDC42 have been shown to play a role in glucose-stimulated insulin secretion.[40–42] Future work will address the effects of upregulated *ARAP1* expression on proinsulin processing and insulin secretion in the pancreatic beta cell. Given that our data suggest that increased *ARAP1* expression is associated with increased risk of T2D, ARAP1 antagonists might hold important therapeutic potential.

Identifying functional variants at T2D GWAS loci and elucidating mechanisms to explain how variants influence expression or function of target genes will help facilitate the discovery of biological gene candidates and novel pathways involved in disease pathology. This work contributes to a growing understanding of the molecular genetic architecture of T2D, provides a model of fine-mapping and experimental analyses that can be used for uncovering functional variants contributing to associations at GWAS loci, and paves the way for future research on the role of *ARAP1* as a diabetes candidate gene in pancreatic islets.

Supplemental Figures and Tables

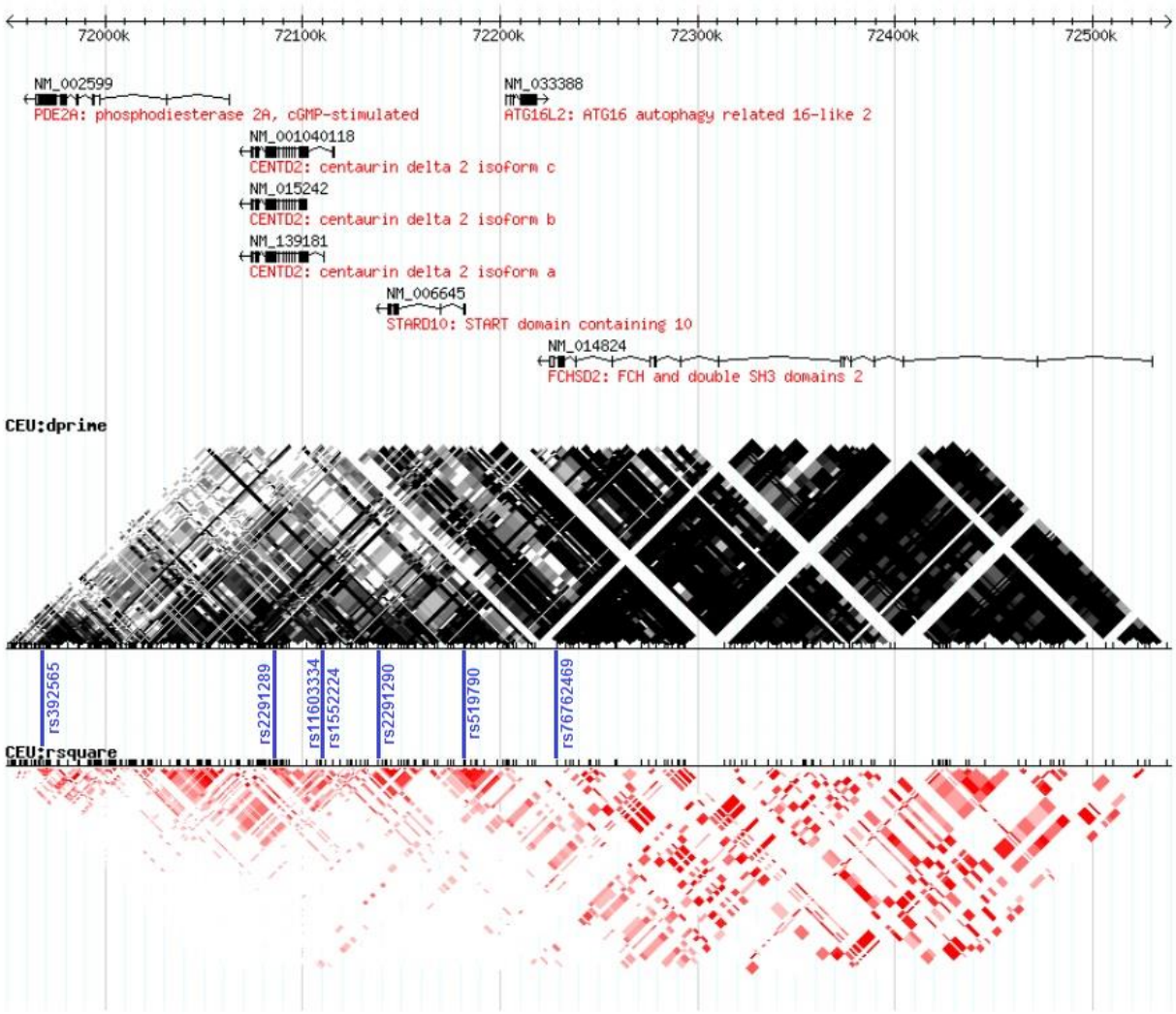

Figure 7. Linkage Disequilibrium near *PDE2A*, *ARAP1/CENTD2*, *STARD10*, *ATG16L2*, and *FCHSD2*

HapMap genome browser (hg18/NCBI 36) LD plot diagrams depicting D' (black) and r^2 (red) for CEU (data release 27 Phase II & III). Six transcribed variants selected for AEI analysis are labeled near their relative genomic positions (blue). In this figure, *ARAP1* is labeled using its previous symbol, *CENTD2*.

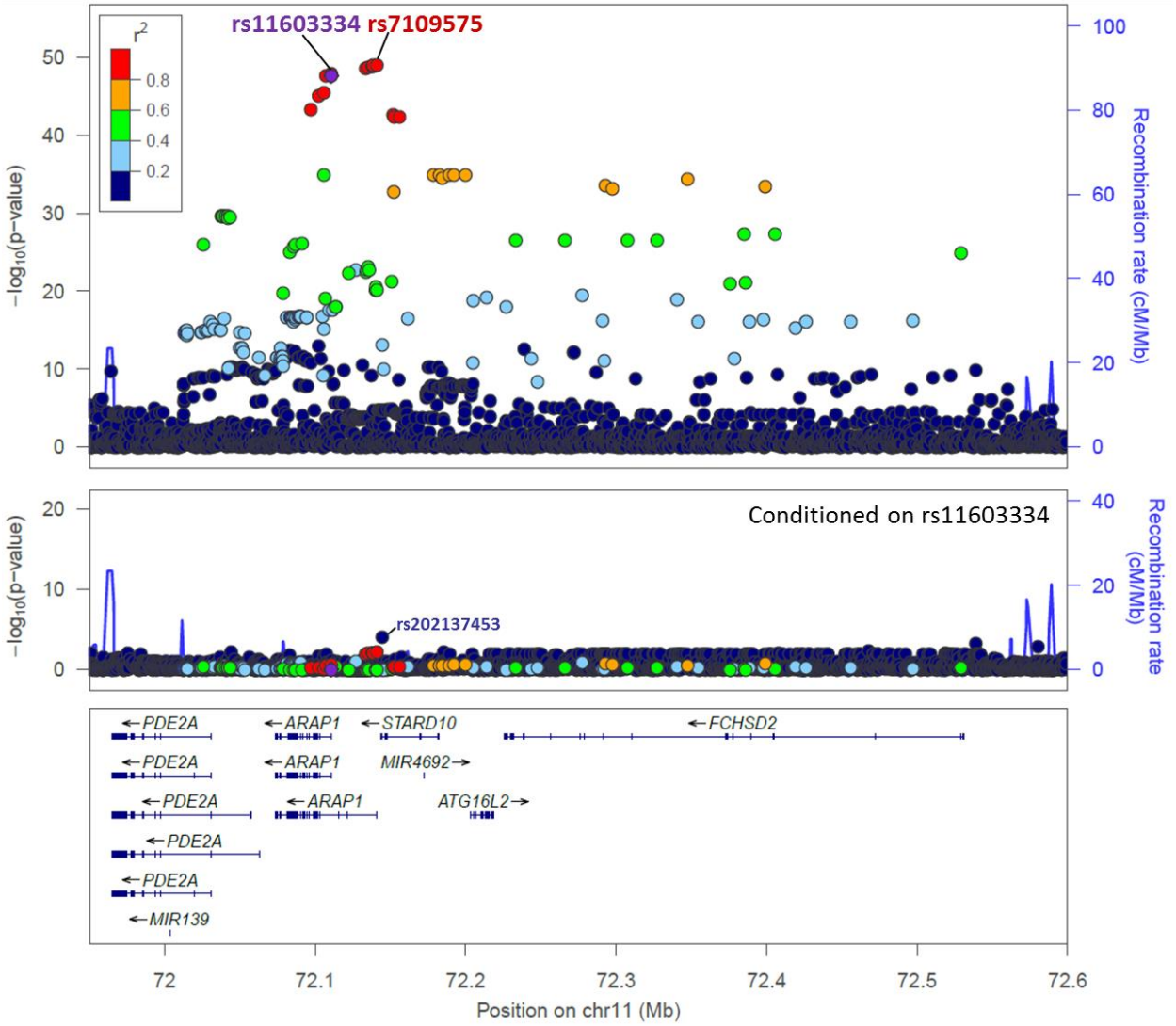

Figure 8. All Variants Associated with Fasting Proinsulin in the METSIM Study
LocusZoom plots of fasting proinsulin association adjusted for age, BMI and fasting insulin. The ~600 kb region containing *PDE2A*, *ARAP1*, *STARD10*, *ATG16L2*, and *FCHSD2* is shown. The robust association with proinsulin levels (top panel: rs7109575, $p = 7.4 \times 10^{-50}$) is attenuated upon inclusion of rs11603334 as a covariate (bottom panel, only showing y-axis to 10^{-20}). Each SNP is colored according to its LD (r^2) with rs11603334. Genomic coordinates refer to hg18.

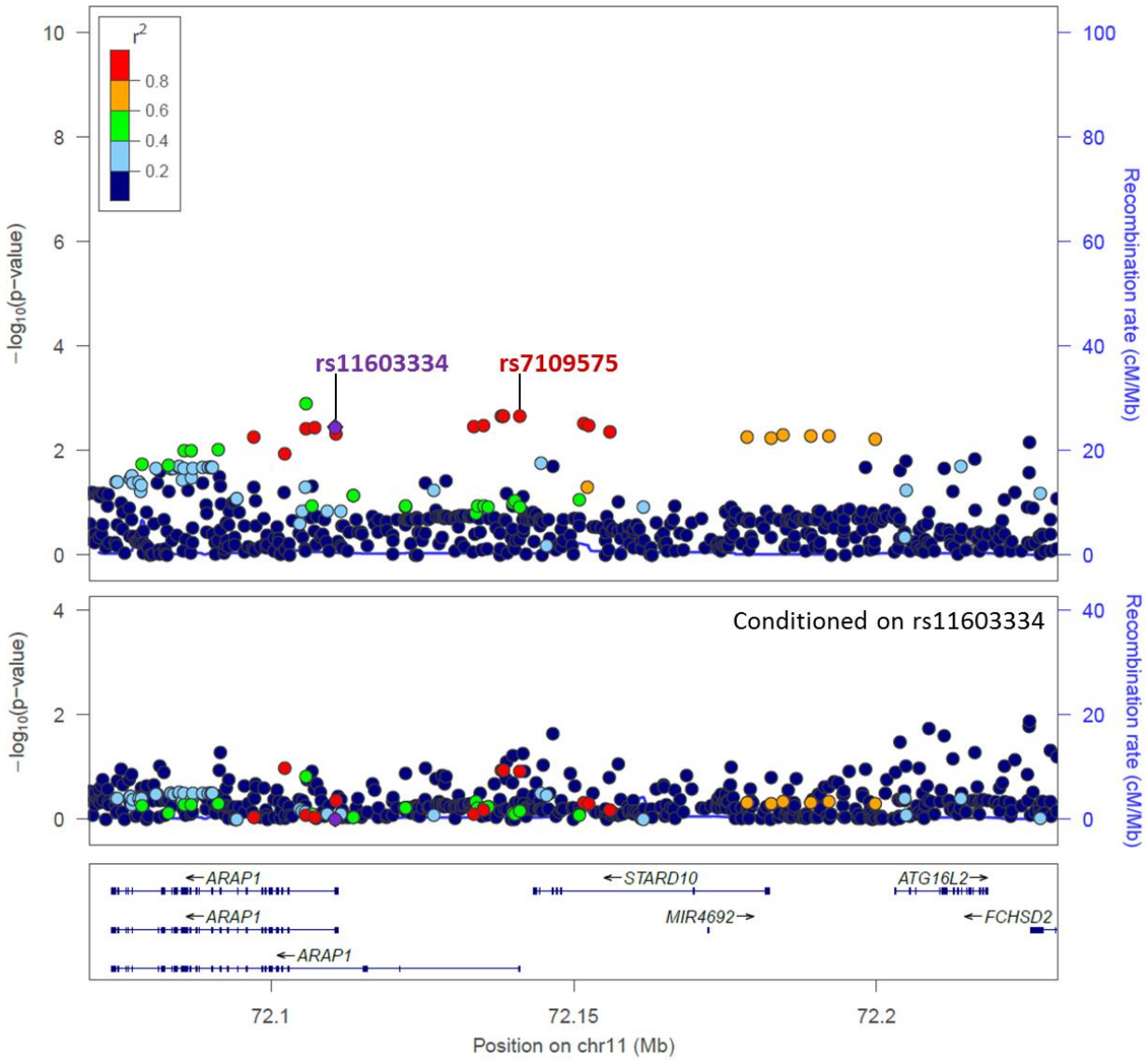

Figure 9. Variants Associated with T2D in the METSIM Study
LocusZoom plots of T2D association in 1,389 case and 5,758 control subjects. Conditional analysis on rs11603334 attenuated association of rs7109575 and LD proxies with T2D (top vs. bottom panels; bottom panel only showing y-axis to 10^{-4}). Each SNP is colored according to its LD (r^2) with rs11603334. Genomic coordinates refer to hg18.

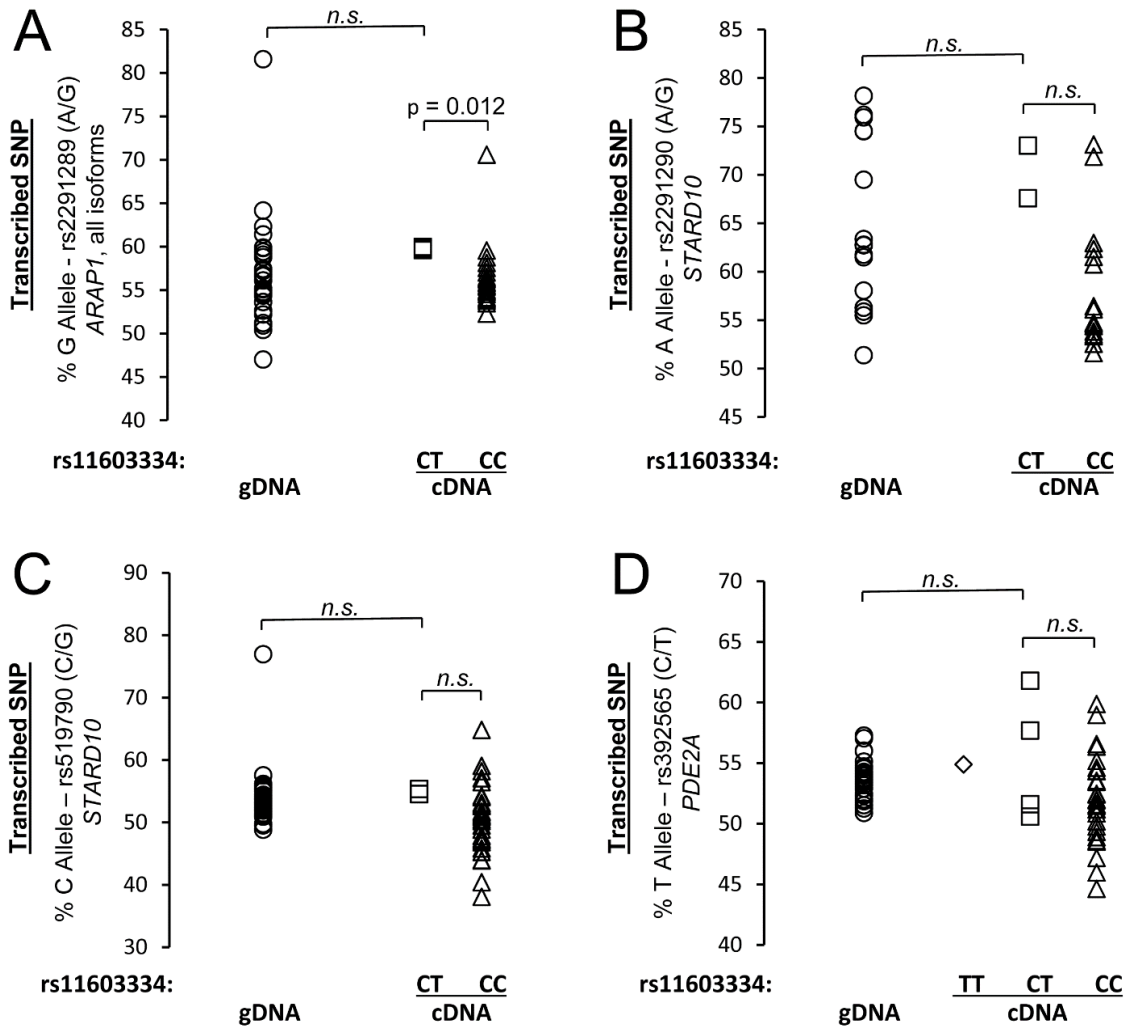

Figure 10. Allelic Expression Analyses for *ARAP1, STARD10* and *PDE2A* in Human Pancreatic Islet gDNA and cDNA

The proportion of total gDNA and cDNA containing the G allele of transcribed *ARAP1* SNP rs2291289 (A), the A allele of transcribed *STARD10* SNP rs2291290 (B), the C allele of transcribed *STARD10* SNP rs519790 (C), and the T allele of transcribed *PDE2A* SNP rs392565 (D) are quantified on the y-axes. All data points are heterozygous for the transcribed SNP. The cDNA data are sub-divided by genotype of lead SNP rs11603334. The p-values for panels A-D were calculated using two-sided t tests when the data were normally distributed. F tests were performed to determine if variances were equal or unequal prior to performing each t test. Non-parametric Wilcoxon pair-wise tests were used in cases where the data were not normally distributed (A: comparisons of both gDNA and rs11603334 homozygous CC cDNA samples to heterozygous CT cDNA samples; B: comparison of homozygous cDNA samples to heterozygous cDNA samples; C: comparison of gDNA samples to heterozygous cDNA samples). For *PDE2A* (D), a one-sided F test was additionally performed to compare the sample variance between cDNA heterozygous for rs11603334 to cDNA homozygous for rs11603334; result was non-significant. A significance level of $p < 0.05$ was used.

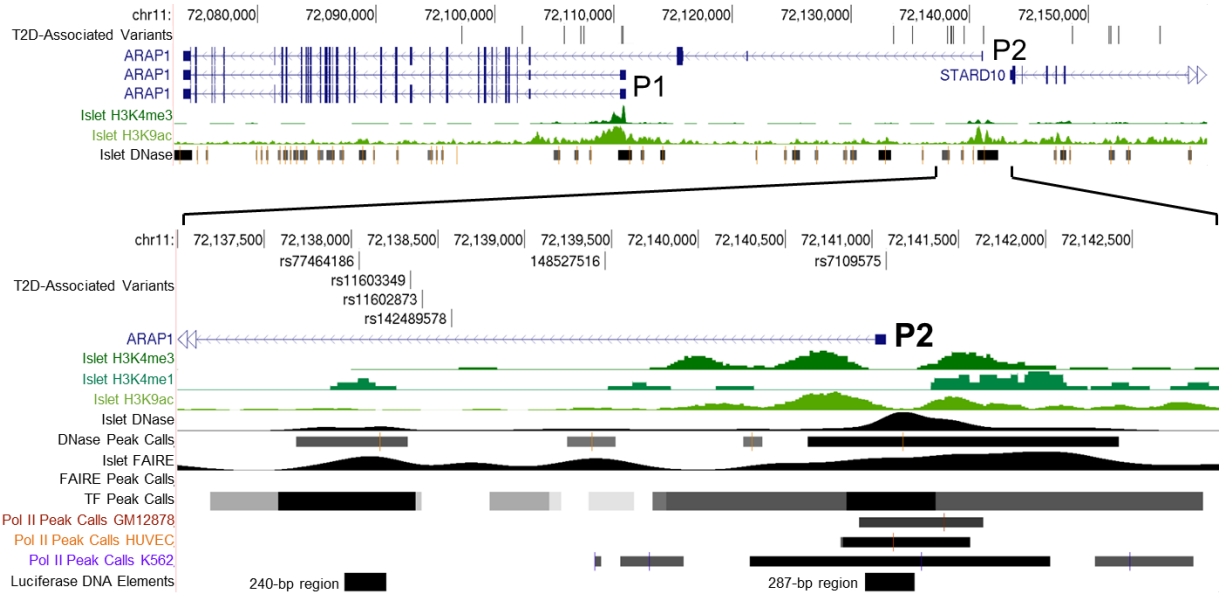

Figure 11. Variants with Regulatory Potential near the *ARAP1* P2 Promoter
UCSC genome browser (hg18) diagram showing six of twenty total variants in high LD ($r^2 \geq 0.8$) with lead SNPs rs11603334 and rs1552224. SNPs rs7109597 and rs77464186 overlap regions of open chromatin in human pancreatic islets detected by DNase hypersensitivity and other transcription factors detected by ChIP-seq across many ENCODE cell lines (Table 6). DNA sequences cloned to evaluate differential allelic transcriptional activity in the dual-luciferase reporter assays are also shown.

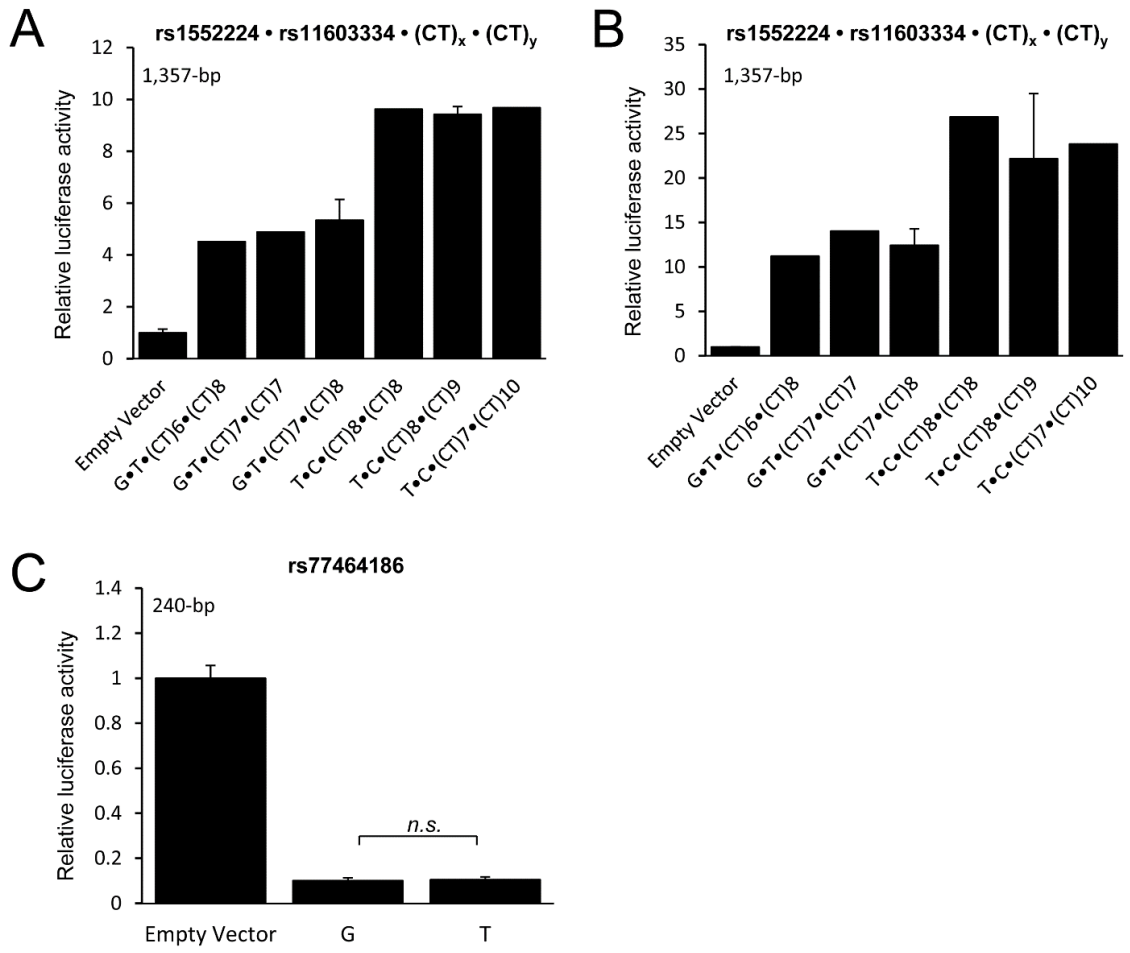

Figure 12. *ARAP1* Variants that Did Not Show Allelic Differences in Transcriptional Activity

Transcriptional activity was evaluated using dual-luciferase reporter assays in the MIN6 and 832/13 β-cell lines 48 hours after transfection with recombinant vector containing selected DNA regions cloned upstream of the luciferase gene. A-B: Variable short CT repeats present within the 1357-bp region at the downstream *ARAP1* promoter do not appear to influence transcriptional activity when tested in MIN6 (A) or 832/13 (B). C: The 240-bp region containing rs77464186 generated less transcription of luciferase compared to empty vector in 832/13; however, the effect was not different between SNP alleles. The p-values were calculated using a two-sided t test. Data are represented as the mean of two to nine clones per allele ± standard deviation. A significance level of $p < 0.05$ was used.

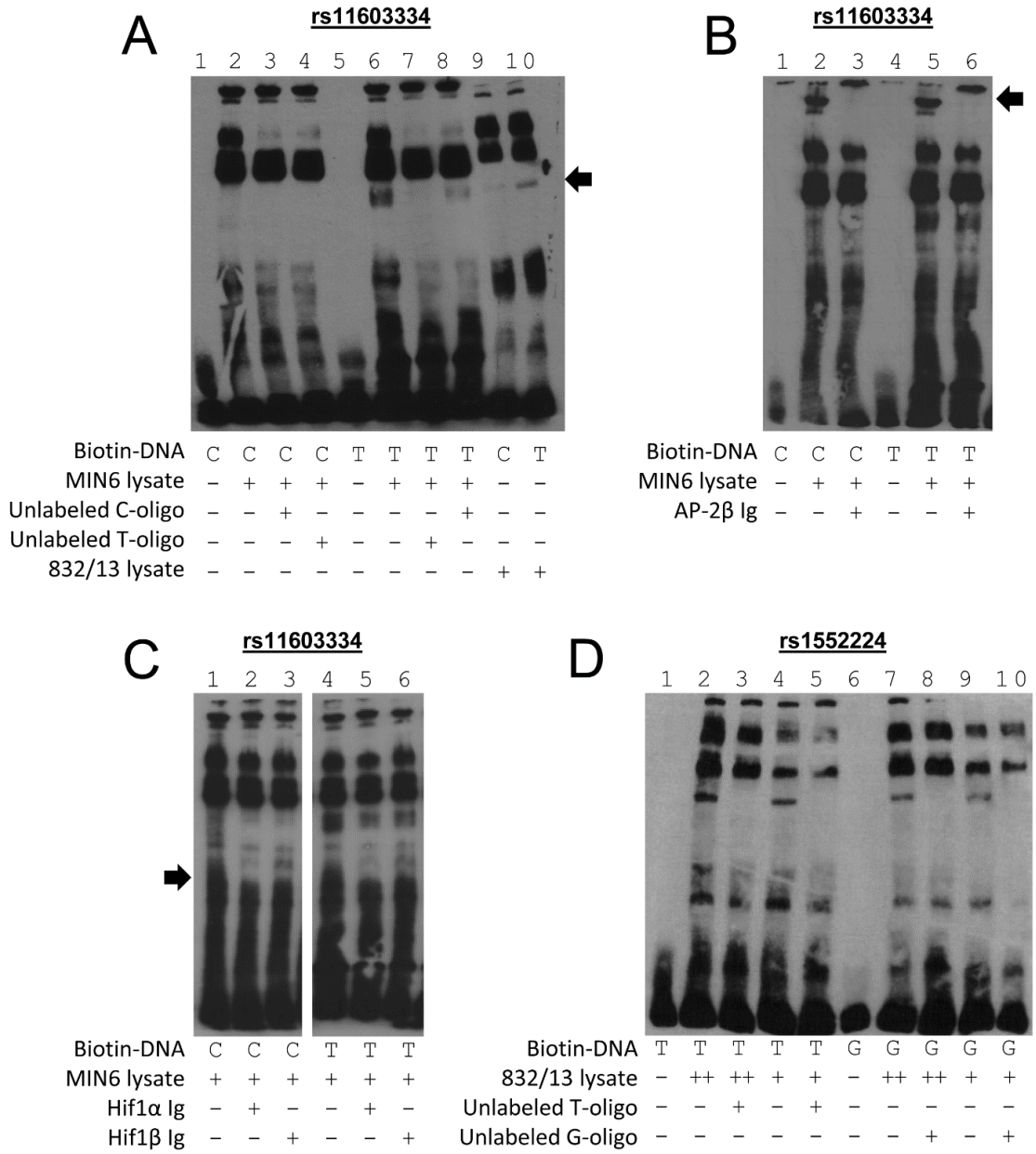

Figure 13. EMSAs Identify Additional Proteins Bound to rs11603334
Differential allelic protein-DNA binding was evaluated *in vitro* using EMSA. Biotin-labeled 21-bp double-stranded DNA oligonucleotide centered on the reference or alternate alleles of rs11603334 (C/T) (A-C) or rs1552224 (T/G) (D) was incubated with MIN6 nuclear lysate (A: lanes 2 and 6; B: lanes 2 and 5; C: lanes 1 and 4) or 832/13 nuclear lysate (A: lanes 9 and 10; D: lanes 2, 4, 7, and 9) and the resulting protein-DNA complexes were separated by electrophoresis. For the competition assays (A and D), EMSA reactions were incubated with 60-fold (A) or 50-fold (D) excess unlabeled DNA oligonucleotide containing rs11603334-T (A: lane 7), rs11603334-C (A: lane 8), rs1552224-T (D: lanes 3 and 5) or rs1552224-G (D: lanes 8 and 10). For the supershift assays (B-C), EMSA reactions were incubated with antibody targeting AP-2β (B: lanes 3 and 6), Hif1α (C: lanes 2 and 5), or Hif1β (C: lanes 3 and 6).

Table 3. Primers Used to Evaluate Allelic Expression Imbalance (AEI)

Gene	SNP	Forward	Reverse	Extension	Amplicon Length
ARAP1	rs11603334	TGCGTATGCCCAAGAATGTC	TGTTTGCGAGACCAGCCCAG	ATGTCGTGTACCTGCCC	71
ARAP1	rs1552224	GTCTGAGGCTGCGTTCTCTT	TCCACAGTTATACACAGAAACAAACA	GCCCAGGTATGGCTTTGC	129
ARAP1	rs2291289	CCGCTCGCCCTGTATGTA	CGCTGCAACTACGGAAACTG	CACCTCCTTCGCTCCAC	107 (cDNA) 479 (gDNA)
STARD10	rs2291290	GATCGTTTATTGGGGCTCTG	CTTCAGGGACGGAGACAGG	GGGCGGCGGCCGCTCCTGC	146
STARD10	rs519790	GAAGAATTGGGAACAGGGACT	GTCAGGTGATTGTAGATTGAGGA	ATCCAACAAAGACACTGACC	247
PDE2A	rs392565	GATCAAAGATGTTGCAGCCG	AGGCACCACTTTGCTCA	GTGGGTGTTGAGGATGGC	61
FCHSD2	rs76762469	TGTGGAAGAGATGTGCGTTAAT	CACAAGGATGTGGCAAAGAT	AGATCATGTTGGCTACATAA	227

Only gene-specific sequences of the primers are shown. All assays could be designed within one exon except rs2291289; this amplicon spans an intron in the gDNA.

Table 4. Primers Used to Generate Clones for Transcriptional Reporter Assays

PCR amplification primers

Variants Contained	Insert Length	Forward	Reverse
rs1552224, rs11603334	245	GTCCTTTCCTCCTGGACAGA	CTGCGTATGCCCAAGAATGT
rs1552224, rs11603334, $(CT)_x$, $(CT)_y$	1357	GCCCAGAGATGCTGAGAAAC	GGAAGAGGGAGGGATGAGAC
rs7109575	287	ACTCCAGGACCGTGACTAGG	GAAACTGCAGGCCAATGTG
rs77464186	240	ACGACCATCACCCTCCAATA	TCTCTACCACCCTCCCAATG

Site-directed mutagenesis primers

SNP	Forward	Reverse
rs11603334	CGGGACCCAGGCTTGCA**T**GGGCAGGTA	TACCTGCCC**A**TGCAAGCCTGGGTCCCG
rs1552224	AAGACATAAATCACAGG**G**GCAAAGCCATACCTGGG	CCAGGTATGGCTTTGC**C**CCTGTGATTTATGTCTT

Only gene-specific sequences are shown. Bold, underlined nucleotides indicate SNP position.

Table 5. Sequences Used to Generate Double-Stranded Oligos for Electrophoretic Mobility Shift Assay (EMSA)

SNP	Allele	5'Biotin labeling	Forward	Reverse
rs11603334	C	Yes	5'Biotin-CAGGCTTGCA**C**GGGCAGGTAC	5'Biotin-GTACCTGCCC**G**TGCAAGCCTG
	T	Yes	5'Biotin-CAGGCTTGCA**T**GGGCAGGTAC	5'Biotin-GTACCTGCCC**A**TGCAAGCCTG
	C	No	CAGGCTTGCA**C**GGGCAGGTAC	GTACCTGCCC**G**TGCAAGCCTG
	T	No	CAGGCTTGCA**T**GGGCAGGTAC	GTACCTGCCC**A**TGCAAGCCTG
rs1552224	T	Yes	5'Biotin-AAATCACAGG**T**GCAAAGCCAT	5'Biotin-ATGGCTTTGC**A**CCTGTGATTT
	G	Yes	5'Biotin-AAATCACAGG**G**GCAAAGCCAT	5'Biotin-ATGGCTTTGC**C**CCTGTGATTT
	T	No	AAATCACAGG**T**GCAAAGCCAT	ATGGCTTTGC**A**CCTGTGATTT
	G	No	AAATCACAGG**G**GCAAAGCCAT	ATGGCTTTGC**C**CCTGTGATTT

Bold, underlined nucleotides indicate SNP position.

Table 6. SNP Association with Fasting Proinsulin and Type 2 Diabetes in the METSIM Study

SNP	Chr11 Position (hg18)	Annotation	EA	NEA	MAF	N	Fasting Proinsulin Association				Type 2 Diabetes Association			
							Unconditioned		Conditioned on rs11603334		Unconditioned		Conditioned on rs11603334	
							β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p
Lead SNPs (top 50 most significant) associated with fasting proinsulin														
rs7109575	**72,141,083**	**Near TSS**	**G**	**A**	**0.250**	**8635**	**-0.270 (0.18)**	**7.43E-50**	**-0.395 (0.14)**	**5.91E-03**	**0.024 (0.01)**	**2.15E-03**	**0.104 (0.07)**	**1.21E-01**
rs11602873	72,138,410	Intron	A	T	0.250	8635	-0.270 (0.18)	9.95E-50	-0.387 (0.15)	8.79E-03	0.024 (0.01)	2.17E-03	0.109 (0.07)	1.16E-01
rs77464186	72,138,046	Intron	A	C	0.250	8635	-0.270 (0.18)	1.02E-49	-0.386 (0.15)	9.13E-03	0.024 (0.01)	2.17E-03	0.109 (0.07)	1.16E-01
rs11603349	72,138,342	Intron	T	C	0.250	8635	-0.270 (0.18)	1.02E-49	-0.386 (0.15)	9.15E-03	0.024 (0.01)	2.16E-03	0.110 (0.07)	1.14E-01
rs74333814	72,135,135	Intron	C	T	0.249	8635	-0.270 (0.18)	1.70E-49	-0.579 (0.22)	7.06E-03	0.023 (0.01)	3.34E-03	0.046 (0.10)	6.55E-01
rs7129793	72,133,533	Intron	C	T	0.249	8635	-0.269 (0.18)	2.30E-49	-0.582 (0.23)	1.11E-02	0.023 (0.01)	3.50E-03	0.028 (0.11)	7.99E-01
rs1552224	**72,110,746**	**5'UTR**	**A**	**C**	**0.253**	**8588**	**-0.268 (0.18)**	**1.06E-48**	**-0.275 (0.23)**	**2.26E-01**	**0.022 (0.01)**	**4.64E-03**	**-0.071 (0.09)**	**4.33E-01**
rs75896506	72,107,227	Intron	G	A	0.249	8635	-0.267 (0.18)	2.11E-48	-3.540 (3.44)	3.04E-01	0.023 (0.01)	3.62E-03	0.162 (1.40)	9.08E-01
rs11603334	**72,110,633**	**5'UTR**	**G**	**A**	**0.252**	**8630**	**-0.267 (0.18)**	**2.23E-48**	-	-	**0.023 (0.01)**	**3.55E-03**	-	-
rs76550717	72,105,820	Intron	A	G	0.247	8635	-0.262 (0.18)	2.90E-46	0.031 (0.10)	7.46E-01	0.023 (0.01)	3.72E-03	0.009 (0.04)	8.29E-01
rs9667947	72,102,282	Intron	T	C	0.255	8635	-0.257 (0.18)	7.09E-46	0.044 (0.09)	6.22E-01	0.020 (0.01)	1.11E-02	-0.064 (0.04)	1.05E-01
rs11605691	72,097,162	Intron	C	T	0.254	8635	-0.252 (0.18)	4.01E-44	-0.030 (0.05)	5.65E-01	0.022 (0.01)	5.36E-03	0.003 (0.02)	8.91E-01
rs79430446	72,151,652	Intron	G	T	0.277	8635	-0.243 (0.17)	2.35E-43	-0.038 (0.05)	3.99E-01	0.022 (0.01)	3.00E-03	0.014 (0.02)	4.64E-01
rs613937	72,152,487	Intron	A	G	0.277	8635	-0.242 (0.17)	3.69E-43	-0.035 (0.05)	4.41E-01	0.022 (0.01)	3.25E-03	0.013 (0.02)	5.03E-01
rs77106637	72,155,968	Intron	G	A	0.277	8635	-0.242 (0.17)	3.85E-43	-0.037 (0.05)	4.12E-01	0.021 (0.01)	4.24E-03	0.009 (0.02)	6.43E-01
rs117819392	72,192,151	Intergenic	C	T	0.304	8635	-0.213 (0.17)	9.51E-36	-0.033 (0.03)	2.61E-01	0.020 (0.01)	5.15E-03	0.009 (0.01)	4.57E-01
rs145398037	72,199,790	Intergenic	G	C	0.304	8635	-0.213 (0.17)	9.61E-36	-0.034 (0.03)	2.42E-01	0.020 (0.01)	5.96E-03	0.008 (0.01)	5.04E-01
rs1059000	72,182,620	Intergenic	G	C	0.303	8635	-0.213 (0.17)	1.01E-35	-0.032 (0.03)	2.62E-01	0.020 (0.01)	5.65E-03	0.009 (0.01)	4.89E-01
rs73541184	72,105,796	Intron	G	A	0.315	8635	-0.212 (0.17)	1.05E-35	-0.014 (0.03)	6.47E-01	0.023 (0.01)	1.25E-03	0.019 (0.01)	1.49E-01
rs76724404	72,189,169	Intergenic	A	C	0.304	8635	-0.213 (0.17)	1.18E-35	-0.032 (0.03)	2.76E-01	0.020 (0.01)	5.30E-03	0.009 (0.01)	4.67E-01
rs3862791	72,178,650	Intron	C	T	0.303	8635	-0.213 (0.17)	1.24E-35	-0.031 (0.03)	2.85E-01	0.020 (0.01)	5.38E-03	0.009 (0.01)	4.74E-01
rs139534838	72,184,528	Intergenic	T	C	0.304	8635	-0.212 (0.17)	3.13E-35	-0.029 (0.03)	3.05E-01	0.020 (0.01)	4.99E-03	0.010 (0.01)	4.43E-01

Table 6. SNP Association with Fasting Proinsulin and Type 2 Diabetes in the METSIM Study

SNP	Chr11 Position (hg18)	Annotation	EA	NEA	MAF	Fasting Proinsulin Association					Type 2 Diabetes Association		
						N	Unconditioned		Conditioned on rs11603334		Unconditioned		Conditioned on rs11603334	
							β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p
rs11605166	72,347,425	Intron	T	C	0.306	8632	-0.212 (0.17)	4.15E-35	-0.030 (0.03)	2.93E-01	0.019 (0.01)	7.51E-03	0.007 (0.01)	5.79E-01
rs11602858	72,292,717	Intron	C	G	0.294	8635	-0.210 (0.17)	2.70E-34	-0.038 (0.03)	1.64E-01	0.020 (0.01)	7.74E-03	0.007 (0.01)	5.23E-01
rs2886599	72,398,919	Intron	C	A	0.295	8635	-0.210 (0.17)	3.48E-34	-0.037 (0.03)	1.75E-01	0.020 (0.01)	7.58E-03	0.008 (0.01)	5.17E-01
rs185815472	72,297,508	Intron	G	A	0.295	8635	-0.209 (0.17)	6.76E-34	-0.035 (0.03)	1.99E-01	0.020 (0.01)	6.80E-03	0.008 (0.01)	4.79E-01
rs7103836	72,152,219	Intron	G	C	0.336	8635	-0.200 (0.16)	1.80E-33	-0.014 (0.03)	6.24E-01	0.014 (0.01)	5.05E-02	-0.008 (0.01)	5.10E-01
rs11602616	72,041,725	Intron	T	C	0.250	8635	-0.210 (0.18)	1.89E-30	-0.017 (0.03)	5.38E-01	0.023 (0.01)	3.53E-03	0.012 (0.01)	2.95E-01
rs11604034	72,041,795	Intron	G	C	0.250	8635	-0.210 (0.18)	1.94E-30	-0.017 (0.03)	5.35E-01	0.023 (0.01)	3.77E-03	0.012 (0.01)	3.09E-01
rs11601122	72,040,366	Intron	A	G	0.250	8635	-0.210 (0.18)	1.99E-30	-0.017 (0.03)	5.40E-01	0.023 (0.01)	3.47E-03	0.012 (0.01)	2.92E-01
rs61895588	72,038,634	Intron	A	G	0.250	8635	-0.210 (0.18)	2.13E-30	-0.017 (0.03)	5.44E-01	0.023 (0.01)	3.44E-03	0.013 (0.01)	2.89E-01
rs55765314	72,038,583	Intron	C	A	0.249	8635	-0.210 (0.18)	2.21E-30	-0.017 (0.03)	5.47E-01	0.023 (0.01)	3.43E-03	0.013 (0.01)	2.89E-01
rs11600831	72,037,868	Intron	C	T	0.249	8635	-0.210 (0.18)	2.28E-30	-0.017 (0.03)	5.49E-01	0.023 (0.01)	3.40E-03	0.013 (0.01)	2.87E-01
rs11604837	72,037,655	Intron	G	A	0.249	8635	-0.210 (0.18)	2.30E-30	-0.017 (0.03)	5.49E-01	0.023 (0.01)	3.39E-03	0.013 (0.01)	2.87E-01
rs72962169	72,043,317	Intron	C	T	0.249	8635	-0.209 (0.18)	2.47E-30	-0.017 (0.03)	5.50E-01	0.024 (0.01)	2.49E-03	0.014 (0.01)	2.26E-01
rs184654	72,042,053	Intron	C	G	0.249	8635	-0.209 (0.18)	3.71E-30	-0.015 (0.03)	5.91E-01	0.023 (0.01)	3.27E-03	0.013 (0.01)	2.78E-01
rs11600585	72,405,488	Intron	A	G	0.340	8635	-0.183 (0.17)	4.06E-28	-0.009 (0.02)	7.06E-01	0.012 (0.01)	8.41E-02	-0.007 (0.01)	5.20E-01
rs7109082	72,385,089	Intron	A	C	0.340	8635	-0.183 (0.17)	4.37E-28	-0.009 (0.02)	7.13E-01	0.012 (0.01)	8.31E-02	-0.007 (0.01)	5.25E-01
rs3814730	72,233,201	Intron	G	A	0.332	8635	-0.181 (0.17)	2.64E-27	-0.016 (0.02)	4.84E-01	0.012 (0.01)	8.44E-02	-0.005 (0.01)	6.21E-01
rs11603313	72,265,825	Intron	G	A	0.332	8635	-0.181 (0.17)	2.78E-27	-0.016 (0.02)	4.86E-01	0.012 (0.01)	8.88E-02	-0.005 (0.01)	5.98E-01
rs17244499	72,307,594	Intron	A	G	0.332	8635	-0.181 (0.17)	2.94E-27	-0.016 (0.02)	4.88E-01	0.012 (0.01)	9.05E-02	-0.005 (0.01)	5.89E-01
rs74439784	72,327,044	Intron	A	T	0.332	8635	-0.181 (0.17)	3.14E-27	-0.016 (0.02)	4.96E-01	0.012 (0.01)	9.04E-02	-0.005 (0.01)	5.89E-01
rs12362059	72,091,332	Intron	C	T	0.353	8635	-0.177 (0.16)	7.68E-27	-0.002 (0.02)	9.30E-01	0.018 (0.01)	9.40E-03	0.007 (0.01)	4.99E-01
rs12575364	72,086,837	Intron	C	T	0.352	8635	-0.176 (0.16)	9.90E-27	-0.001 (0.02)	9.78E-01	0.018 (0.01)	9.71E-03	0.007 (0.01)	5.13E-01
rs61895574	72,025,636	Intron	A	T	0.228	8635	-0.201 (0.19)	1.10E-26	-0.021 (0.03)	4.16E-01	0.022 (0.01)	5.61E-03	0.011 (0.01)	3.09E-01

Table 6. SNP Association with Fasting Proinsulin and Type 2 Diabetes in the METSIM Study

SNP	Chr11 Position (hg18)	Annotation	EA	NEA	MAF	N	Fasting Proinsulin Association				Type 2 Diabetes Association			
							Unconditioned		Conditioned on rs11603334		Unconditioned		Conditioned on rs11603334	
							β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p
rs56200889	72,085,703	Missense ARAP1 Gln802Glu	G	C	0.355	8621	-0.175 (0.16)	2.07E-26	0.002 (0.02)	9.48E-01	0.018 (0.01)	1.01E-02	0.007 (0.01)	5.25E-01
rs10898864	72,083,098	Intron	C	T	0.360	8630	-0.172 (0.16)	7.61E-26	0.002 (0.02)	9.39E-01	0.016 (0.01)	1.89E-02	0.003 (0.01)	7.40E-01
rs1783598	72,529,111	Intron	T	C	0.363	8635	-0.172 (0.16)	1.04E-25	-0.011 (0.02)	6.17E-01	0.016 (0.01)	2.24E-02	0.003 (0.01)	7.33E-01
rs12292949	72,135,179	Intron	C	T	0.416	8635	-0.161 (0.16)	7.58E-24	-0.004 (0.02)	8.43E-01	0.011 (0.01)	1.14E-01	-0.005 (0.01)	5.89E-01
rs67260737	72,135,844	Intron	A	G	0.415	8635	-0.160 (0.16)	1.56E-23	-0.003 (0.02)	8.99E-01	0.011 (0.01)	1.21E-01	-0.005 (0.01)	5.65E-01
Top SNPs (p < 0.001) associated with fasting proinsulin after conditional analysis on rs11603334														
rs202137453	72,144,438	Missense STARD10 Pro196Ser	G	A	0.003	8621	0.319 (0.13)	1.54E-02	0.514 (0.13)	8.46E-05	0.062	2.72E-01	0.045	4.29E-01
rs191209601	72,538,773	Intergenic	T	C	0.008	8635	-0.252 (0.10)	8.28E-03	-0.327 (0.09)	5.20E-04	-0.03	4.38E-01	-0.024	5.38E-01
Additional genotyped nonsynonymous SNPs within 500 kb of rs11603334														
rs200273651	72,147,966	Missense STARD10 Ala106Thr	C	T	0.003	8621	0.189 (0.15)	1.99E-01	0.110 (0.15)	4.46E-01	0.025	6.98E-01	0.11	4.46E-01
rs139124425	72,084,464	Missense ARAP1 Lys878Gln	T	G	0.005	8621	0.077 (0.11)	4.88E-01	0.007 (0.11)	9.49E-01	0.038	4.21E-01	0.007	9.49E-01
rs143644813	72,083,694	Missense ARAP1 Val980Ile	C	T	0.001	8621	0.115 (0.23)	6.18E-01	0.089 (0.23)	6.97E-01	-0.07	5.14E-01	0.089	6.97E-01
rs148377320	72,099,229	Missense ARAP1 Arg177His	C	T	0.020	8619	0.027 (0.06)	6.26E-01	-0.025 (0.06)	6.54E-01	-0.009	7.05E-01	-0.025	6.54E-01
rs200662872	72,101,242	Missense ARAP1 Arg11Gln	C	T	0.000	8621	-0.216 (0.45)	6.29E-01	-0.109 (0.44)	8.04E-01	-0.16	4.71E-01	-0.109	8.04E-01
rs149248685	72,099,211	Missense ARAP1 Ser183Phe	G	A	0.002	8621	0.059 (0.16)	7.05E-01	-0.012 (0.15)	9.39E-01	-0.036	5.80E-01	-0.012	9.39E-01

Table 6. SNP Association with Fasting Proinsulin and Type 2 Diabetes in the METSIM Study

SNP	Chr11 Position (hg18)	Annotation	EA	NEA	MAF	Fasting Proinsulin Association					Type 2 Diabetes Association		
						N	Unconditioned		Conditioned on rs11603334		Unconditioned		Conditioned on rs11603334	
							β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p
rs150013910	72,091,645	Missense ARAP1 Ala462Val	G	A	0.001	8621	-0.083 (0.27)	7.58E-01	-0.120 (0.27)	6.53E-01	-0.109	3.33E-01	-0.12	6.53E-01
rs143097446	71,761,685	Missense CLPB Ser223Asn	C	T	0.006	8621	-0.023 (0.10)	8.17E-01	-0.095 (0.10)	3.26E-01	-0.024	5.65E-01	-0.095	3.26E-01
rs341047	71,979,151	Missense PDE2A Thr224Ile	G	A	0.158	8621	-0.004 (0.02)	8.44E-01	0.020 (0.02)	3.43E-01	-0.008	3.68E-01	0.02	3.43E-01
rs150343959	71,682,301	Missense CLPB Arg628Cys	G	A	0.001	8621	0.002 (0.30)	9.96E-01	-0.035 (0.30)	9.07E-01	-0.004	9.76E-01	-0.035	9.07E-01

Evidence of SNP association with fasting proinsulin and type 2 diabetes in up to 8,635 non-diabetic individuals (fasting proinsulin), and in 1,389 cases and 5,748 normal glycemic controls (type 2 diabetes) from the METSIM study. Beta coefficients (β) for proinsulin represent the change in inverse-normalized fasting proinsulin levels after adjustment for age, BMI, and fasting insulin. Coefficients for type 2 diabetes are not directly interpretable; however, the signs indicate direction of effect, i.e., a positive sign indicates that the effect allele increases type 2 diabetes risk. SNPs were either directly genotyped or imputed using a Central-Northern European reference panel (see Methods). Call rates for directly genotyped SNPs exceeded 0.995 and were consistent with Hardy-Weinberg equilibrium ($p > 0.01$). Imputation quality scores exceeded 0.990. Positions on chromosome 11 are shown in hg18 coordinates; SNP names are from dbSNP 137. SNPs in **bold** represent lead associated SNPs in this and previous association studies. EA, effect allele (major allele); NEA, non-effect allele (minor allele); MAF, minor allele frequency.

Table 7. Twenty Variants in Strong Linkage Disequilibrium ($r^2 \geq 0.8$) with rs11603334 and rs1552224

Variant	Chr11 Position (hg18)	Genomic Annotation	r^2 [a]	Open Chromatin in Human Pancreatic Islets[b]	Histone Modifications in Human Pancreatic Islets[c]	Transcription Factor Binding[b]
rs11605691	72,097,162	Intron of *ARAP1*	0.85	-	-	-
rs9667947	72,102,282	Intron of *ARAP1*	0.88	-	-	-
rs76550717	72,105,820	Intron of *ARAP1*	0.99	-	-	-
rs75896506	72,107,227	Intron of *ARAP1*	1.0	-	-	EBF (GM12878)
rs140717197[d]	72,107,445-7	Intron of *ARAP1*	1.0	-	-	EBF (GM12878)
rs11603334	**72,110,633**	**5'UTR near *ARAP1* P1 promoter**	**1.0**	**DNase**	**H3K9ac**	**NFKB (GM12878); Ini1 (HeLa-S3)**
rs1552224	**72,110,746**	**5'UTR near *ARAP1* P1 promoter**	**1.0**	**DNase**	**H3K4me3, H3K9ac**	**NFKB (GM12878)**
rs7129793	72,133,533	Intron of *ARAP1*	0.92	-	-	-
rs74333814	72,135,135	Intron of *ARAP1*	0.95	-	-	-
rs77464186	**72,138,046**	**Intron of *ARAP1***	**0.95**	**DNase**	-	**HNF4α, RXRα, p300 (HepG2); JunD (HepG2, GM12878); NFKB, IRF4, BCL11A, BATF, POU2F2, SP1 (GM12878)**
rs11603349	72,138,342	Intron of *ARAP1*	0.91	-	-	IRF4, POU2F2 (GM12878)
rs11602873	72,138,410	Intron of *ARAP1*	0.95	-	-	-
rs142489578[d]	72,138,578-9	Intron of *ARAP1*	0.95	-	-	-
rs148527516[d]	72,139,460-1	Intron of *ARAP1*	0.95	DNase	-	HEY1 (HepG2)
rs7109575	**72,141,083**	**TSS near *ARAP1* P2 promoter**	**0.94**	**DNase**	-	**Ini1, BAF155, BAF170, c-Myc (HeLa-S3); HNF4α, FOSL2, CEBPβ, Sin3Ak-20 (HepG2); PAX5, TCF12, SP1 (GM12878); HEY1 (HepG2, K562); PU.1(GM12891); TAF1 (HepG2, H1-hESC); MAX (HeLa-S3, K562, HUVEC); NFKB (GM15510, GM12891); USF-1 (GM12878, HepG2, K562, A549)**
rs140130268[d]	72,148,564-7	Intron of *STARD10*	0.92	-	-	-

Table 7. Twenty Variants in Strong Linkage Disequilibrium ($r^2 \geq 0.8$) with rs11603334 and rs1552224

Variant	Chr11 Position (hg18)	Genomic Annotation	r^2 [a]	Open Chromatin in Human Pancreatic Islets[b]	Histone Modifications in Human Pancreatic Islets[c]	Transcription Factor Binding[b]
rs79430446	72,151,652	Intron of *STARD10*	0.82	-	-	EBF, NFKB, PU.1, BCL3 (GM12878)
rs140735484[d]	72,151,830	Intron of *STARD10*	0.82	DNase	H3K4me1	EBF, NFKB, PU.1, BCL3, BCL11A, PAX5, TCF12 (GM12878)
rs613937	72,152,487	Intron of *STARD10*	0.80	-	-	PAX5, NFKB (GM12878)
rs77106637	72,155,968	Intron of *STARD10*	0.80	-	-	HNF4α (HepG2)

SNPs in **bold** were evaluated for differential allelic transcriptional activity in this study. TSS, transcription start site.
[a]with rs11603334, EUR, 1000 Genomes, Phase 1. [b]ENCODE peak calls, hg18. [c]Human Epigenome Atlas, Release 1 (hg18); called if signal > ½ locus maximum. [d]Indels

CHAPTER 3: ARAP1 REGULATES GLUCOSE-STIMULATED PROINSULIN SECRETION AND ARF6 GTPASE ACTIVITY IN HUMAN PANCREATIC ISLETS[iii]

Introduction

Previously, we presented evidence obtained from allelic expression imbalance and transcriptional reporter assays to suggest that the C allele of 5'UTR variant rs11603334, associated with decreased fasting proinsulin levels and increased risk of T2D, may function to upregulate expression of *ARAP1* [MIM 606646] isoforms transcribed from the P1 promoter.[93] *ARAP1* RefSeq isoforms NM_015242 and NM_001135190, corresponding to proteins NP_056057 and NP_001128662, respectively, are transcribed from the P1 promoter and TSS located at 72.11 Mb (hg18) on chromosome 11. ARAP1 protein isoform NP_001128662 differs from NP_056057 in that it contains an altered ARF-GAP domain and an altered C-terminal PH domain. A third *ARAP1* RefSeq isoform, NM_001040118, corresponding to protein NP_001035207, is transcribed from the P2 promoter and TSS located at 72.14 Mb. To date, ARAP1 protein localization and isoform expression in islets have not been characterized, and the biological role of the ARAP1 isoform we previously implicated as a functional candidate remains uninvestigated with respect to proinsulin and insulin secretion from the islet.

ARAP1 (ARF-GAP with RHO-GAP domain, ankyrin repeat and PH domain 1) is a GTPase activating protein (GAP) that regulates the activity of ARF and RHO family GTPases by helping to

[iii]Jennifer Kulzer generated the adenoviral vectors; designed, performed and analyzed the glucose-stimulated secretion assays, ELISAs, G-LISA GTPase assays, and the immunofluorescence studies; designed and analyzed the Western blots. Rani Vadlamudi performed Western blots, ELISAs, and G-LISA GTPase assays. Rachel McMullan performed Western blots.

catalyze the ability of these GTPases to hydrolyze bound GTP to GDP.[18] Since GTP-bound GTPases are usually considered 'active', while GDP-bound GTPases are usually considered 'inactive', GAP regulation typically leads to GTPase inactivation. ARF and RHO family GTPases are known to regulate Golgi transport, membrane trafficking, and actin cytoskeleton dynamics, broad cellular processes of established importance to the transport, processing, and secretion of insulin in the pancreatic beta cell.

Endogenous ARAP1 has been located near the Golgi in resting HeLa cells[18,34] and has been observed to relocate to the cytoplasm and at the plasma membrane in spreading HeLa cells and NIH3T3 mouse fibroblasts;[18] however, its location in the beta cell has not been examined. *In vitro* or in cell culture using other cell types, others have demonstrated the ability of ARAP1 to help regulate GTP-to-GDP hydrolysis for five GTP-binding proteins: ARF1, ARF5, ARF6, CDC42, and RHOA;[18,32] however, ARAP1 remains an understudied protein and a consensus for its ubiquitous activity on these substrates has not been reached. Specifically, the GTPases regulated by ARAP1, if any, in the pancreatic beta cell and islet are unknown. Of the reported ARAP1 ARF and RHO GTPase substrates in other cell types, ARF6 and CDC42, which are both activated within 3 minutes of glucose stimulation,[40, 42] have been shown to play roles in glucose-stimulated insulin secretion.[40–42] However, their roles in proinsulin secreted have not been investigated. Further, the GAPs that regulate ARF6 and CDC42 in in the beta cell or islet have not been identified.[94]

Here, we characterize ARAP1 localization, activity, and isoform expression in human pancreatic islets, and we are the first to initiate experiments to test the hypothesis that increasing ARAP1 (NP_056057) expression in pancreatic beta cells might lead to decreased glucose-stimulated proinsulin and/or insulin secretion.

Materials and Methods

Human Islets and Cell Culture

Human islets from non-diabetic organ donors were obtained from the National Disease Research Interchange (NDRI) and the Islet Cell Resource/Integrated Islet Distribution Program (ICR/IIDP). Upon arrival, islets were equilibrated for one hour at 37°C and 5% CO_2 prior to being counted and partitioned for use. Batches of islets that were used for DNA and protein lysate samples were rinsed with phosphate buffered saline (PBS), flash frozen in a dry ice and ethanol bath, and stored at -80°C until use. Intact islets used for functional experiments were transferred to Connaught Medical Research Laboratories (CMRL) medium supplemented with 10% FBS, 2 mM L-glutamine, 10 mM HEPES, 100 IU penicillin, and 100 μg/mL streptomycin, and cultured for 1-2 days at 37°C and 5% CO_2 prior to use. Islets used as dispersed cells for glucose-stimulated insulin secretion experiments were cultured in 7 mM glucose RPMI medium supplemented with 10% FBS, 100 IU penicillin, and 100 μg/mL streptomycin for 6 days prior to dispersion. The MIN6 mouse insulinoma beta cell line[66] was grown in DMEM medium (Sigma, St. Louis, MO) supplemented with 10% FBS, 1 mM sodium pyruvate, and 100 μM 2-mercaptoethanol. INS-1 and INS-1-derived 832/13[67] rat insulinoma beta cell lines (from C.B. Newgard, Duke University, Durham, NC) were grown in RPMI 1640 culture medium (Corning Cellgro, Mannassas, VA) supplemented with 10% FBS, 2 mM L-glutamine, 1 mM sodium pyruvate, 10 mM HEPES, and 0.05 mM 2-mercaptoethanol. Cell lines and islets were maintained and treated at 37°C and 5% CO_2.

Determination of Endogenous ARAP1 Isoforms

DNA samples were extracted from 23 primary human pancreatic islet samples using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI). Samples were genotyped using the validated TaqMan SNP Genotyping Assay for rs1552224 (Life Technologies, Grand Island, NY) according to the manufacturer's instructions. Protein lysates were collected using batches of 2000-4000 islets from five rs1552224 major-allele homozygotes and five heterozygotes. Islets were lysed in ice-cold NP-40 lysis buffer and passed through a syringe equipped with several increasing-gauge needles to mechanically break up the islet cells. Lysates were then clarified at 16,000 x g in a 4°C microcentrifuge and protein concentrations were determined using the Pierce BCA Protein Assay (Thermo Scientific, Waltham, MA). Islet cell lysates were flash frozen in a dry ice and ethanol bath and stored at -80°C until use. Western blots were carried out using the NuPAGE Novex System coupled with the NuPAGE Novex 3-8% Tris-Acetate Gels and the NuPAGE Novex Tris-Acetate SDS Running Buffer (Life Technologies) to resolve the higher molecular weight ARAP1 protein isoforms. Proteins were transferred onto Invitrolon polyvinylidene difluoride (PVDF) membranes using NuPAGE Transfer Buffer (Life Technologies). Polyclonal antibody to ARAP1 (ab5912) was purchased from Abcam (Cambridge, England) and used at a final concentration of 1 μg/ml. After immunodetection of ARAP1, the membrane was washed twice for 10 minutes in tris-buffered saline (TBS)-Tween containing 1.5% Glycine, 0.1% SDS, and 1% Tween-20 on a shaker to strip primary and secondary antibodies. Following stripping, the membrane was re-probed to detect low molecular weight control proteins. Antibodies to GAPDH (Santa Cruz Biotechnology, Dallas, TX) or β-Actin (Novus Biologicals, Littleton, CO) were used at a final concentrations of 1 μg/ml. Chemiluminescent immunodetection of secondary antibodies conjugated with horseradish peroxidase (HRP) was performed using

WesternBright Quantam and WesternBright Peroxide (Advansta, Menlo Park, CA). Membranes were developed on film or scanned using LI-COR C-DiGit Blot Scanner and densitometry analyses were performed using Image Studio software (LI-COR Biosciences, Lincoln, NE).

Adenoviral Vector Preparation

Full-length human *ARAP1* corresponding to the RefSeq isoform NM_015242 transcribed from the *ARAP1* P1 promoter was purchased in the Gateway Ultimate ORF pENTR221 entry clone backbone (Life Technologies, Grand Island, NY). Enhanced green fluorescent protein (EGFP) in the pENTR4 entry clone backbone was a gift from V. Bautch (University of North Carolina at Chapel Hill). We used site-directed mutagenesis (QuikChange II XL Site-Directed Mutagenesis Kit, Agilent, Santa Clara, CA) to change one or both of the previously published[18] catalytically active arginine residues to lysine residues in each of the ARAP1 GAP domains (positions 333 in the ARF-GAP domain and 748 in the RHO-GAP domain of isoform NP_056057, corresponding to positions 338 and 753, respectively, of the prior ARAP1 GenBank accession number AY049732). We also changed the terminal stop codon (position 1206) following the final coding amino acid, a valine (position 1205) to an alanine to allow translation through the C-terminal V5 tag that is included in the cytomegalovirus (CMV) promoter-driven Gateway pAd/CMV/V5-DEST destination plasmids. Gateway recombination was performed using LR Clonase II (Life Technologies), and recombined plasmids were most efficiently transformed using electroporation (GenePulsar II, Biorad, Hercules, CA) into ElectroMAX Stbl4 Competent Cells (Life Technologies). For each construct, several independent plasmid clones were selected. The genotype and fidelity of each cloned ARAP1 variant was verified by sequencing. Adenoviral vectors (Ad-CMV) expressing human ARAP1(wt), ARAP1(R333K),

ARAP1(R748K), ARAP1(R333K•R748K), ARAP1(wt)-V5, ARAP1(R333K)-V5, ARAP1(R748K)-V5, ARAP1 R333K•R748K)-V5, and EGFP were generated from the adenoviral plasmids in 293A cells with the ViraPower Adenoviral Expression System according to the manufacturer's instructions (Life Technologies). At least two independent adenoviral vectors were prepared from independent plasmid clones for each ARAP1 variant. For EGFP, two independent adenoviral preparations were generated from the same plasmid clone. After final amplification, all adenoviral vectors were simultaneously titrated using a plaque formation assay and then screened for the presence of wild-type replication-competent adenovirus using a supernatant rescue assay.[95]

Adenoviral Vector Transduction

Intact human islets were incubated overnight in 3 ml of normal growth medium containing 10 million plaque-forming units (PFU)/ml adenoviral vector and 10 μg/ml polybrene in separate 35 mm non-treated dishes. Dispersed human islet cells used for glucose-stimulated insulin secretion experiments were dissociated from whole islets with a 10 minute incubation in Accutase (Life Technologies, Grand Island, NY) in a 37°C water bath, followed by gentle pipetting. The cells were then collected by centrifugation at 1000 rpm for 5 minutes, resuspended in complete CMRL human islet culture medium, and seeded at a density of 5,000 cells per well in non-treated Nunc 96-Well Polystyrene Conical Bottom MicroWell Plates (Thermo Scientific, Waltham, MA). To transduce, dispersed cells were incubated overnight in 100 μl of normal growth medium containing adenoviral vector at a concentration of 1 million PFU/ml.

MIN6 cells were seeded at a density of 150,000 cells per well in 24-well plates or 1.5 million cells per well in 6-well plates and 832/13 cells were seeded at a density of 300,000 cells per well in 24-well plates one day prior to transduction. To transduce, MIN6 cells were incubated overnight in normal growth medium containing adenoviral vector at a concentration of 5 million PFU/ml; 832/13 cells were incubated for one to two hours in normal growth medium containing 2.5 million PFU/ml. After transduction was complete in each tissue or cell type, the media were replaced with fresh growth media. Two days after transduction, consistency of transduction efficiency was verified by inspection of EGFP positive cells and islets under an inverted fluorescent microscope. Using the above conditions, over 75% of MIN6, 832/13, and dispersed human islet cells were transduced; all intact human islets were transduced, each at approximately 50-70%.

Glucose-Stimulated Secretion of Insulin and Proinsulin

For adenoviral vector transduction of human islets presented in Figure 16, two replicate samples of 50 intact islets each were counted into separate Corning Transwell Polycarbonate Membrane Cell Culture Inserts (Sigma-Aldrich, St. Louis, MO) fitted in 24-well plates. Islets were rinsed with SAB buffer containing 114 mM NaCl, 4.7 mM KCl, 1.2 mM KH_2PO_4, 1.16 mM $MgSO_4$, 20 mM HEPES, 2.5 mM $CaCl_2$, 25.5 mM $NaHCO_3$, and 0.2% BSA, pH 7.2. After rinsing, islets were then pre-incubated in fresh SAB buffer for two hours prior to treatment with 500 µl of 3 mM or 20 mM glucose for 1 hour. After each incubation step, islets were spun twice for two minutes at 1000 rpm to allow buffer or treatments containing secreted insulin to flow through the membrane and into the wells of the plate for removal and collection. For adenoviral vector transduction of human islets presented in Figure 20, four replicate samples of

approximately 100 intact islets each were aliquoted into 1.5 ml microcentrifuge tubes, rinsed with SAB buffer, and then pre-incubated in 2.5 mM glucose in SAB buffer for one hour. After pre-incubation, islets were first treated with 500 µl of fresh 2.5 mM glucose for one hour and then treated with 500 µl of 16.7 mM glucose for one hour. After each incubation step, islets were spun for 5 minutes at 500 rpm, the buffer or treatments were removed or collected from the islets via careful pipetting, and the islets were resuspended in the next treatment. After all treatments were complete, islets were rinsed with PBS and then lysed in 40 µl NP-40 lysis buffer with protease inhibitor for measurement of total cellular insulin. Dispersed islet cells were pre-incubated with 2.5 mM glucose in SAB buffer for one hour prior to treatment with 100 µl of fresh 2.5 or 16.7 mM glucose for one hour. Before and after each incubation step, the "V" bottom 96-well plates were spun at 1000 rpm for 5 minutes to reaggregate the islet cells. Four replicate wells per glucose treatment were used for each adenoviral vector transduction. Transduced MIN6 and 832/13 cells were rinsed with SAB buffer and pre-incubated in fresh buffer for two hours prior to treatment with 500 µl of 3 mM or 20 mM glucose for 2 hours. Insulin and proinsulin levels were quantified using human- or rodent-specific enzyme-linked immunosorbent assays (ELISA) (Mercodia, Uppsala, Sweden), according to the manufacturer's instructions.

GTPase Activity Assays

Aliquots of six hundred human islets were used per sample to determine GTPase activity levels. For MIN6, two wells of a 6-well plate, or four wells of a 24-well plate were used per sample. Cell lines and islets were first rinsed with SAB buffer and then incubated in fresh SAB buffer for two hours prior to treatment with 20 mM glucose in SAB buffer for precisely 3

minutes. Sample lysates were collected immediately after 3 minutes using G-LISA Activation Assay Kit Cell Lysis Buffer and Protease Inhibitor Cocktail according to manufacturer's instructions (Cytoskeleton, Denver, CO). Sample collection for each cell type was performed in two batches in order to quickly and reliably maintain the 3 minute timepoint; batches contained one sample each per exogenously expressed construct (i.e., untranduced cells or islets, Ad-EGFP, and each Ad-ARAP1 variant) to avoid potentially confounding batch effects. Each of the islet lysate samples were also briefly passed through a syringe equipped with increasing-gauge needles to mechanically break up the islet cells. After clarification, lysates were flash frozen in liquid nitrogen and stored at -80°C prior to use. ARF6, CDC42, ARF1, and RHOA activity levels were determined using corresponding absorbance-based G-LISA Activation Assay Kits according to manufacturer's instructions (Cytoskeleton), with one deviation: CDC42 samples were collected and diluted with the ARF1 or RHOA G-LISA Cell Lysis Buffers due to excessive lysing of the beta cell lines using the doubly concentrated CDC42 G-LISA Cell Lysis Buffer. Separate aliquots of 400 human islets transduced with each adenoviral vector were rinsed in PBS, lysed in NP-40 lysis buffer with Halt Protease Inhibitor Cocktail (Thermo Scientific, Waltham, MA), snap-frozen, and reserved for Western blot determination of exogenous ARAP1-V5 and total endogenous ARF6 levels. Western blots were performed as described above; primary antibody to V5 (Life Technologies, Grand Island, NY) was used at a 1:5,000 dilution.

Immunofluorescence

Human islets were dispersed in batches of 500-1000 by incubation with 1 ml Trypsin-EDTA (Sigma, St. Louis, MO) for 5-7 minutes at 37°C and 5% CO_2 with occasional passage through a 1 ml pipette tip. Dispersed human islet cells were collected by centrifugation at 500

rpm and resuspended in supplemented CMRL human islet culture medium. The dispersed islet cells were seeded onto Nunc Lab-tek II CC2 Chamber Slides (Thermo Scientific, Waltham, MA) and incubated in 500 µl human islet culture medium for 48 hours at 37°C and 5% CO_2 prior to treatment and fixation. On the day of fixation, the culture media was replaced with 300 µl fresh media and islet cells were pre-fixed for 5 minutes through addition of 300 µl 4% paraformaldehyde (PFA) directly into the each chamber containing culture media. The PFA-media solution was removed and the islet cells were fixed for another 15 minutes in 300 µl 4% PFA. Following fixation, islet cells were rinsed and then washed 3 times for 5 minutes each in 500 µl PBS. Islet cells were blocked for 1 hour at room temperature in 500 µl of a PBS solution containing 0.1% Triton-X, 1% BSA, and 5% donkey serum. Antibody incubations were performed in 300 µl of a PBS solution containing 0.1% Triton-X and 1% BSA. Islet cells were incubated with primary antibodies overnight at 4°C. The next day, islet cells were washed 3 times in PBS and then incubated with secondary antibodies at room temperature for 1 hour in the dark. Islet cells were again washed 3 times in PBS before the chambers were removed and 1.5 thickness cover slides (Thermo Scientific) were mounted using Prolong Gold Anti-Fade Reagent with DAPI (Life Technologies, Grand Island, NY). Polyclonal antibody to ARAP1 (ab5912) was purchased from Abcam (Cambridge, England) and used at a final concentration of 1 µg/ml. Antibody to ARF6 (ARF06) was purchased from Cytoskeleton (Denver, CO) and used at a dilution of 1:200. Antibody to insulin (C27C9) was purchased from Cell Signaling Technology and used at a dilution of 1:400. Images were obtained using an Olympus FV1000 Multiphoton Confocal microscope and the Olympus FV10-ASW software.

Results

Endogenous Expression of ARAP1 Isoforms in Human Islets

We examined endogenous expression of ARAP1 protein isoforms in primary human islets using Western Blot and a previously published,[96] commercially prepared antibody against ARAP1. In HepG2 and 293T human cell lines, the strongest band observed (Figure 14, band b) is larger than 150 kDa, consistent with the 162 kDa predicted molecular weight of the ARAP1 isoform that is expressed from P2 and contains the alternative SAM domain. In human islets and in mouse and rat beta cell lines, we observed a primary band consistent with the 136 kDa predicted molecular weight of a shorter ARAP1 isoform expressed from P1 and lacking the SAM domain (Figures 14 and 15, band a). In MIN6 cells only, a band is also detected close to 250 kDa. In the larger set of human islet samples, a strong third band was also detected around 200 kDa (Figure 15, band c); this band could be an artifact or could correspond to an unannotated isoform of ARAP1.

To determine if the relative strength of the bands differed by genotype at a T2D risk variant, we examined the relative levels of expression of ARAP1 isoforms between the T2D-risk increasing and risk decreasing alleles in 10 primary human islet samples: five samples homozygous for the C allele and five heterozygous samples at rs11603334. Relative to the consistent GAPDH expression across islet samples, ARAP1 expression was highly variable. We observed no correlation between the levels of relative expression of ARAP1 isoforms and genotype (Figure 15). Specifically, the band intensities relative to GAPDH for any of the ARAP1 isoforms, or for all three isoforms combined, were not significantly different between genotypes (all $p \geq .1$).

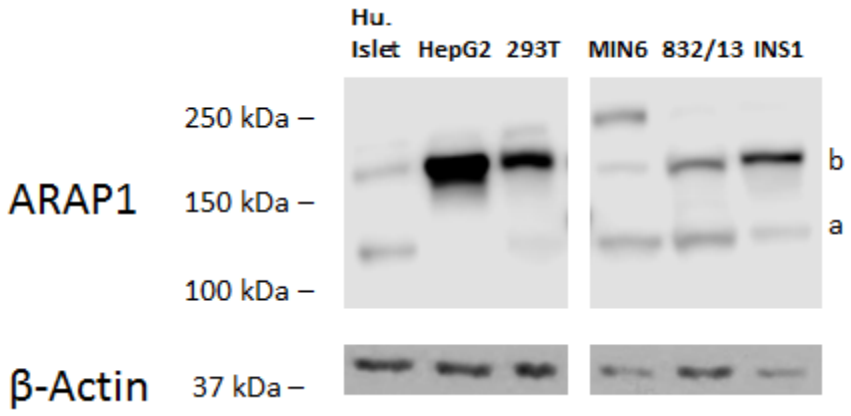

Figure 14. Endogenous ARAP1 Isoform Expression across Tissues and Species
Western blot was used to determine the endogenous levels of ARAP1 isoform expression in primary human islets, the HepG2 human hepatocyte cell line, the 293T human embryonic kidney cell line, the MIN6 mouse beta cell line and the 832/13 and INS1 rat beta cell lines. Each lane contained 100 µg total lysate.

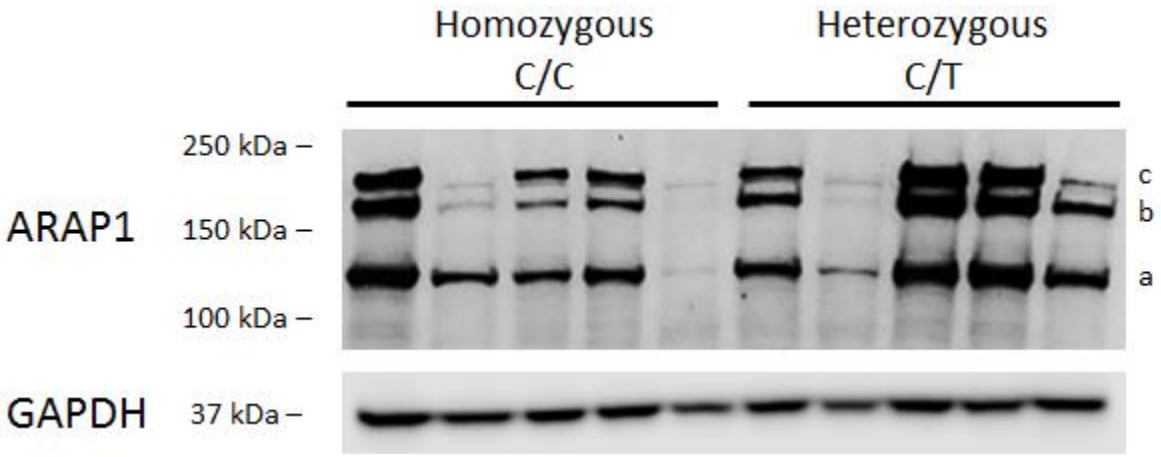

Figure 15. Endogenous Levels of ARAP1 Expression in Human Islet Samples
Western blot was used to determine the endogenous levels of ARAP1 isoform expression in 10 primary human islet samples. Five samples homozygous for the major allele (C) of rs11603334 were compared to five heterozygous samples. Each lane contained 30 µg total lysate.

ARAP1 Overexpression Decreases Glucose-Stimulated Proinsulin Secretion in Human Islets

Our previous study[93] suggested that the rs11603334 C allele, associated with increased T2D risk and decreased plasma proinsulin levels, may function to increase transcription of ARAP1 isoforms expressed from the P1 promoter. Based on this work, we hypothesized that increased expression of the ARAP1 RefSeq isoform NP_056057 (136 kDa), corresponding to P1 promoter-transcribed NM_015242, may lead to decreased glucose-stimulated proinsulin secretion in pancreatic beta cells and islets. Exogenous overexpression by plasmid transfection was inefficient in primary islets and highly variable in MIN6 and 832/13 rodent pancreatic beta cell lines (5-50% and 1-20%, respectively, data not shown). Thus, we generated adenoviral vectors expressing ARAP1 with or without a 3' V5 tag under the control of a CMV promoter. Control adenoviral vectors expressing EGFP were able to highly and reproducibly transduce both intact islets and beta cells (see Methods). The ARAP1 adenoviral vectors were used to transduce intact primary human islets, dispersed primary human islet cells, and MIN6 and 832/13 cell lines, and the levels of proinsulin and insulin secreted in response to treatment with 20 mM glucose were measured. Using intact human islets, we observed a decrease in levels of glucose-stimulated proinsulin but not insulin secretion from islets expressing exogenous ARAP1 and ARAP1-V5 compared to untransduced islets and islets transduced with EGFP ($p = .02$, Figure 16).

To determine whether the ARF-GAP or RHO-GAP domain of ARAP1 is responsible for regulating proinsulin secretion, we overexpressed ARAP1 constructs that are mutated at one or both of the previously described catalytic residues (see Methods) in these domains. We hypothesized that if increased presence and thus activity of an ARAP1 GAP domain was responsible for the decrease in proinsulin secretion, the corresponding ARAP1 GAP domain mutant should rescue levels of glucose-stimulated proinsulin secretion. Preliminary results show

a trend suggesting that mutating the catalytic residue R333K in the ARAP1 ARF-GAP domain may partially rescue glucose-stimulated proinsulin secretion; however, these results were not statistically significant ($p < .05$) (Figure 20). The R748K mutation in the ARAP1 RHO-GAP domain did not appear to rescue levels of glucose-stimulated proinsulin secretion.

Levels of glucose-stimulated insulin secretion upon ARAP1 and ARAP1-V5 over-expression did not differ significantly from intact islet controls. Glucose-stimulated insulin and proinsulin levels from dispersed primary human islet cells, MIN6, and 832/13 were not observed to be affected by adenoviral vector-mediated over-expression of ARAP1 (Figure 21, and data not shown).

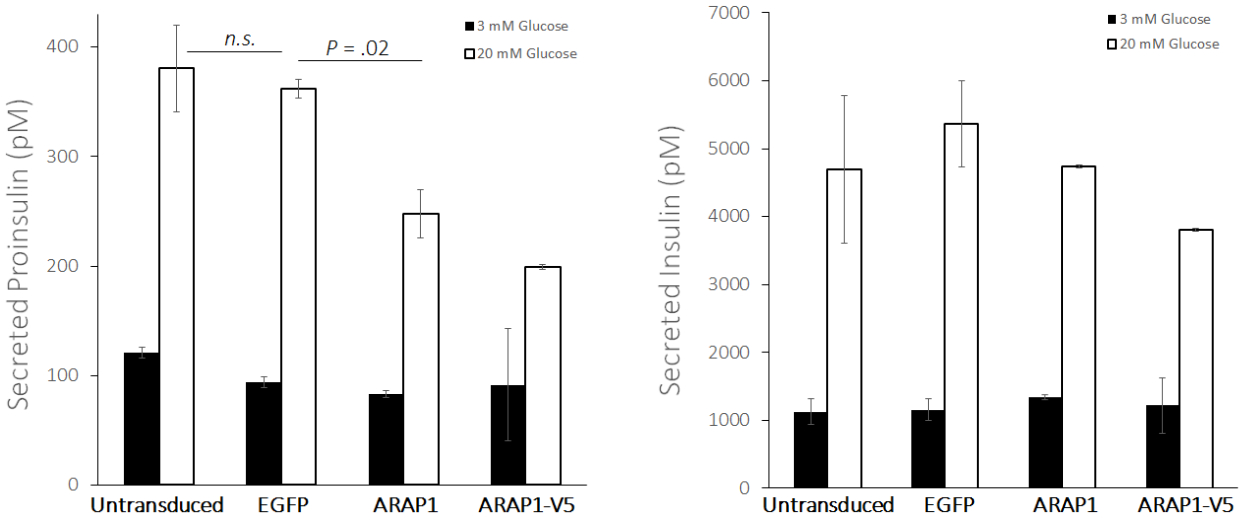

Figure 16. Proinsulin and Insulin Secretion upon ARAP1 Overexpression in Intact Human Islets

Bar charts show the concentrations of proinsulin and insulin secreted into 500 µl treatment buffer from 50 transduced islets after incubation for 1 hour at low (3 mM) and high (20 mM) glucose treatments. Each bar represents the average and standard deviation of two independent preparations of adenoviral vector (except untransduced islets); the average of each vector was determined from two biological replicate samples. The p-values were calculated with a two-sided t test. A significance threshold of $p < .05$ was used; n.s., not significant.

ARAP1 Regulates Glucose-Stimulated ARF6 GTPase Activity through its ARF-GAP Domain

We next hypothesized that increased ARAP1 expression may lower the glucose-stimulated activity of one or more small GTPases in the beta cell by catalyzing conversion of bound GTP to GDP. To identify GTPases regulated by ARAP1, we overexpressed ARAP1-V5 in MIN6 determined the levels of GTP-bound ARF6, CDC42, ARF1, and RHOA following 3 minute treatments with 20 mM glucose. ARAP1 overexpression did not appear to decrease GTP-bound levels of CDC42, ARF1, or RHOA (Figure 17). For ARF6, we observed a reproducible but non-significant trend such that cells with exogenously expressed ARAP1-V5 contained slightly decreased levels of ARF6-GTP upon glucose stimulation relative to untransduced cells and cells transduced with EGFP.

To validate regulation of ARF6-GTP levels by ARAP1, we repeated the experiment in human islets. We observed that exogenous ARAP1 expression decreased glucose-stimulated ARF6-GTP levels ($p = .02$, Figure 18, left panel). Expression of the ARAP1 ARF-GAP domain mutant, R333K, rescued levels of glucose-stimulated ARF6-GTP ($p = .02$) while the RHO-GAP domain mutant, R748K, did not. For the ARF-GAP/RHO-GAP double mutant, R333/748K, we observed a non-significant trend suggesting rescue. These data suggest that ARAP1 regulates ARF6-GTP levels in human islets through its ARF-GAP domain. For all transduced islets, the total level of endogenous ARF6 were similar (Figure 18, right panel), demonstrating that increased ARAP1 expression does not affect the total expression of ARF6.

To determine whether ARAP1 and ARF6 are located together within the beta cell, we assessed cellular localization using immunocytochemistry. In dispersed primary human islet beta cells, ARAP1 and ARF6 partially colocalized together in the cytoplasm (Figure 19A-C). In some instances, these coincident signals also overlapped discrete areas marked by the insulin

antibody; however, in most cells, insulin was fairly diffuse (Figure 19D-E). Taken together with the assays of ARF6-GTP levels, these results suggest that ARAP1, through its ARF-GAP domain, regulates glucose-stimulated ARF6 activity in primary human islets.

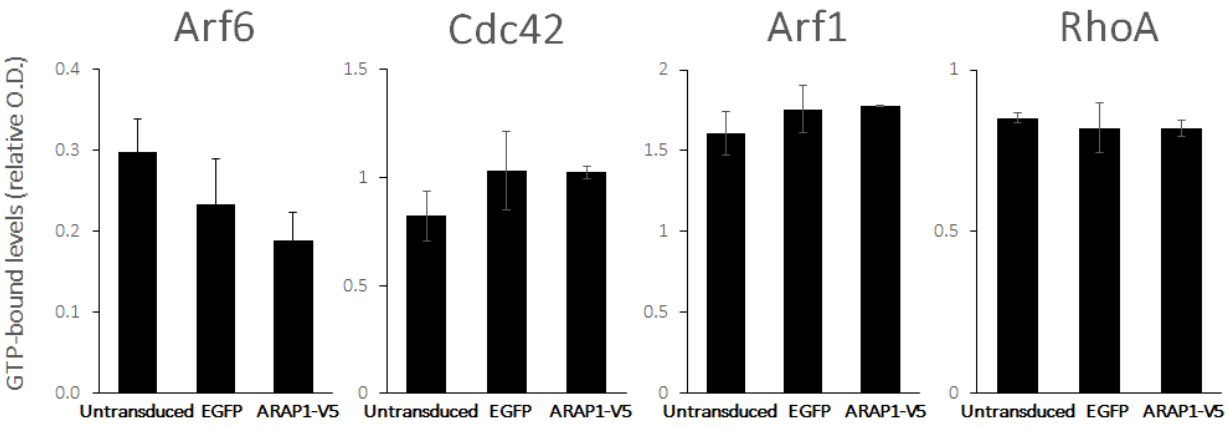

Figure 17. Effects of ARAP1 Overexpression on GTP-Bound Levels of Several Small GTPases upon Glucose Treatment in MIN6

Bar charts show the amounts of GTP-bound small GTPase contained in MIN6 cell lysates following incubation for 3 minutes in high (20 mM) glucose. GTP-bound levels were measured as relative O.D. units at 490 nm upon completion of the G-LISA GTPase activity assays. Each bar represents the average and standard deviation of two independent preparations of adenoviral vector (except untransduced islets). For each adenoviral vector, an average was first determined from two biological replicate samples, which were each technically assayed twice using G-LISA. The entire experiment was carried out twice for ARF6 and CDC42, and once for ARF1 and RHOA.

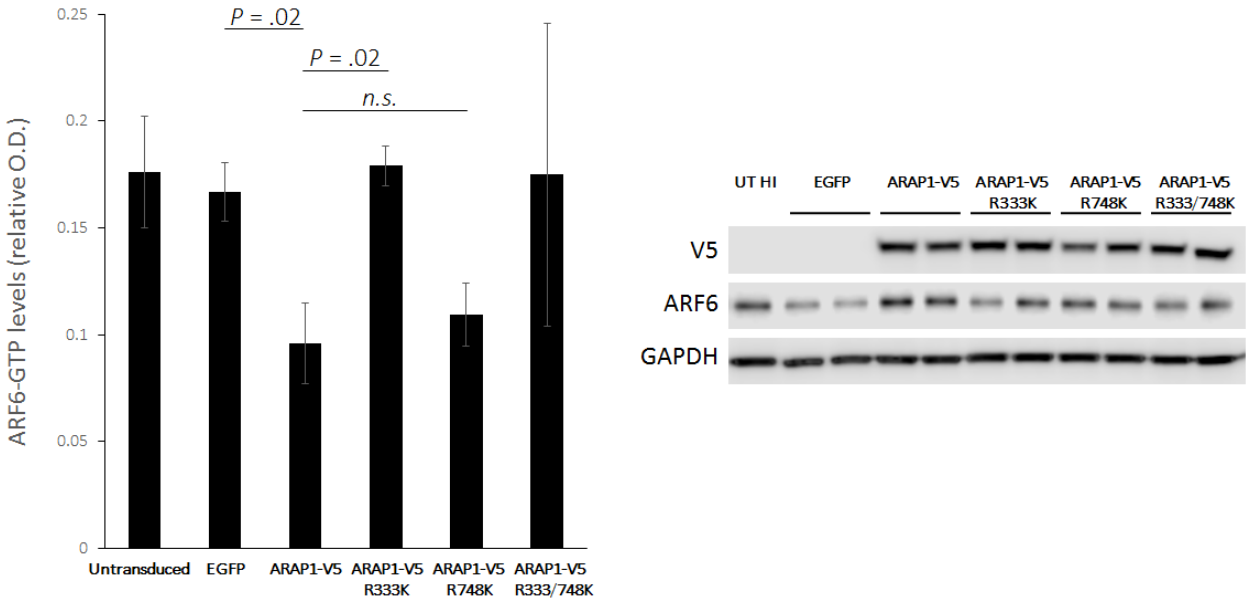

Figure 18. ARAP1 Regulates ARF6 Activity in Human Islets through its ARF-GAP Domain

The bar chart (left) shows the relative amounts of ARF6-GTP contained in lysate from 600 transduced islets following incubation for 3 minutes in high (20 mM) glucose. GTP-bound levels were measured as relative O.D. units at 490 nm upon completion of the G-LISA GTPase activity assays. Each bar represents the average and standard deviation of two independent preparations of adenoviral vector (except untransduced islets). For each adenoviral vector, an average was first determined from two biological replicate samples, which were each technically assayed twice using G-LISA. The ARAP1 R333K mutation alters the catalytic residue in the ARF-GAP domain; the R748K mutation alters the catalytic residue in the RHO-GAP domain. The p-values were calculated with a one-sided t test. A significance threshold of $p < .05$ was used; n.s., not significant. The Western blot (right) shows relative islet levels of exogenously expressed ARAP1-V5 and levels of total endogenous ARF6 and endogenous GAPDH. Each lane contained 10 µg total lysate from islets transduced with independent preparations of adenoviral vector.

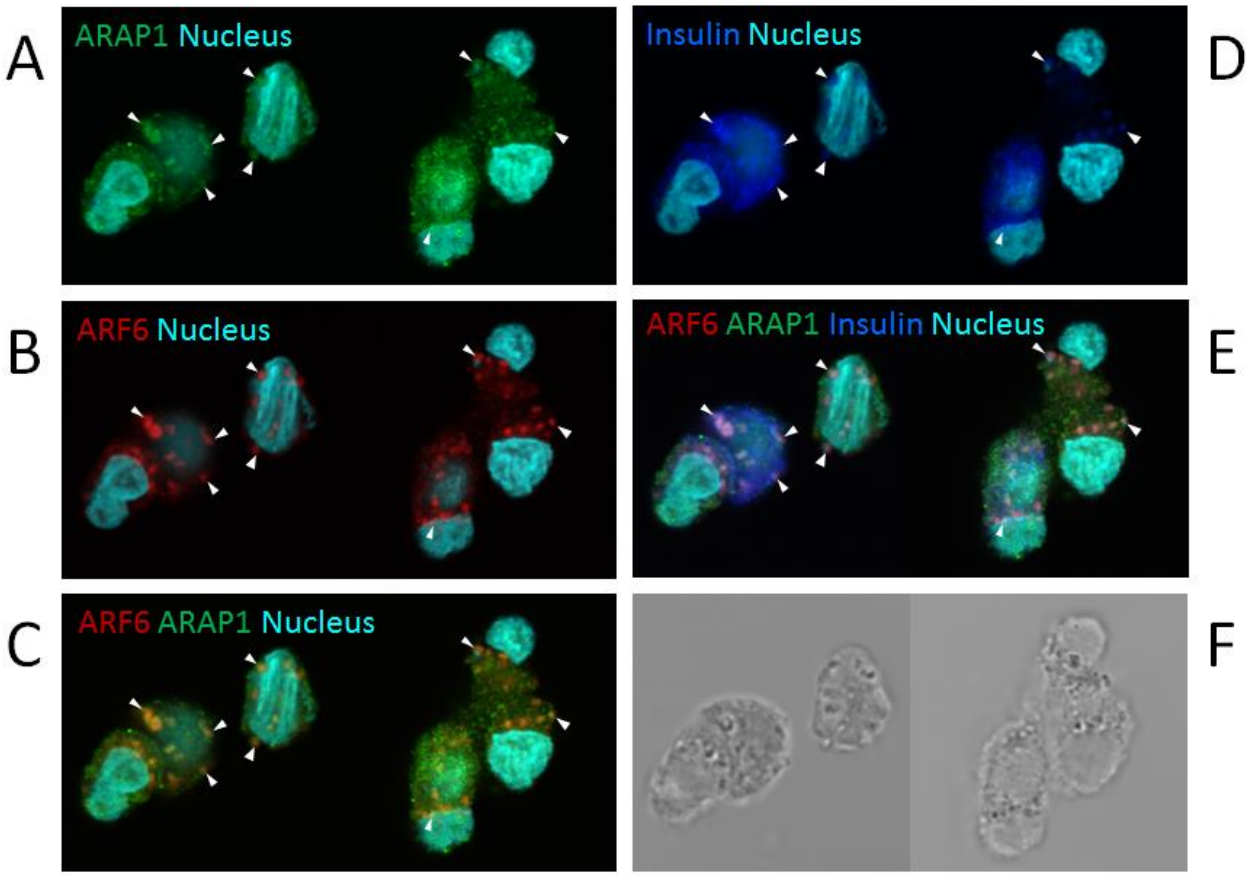

Figure 19. ARAP1 and ARF6 Partially Colocalize in the Cytoplasm in Human Beta Cells
Immunocytochemistry was used to visualize localization of endogenous ARAP1 and ARF6 in dispersed primary human beta cells. Each frame above (A-F) shows two independently captured images. Example areas of ARAP1 and ARF6 colocalization are highlighted with white arrows.

Discussion

In this study, we demonstrate that ARAP1 may play a functional role at the locus associated with type 2 diabetes and proinsulin on 11q13.4. We characterized localization and isoform expression of ARAP1 in the human pancreatic islet and showed that the 136 kDa ARAP1 RefSeq isoform NP_056057 can regulate glucose-stimulated proinsulin secretion in intact primary islets. We also identified this ARAP1 isoform as a GAP for ARF6 in islets, and showed that its ability to regulate glucose-stimulated ARF6-GTP levels is dependent on the activity of its ARF-GAP domain.

Upon examination of ARAP1 protein isoforms in human islets using an antibody targeting the C-terminal end of ARAP1, we observed two to three major bands. Two of these bands match predicted sizes of ARAP1 protein isoforms in the RefSeq database that start with exons at the P1 (136 kDa) and P2 (162 kDa) promoters. The smaller band labeled 'a' was a primary one in human islets as well as in MIN6 mouse and 832/13 and INS1 rat beta cell lines. This band may represent either or both of the similar ARAP1 protein isoforms NP_056057 and NP_001128662, which differ by two codons containing 72 amino acids, or about 8 kDa. The larger band labeled 'b', which may represent the ARAP1 protein isoform NP_001035207, was most important in human HepG2 hepatocyte and in HEK293T embryonic kidney cells, suggesting different regulation of isoform use across tissues.

The identity of the third band that we observed around 200 kDa is unknown; however, it was strong in 6 of 10 human islet samples we examined. This 200 kDa band may be an experimental artifact, or may represent a closely related protein detected by the ARAP1 antibody. Alternatively, other unannotated isoforms of ARAP1 may exist. Two other human genome references, available on the University of California Santa Cruz (UCSC) Genome Browser and from the GENCODE[97] consortium annotate a predicted non-RefSeq ARAP1 isoform that is transcribed from the STARD10 promoter on the same strand as STARD10, but contains no STARD10 coding exons. The molecular weight of this isoform is predicted to be only 161 kDa (UCSC, uc001osv.3; GENCODE, ENST00000359373.5), so this predicted isoform does not seem to explain the origin of a 200 kDa protein. Noting that ARAP1 may also be transcribed from the STARD10 promoter and that the predicted molecular weight of STARD10 is 33 kDa, we hypothesize that a 200 kDa band, if confirmed to exist, may represent a potential STARD10-ARAP1 fusion protein containing the full

sequence of both genes, which would have a theoretical size of 195 kDa. Additional experiments would be required to determine the existence of any unannotated ARAP1 protein isoforms.

We did not observe a significant difference in the relative levels of ARAP1 136 kDa isoform expression between T2D risk-increasing and T2D risk-decreasing genotypes. Our analysis of only 10 islet samples may be underpowered to detect the small predicted difference in expression of the shorter ARAP1 isoform across genotypes based on our previous allelic imbalance findings in human islet mRNA. However, the intensities of ARAP1 isoform expression relative to GAPDH differed substantially across samples, suggesting that other factors may influence ARAP1 expression or relative isoform use in human islets.

In support of our previous report implicating a possible role for increased expression of ARAP1 in decreased fasting proinsulin levels, we provide here the first experimental evidence demonstrating that ARAP1 (NP_056057) over-expression decreases the amount of proinsulin secreted by intact primary human pancreatic islets in response to a 1 hour glucose treatment. Preliminary findings also suggest that ARAP1 might exert this effect through the activity of its ARF-GAP domain, although our results were not statistically significant ($p > .05$); this experiment is worth repeating when islet resources allow. We did not observe ARAP1 over-expression to influence glucose-stimulated proinsulin secretion in MIN6 or 832/13 rodent pancreatic beta cell lines, nor did we observe an effect in dispersed primary human islets. These results suggest that ARAP1 may regulate proinsulin secretion through a pathway mediated by cell-to-cell interactions in the islet, highlighting the importance of examining beta cell function in the context of intact islets. Alternatively, species- and/or cell-line differences may explain the lack of effect in MIN6 and 832/13, and possible adenoviral vector toxicity to freshly dispersed islet cells may have obscured our ability to observe genetic effects that are expected to be more subtle. Since proinsulin is secreted in

small amounts along with mature insulin from exocytosed insulin secretory granules in the beta cell, we hypothesized that ARAP1 over-expression may also decrease insulin secretion in intact primary human islets. We did not observe a clear effect of ARAP1 over-expression on 1 hour glucose-stimulated insulin secretion; however, further investigation is warranted.

We show that adenoviral vector-mediated ARAP1 (NP_056057) over-expression decreases glucose-stimulated ARF6-GTP levels in human pancreatic islets and demonstrate that this effect is dependent on the activity of the ARAP1 ARF-GAP, but not the RHO-GAP, domain. Lending further support, we observed that ARAP1 and ARF6 partially colocalize together in the cytoplasm of dispersed human pancreatic beta cells. ARAP1 has been identified as a GAP for ARF6 *in vitro*[18] and in BHK cells,[32] and we now validate that it has GAP activity for ARF6 in human islets. Given the possibility that ARAP1 may also regulate glucose-stimulated proinsulin secretion through its ARF-GAP domain, our results suggest that ARF6 could be one GTPase through which ARAP1 may act to regulate proinsulin secretion; this would implicate a potential role for ARF6 in proinsulin secretion as well. A full investigation of ARAP1's activity for all known ARF- and RHO-GTPases, combined with a set of experiments investigating the proinsulin-specific effects of knocking down ARF6 or decreasing the glucose-stimulated activity of ARF6 may provide further insight.

This work contributes to understanding of ARAP1's biological role as a functional candidate gene in the pancreatic islet without excluding roles for ARAP1 in other T2D-relevent tissues or roles for other genes at the 11q13.4 locus. Given that expression of the 'long' ARAP1 SAM domain-containing isoform is highly preferred in HEPG2 cells, a mechanism that leads to low but aberrant expression of the potentially more active, SAM domain-lacking, 'shorter' ARAP1 isoform(s) in human hepatocytes may be biologically disruptive to liver-specific pathways. Possible involvement of a T2D-associated gene in biological pathways related to both the pancreas and the

liver is not without precedent: this may be the case with TCF7L2.[98] In the pancreas and islet, STARD10 (StAR-related lipid transfer domain containing 10) is another strongly expressed gene at the T2D-associated 11q13.4 locus, making it another logical candidate worth examining for a possible role in insulin secretion.

Supplemental Figures

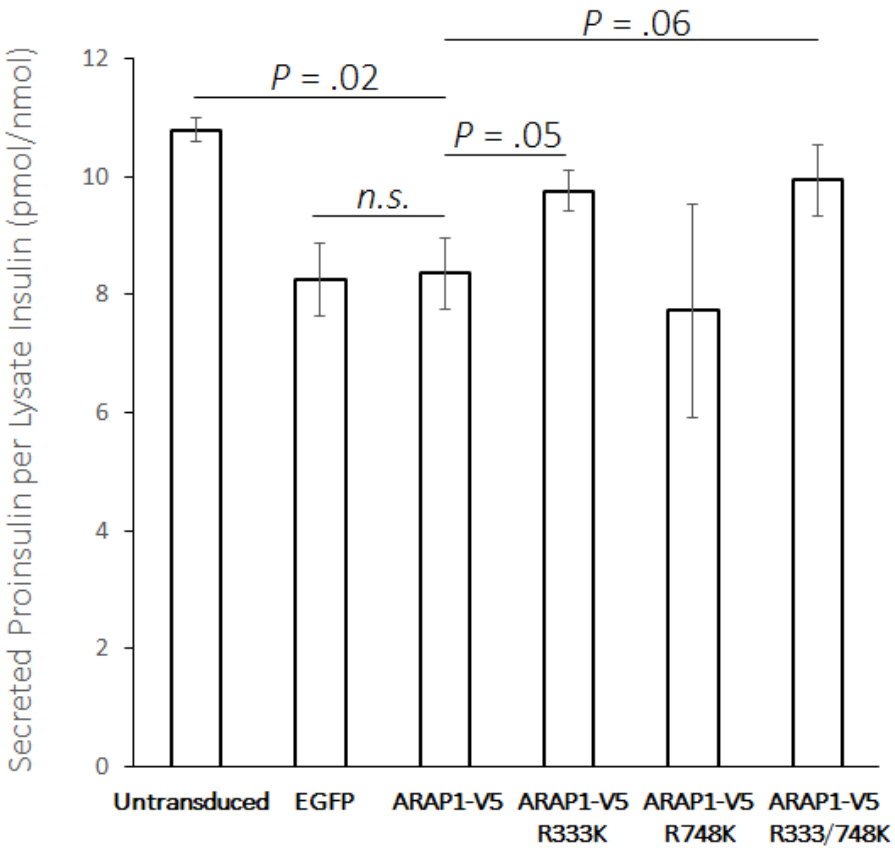

Figure 20. Glucose-Stimulated Proinsulin Secretion in Islets Transduced with ARAP1-Expressing Adenoviral Vectors
Bar charts show the levels of secreted proinsulin, normalized to total islet lysate insulin levels, from 100 transduced islets after incubation for 1 hour at high (16.7 mM) glucose. Each bar represents the average and standard deviation of two independent preparations of adenoviral vector (except untransduced islets). For each adenoviral vector, an average was first determined from four islet samples. The ARAP1 R333K mutation alters the catalytic residue in the ARF-GAP domain; the R748K mutation alters the catalytic residue in the RHO-GAP domain. The p-values were calculated with a one-sided t test; n.s., $p > 0.1$.

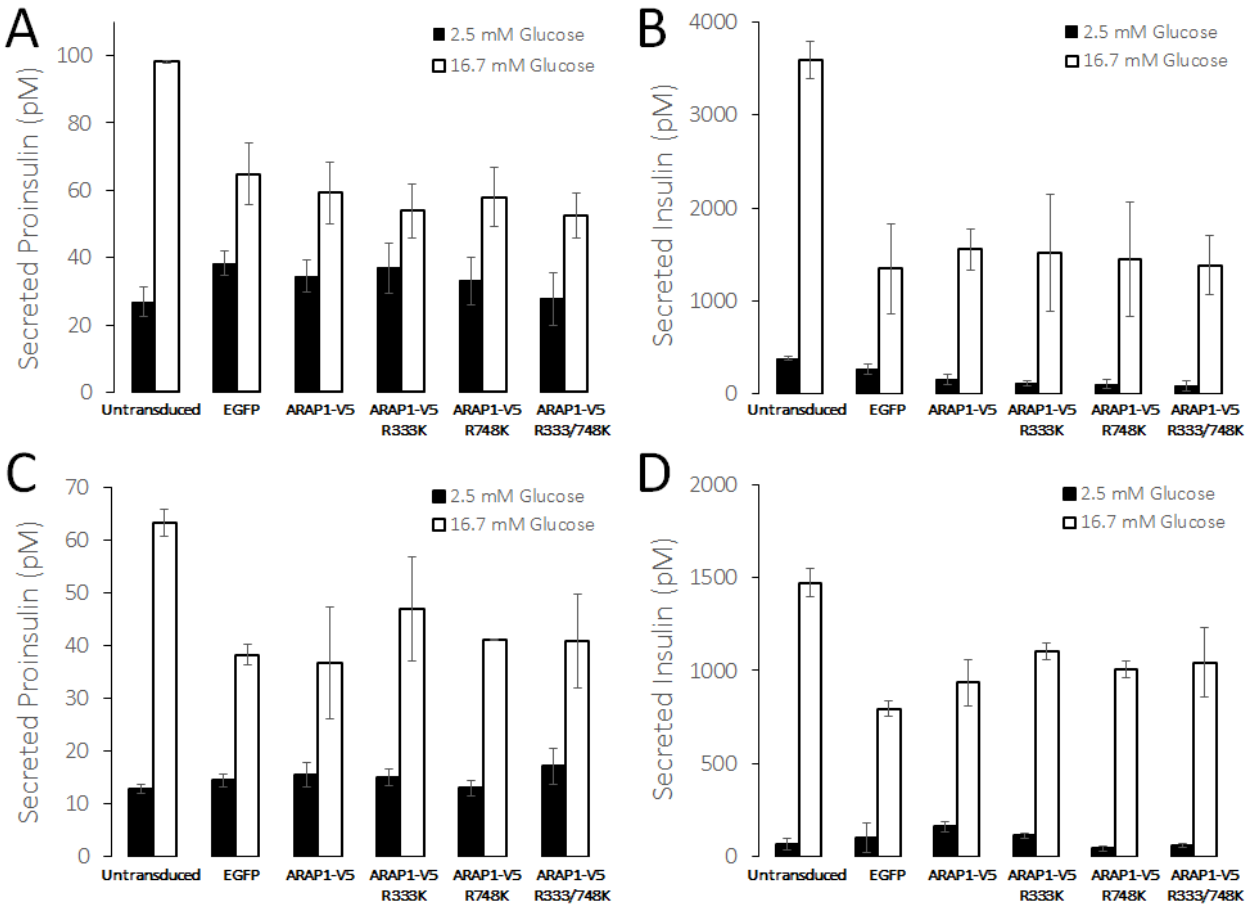

Figure 21. Glucose-Stimulated Proinsulin and Insulin Secretion in Dispersed Primary Human Islet Cells Transduced with Wild Type and Mutant ARAP1

Bar charts show the concentrations of proinsulin and insulin secreted into 50 µl treatment buffer from 5,000 dispersed islet cells after incubation for 1 hour at low (2.5 mM) and high (16.7 mM) glucose treatments. Two sets of islet cells were used; the results for the first set appear in A-B and the results for the second set appear in C-D. Each bar represents the average and standard deviation of two independent preparations of adenoviral vector (except untransduced islets). For each adenoviral vector, an average was first determined from four dispersed islet cell samples. The ARAP1 R333K mutation alters the catalytic residue in the ARF-GAP domain; the R748K mutation alters the catalytic residue in the RHO-GAP domain.

CHAPTER 4: DISCUSSION

In this dissertation, I investigated and described functional molecular and biological mechanisms that may underlie the genome-wide association of the *ARAP1* locus with proinsulin levels and T2D. Here, I discuss the implications of these results, consider the possible role of increased ARAP1 expression in T2D risk, and outline other potential mechanisms.

The rs11603334-C allele at the *ARAP1* locus is robustly associated with decreased fasting plasma proinsulin levels ($p = 3.2 \times 10^{-102}$, $n > 19{,}000$)[2] but only nominally associated with levels of decreased plasma insulin during an OGTT ($p = 0.001$, $n = 5{,}529$).[3] Recent results also show a very strong association between the *ARAP1*-increasing,[93] rs11603334-C allele and decreased plasma proinsulin levels in human subjects 30 minutes after introduction of oral glucose ($p = 4.6 \times 10^{-54}$, $n \sim 8{,}600$ Finnish males, unpublished data, Markku Laakso, Johanna Kuusisto, Alena Stančáková, Henna Cederberg). Consistent with these dramatically different strengths of association, we observed that experimental overexpression of ARAP1 in cultured intact human pancreatic islets led to decreased 1 hour glucose-stimulated proinsulin secretion, but we did not find a significant effect of ARAP1 overexpression on 1 hour glucose-stimulated insulin secretion.

In addition to the striking disparity between the strength of the proinsulin and insulin associations, the directions of effect for the T2D and plasma proinsulin associations at the *ARAP1* locus differ from other GWAS loci in an unconventional way: the allele associated with decreased measures of insulin secretion and increased risk of T2D (rs11603334-C) is associated with decreased

plasma proinsulin levels. At the *TCF7L2*, *SLC30A8*, and *VPS13C* loci, the T2D risk-increasing alleles are associated with increased proinsulin levels and are thought to interfere with insulin processing in the beta cell such that higher amounts of proinsulin are secreted in place of insulin.[2] The *ARAP1* locus phenotype associations also stand in contrast to the epidemiological observation that elevated plasma proinsulin associates with an increased risk of T2D.[99] To our knowledge, *ARAP1* is the only GWAS locus associated with both T2D and proinsulin that exhibits this unusual direction of effect.

We and others[2] have proposed that the biological mechanisms underlying these *ARAP1* locus associations might contribute to an overall defect in insulin packaging, trafficking, or exocytosis, such that the release of all forms of insulin hormone would be impaired. Such a defect would logically be expected to contribute to increased T2D susceptibility, as it would directly affect the ability of the beta cell to secrete insulin. We initially hypothesized that, in cultured β-cells or islets, glucose-stimulated secretion of both insulin and proinsulin may be decreased by experimentally increasing ARAP1 expression, postulating that the associations for fasting and OGTT proinsulin could be stronger than those of insulin as a simple consequence of proinsulin's longer half-life in the bloodstream. In this model, we expected to observe statistically significant decreases in both proinsulin and insulin secretion by islets *ex vivo* into a cell culture treatment buffer that has no ability to clear insulin. Further, since imminently exocytosing insulin granules contain much more mature insulin than immature proinsulin, the net loss in secreted insulin arising from potential global trafficking or exocytosis defects might be predicted to be even more apparent than that of secreted proinsulin. But, when we overexpressed ARAP1 in cultured intact islets, we observed a significant decrease in secreted proinsulin levels only.

What does it mean that we did not observe ARAP1 overexpression to also affect insulin secretion in intact islets? First, the experiment is worth repeating in additional human islet samples in order to validate this negative finding.

One alternative hypothesis may be that increased ARAP1 expression could cause a short delay in plasma membrane recruitment and/or exocytosis for some insulin secretory granules. This delay might afford increased opportunity for late proinsulin-to-insulin processing within the granules; and a consequence of this could be that, over a cumulative period of 30 minutes or longer, relative proinsulin levels may be decreased to a greater extent than insulin. In this model, proinsulin secretion would be decreased due to both the delay and the processing; insulin secretion would be decreased early on by the delay, but subsequently recover to a certain extent from increased processing. It is unclear how long such a delay would have to be and what proportion of insulin granules would need to be affected to demonstrate these disparate phenotypes; too long of a delay or the involvement of too many insulin granules would greatly affect insulin secretion as well as proinsulin; the kinetics may be delicately balanced. Measuring insulin and proinsulin secretion at earlier, more precise time points in the clinic and in cultured human islets *ex vivo* using islet perfusion studies may help validate a hypothesis implicating delayed granule exocytosis. On the cellular level, the kinetics of exocytosis could be examined and quantified more precisely using measurements of beta cell plasma membrane capacitance. As with a theoretical global defect in insulin granule trafficking or exocytosis, a hypothetical persistent delay in insulin and proinsulin secretion immediately upon feeding, when glucose levels are highest, would subject the body to recurrent, slightly longer periods of elevated glucose exposures and could also conceivably contribute to risk of developing T2D, especially for insulin-resistant individuals.

In human pancreatic islets, we also discovered that increased ARAP1 expression decreases the levels of glucose-stimulated ARF6-GTP. We found that the regulation of ARF6 by ARAP1 is dependent on ARAP1 ARF-GAP domain activity and noted a non-significant trend suggesting that ARAP1 regulation of proinsulin secretion might also be dependent on the ARF-GAP domain. While ARAP1 regulation of other GTPases in the human islet has not been ruled out, the above findings lead us to postulate that ARAP1 could partially influence proinsulin secretion through regulation of ARF6.

A regulatory role of ARF6 in proinsulin secretion has not yet been specifically demonstrated; however, other groups have reported that MIN6[41] and 832/13[40] cells expressing a dominant negative, non-GTP binding ARF6(T27N) mutant have reduced glucose-stimulated insulin secretion. Examining the effects of this ARF6(T27N) mutant, or the effects of ARF6 knock-down, on glucose-stimulated proinsulin levels would help determine if it is conceivable that ARAP1 could regulate proinsulin secretion through ARF6.

ARF6 has been previously linked to late-stage exocytosis, suggesting that an ARAP1/ARF6-mediated biological mechanism could be compatible with the 'delay' hypothesis. In PC12 neuroendocrine cells, ARF6 activity regulates K^+-stimulated growth hormone secretion through its ability to activate downstream effector phospholipase D1 (PLD1),[100–102] an enzyme that is reported to stimulate exocytosis by generating 'fusogenic' lipid second messenger phosphatidic acid (PA) from phosphatidylcholine, which may aid in granule fusion with the plasma membrane.[103] As others have previously proposed[94] in light of evidence in MIN6 cells that PLD1 activity is required for proper glucose stimulated insulin secretion,[104] ARF6-GTP may regulate late-stage insulin granule exocytosis through its activation of PLD1. As an upstream regulator of ARF6, ARAP1 may modulate this pathway (Figure 22). Changes in PLD activity upon ARAP1 overexpression could be

determined by using a radiolabeled phospholipid biochemical assay such as transphosphatidylation[105,106] or by using a chemically coupled commercial assay, such as the Amplex Red Phospholipase D Assay Kit (Life Technologies, Grand Island, NY).

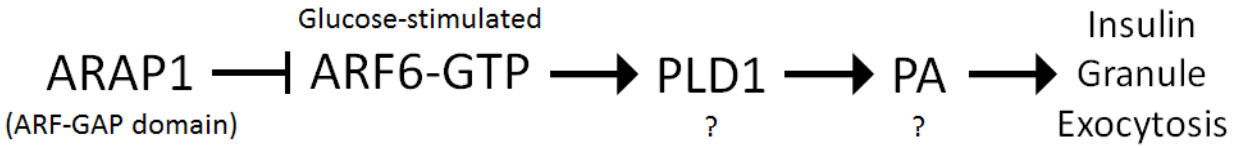

Figure 22. Proposed Model for ARAP1 regulation of Insulin Granule Exocytosis
Increased ARAP1 expression decreases glucose-stimulated ARF6-GTP levels and 1 hour proinsulin but not insulin secretion in human islets, suggesting a functional role for ARAP1 in regulating the kinetics of insulin granule exocytosis, possibly through an ARF6/PLD1/PA pathway.

A philosophical counterargument to our proposed ARAP1/ARF6-mediated 'delay' hypothesis arises when we consider that we did not observe a significant effect of ARAP1 overexpression on glucose-stimulated insulin secretion in human pancreatic islets, even though disrupting ARF6-GTP activity has been shown to decrease glucose-stimulated insulin secretion in MIN6[41] and 832/13.[40] Species differences in the nature of GTPase regulation by ARAP1 may be important considerations here, as we did not observe ARAP1 overexpression to decrease glucose-stimulated proinsulin or insulin secretion in MIN6 or 832/13. Also importantly, in the previous studies, ARF6 siRNA knockdown or overexpression of a dominant negative, non-GTP binding ARF6(T27N) mutant were used to show an effect. Relative to increased presence of ARAP1, our proposed ARF6 GAP, these ARF6-specific disruptions are fairly severe and mechanistically more proximal; they may be expected to affect the recruitment and/or exocytosis of a majority of insulin secretory granules. As discussed above, in the 'delay' hypothesis, we would anticipate that the involvement of many insulin granules would greatly affect insulin secretion as well as proinsulin.

Alternatively, increased ARAP1 expression may delay insulin granule recruitment or exocytosis through regulation of another small GTPase. ARAP1's ability to help catalyze GTP to GDP conversion for additional GTPases could be tested using G-LISA activity assays, if human islet resources allow.

There could also be more than one biological mechanism relating increased ARAP1 expression to plasma proinsulin, plasma insulin, and/or T2D susceptibility. In the islet, ARF6, or another GTPase, could be found to mediate a potential effect of increased ARAP1 expression on secretion of proinsulin in the minor constitutive pathway,[107] distinct from the role of ARF6 on glucose-stimulated insulin secretion. Meanwhile, ARAP1 could influence T2D risk by decreasing levels of one or more GTP-bound GTPases in other tissues. CDC42, for example has been demonstrated to play a role in insulin-stimulated glucose transporter 4 (GLUT4) translocation to the plasma membrane in 3T3-L1 mouse adipocytes.[43] Defects in the GLUT4 translocation/glucose uptake part of the insulin signaling pathway could contribute to insulin resistance. Examining the ability of ARAP1 to regulate levels of CDC42-GTP and other GTP-bound GTPases and measuring the effects of increased ARAP1 expression on insulin-stimulated GLUT4 translocation and glucose uptake in adipocytes, hepatocytes, and skeletal muscle cells may shed light on a potential role for ARAP1 in insulin resistance. Additionally, whole-organism ARAP1 knockout, transgenic, or *ARAP1* locus humanized mouse models would allow ascertainment of the effects of altered ARAP1 expression across T2D-relevant tissues an in conjunction with a diabetogenic diet.

We cannot rule out possible roles for other isoforms of ARAP1 or for other genes at the *ARAP1* locus in influencing proinsulin levels or T2D risk. We chose to investigate the full length ARAP1 protein isoform translated from P1 promoter-transcribed mRNA, NP_056057, rather than its closely related alternative ARAP1 protein isoform, NP_001128662, which is predicted to contain

modified ARF-GAP and C-terminal PH domains due to two missing in-frame exons in the mRNA sequence. The results from our first study examining a possible molecular mechanism underlying the *ARAP1* locus did not determine that mRNA transcript level for the 'long' SAM-domain containing isoform of ARAP1 is necessarily regulated by rs11603334; we did not test the effects of its altered expression in our follow-up study. The 'long' isoform of ARAP1 does not catalyze ARF6-GTP hydrolysis in BHK cells,[32] suggesting that the SAM domain may function as a negative regulator of ARAP1 ARF-GAP activity. For this reason, it seems less likely that this isoform would play as important a role in ARAP1-regulated phenotypes that are dependent on activity of the ARF-GAP domain, and in particular those mediated by ARF6. All of the other genes within the association signal are similarly understudied; none have yet been clearly implicated in T2D to date. STARD10 (StAR-related lipid transfer domain containing 10) binds and transfers phospholipids between membranes.[19] The particularly strong expression of *STARD10* in pancreas and islets makes it another logical candidate worth examining for a possible role in insulin granule exocytosis.

For the vast majority of the thousands of complex-trait loci identified through GWAS, the molecular and biological mechanisms underlying these associations are yet to be determined. Investigation of the molecular mechanisms at these loci will help to identify the target genes and establish the direction of their regulation in individuals at increased risk of disease. Exploration of the biological mechanisms will help us discover novel pathways involved in disease pathology that could be used to develop therapies and interventions to advance human health.

Post-GWAS functional investigation presents a multifaceted challenge: many complex trait loci consist of dozens of non-coding variants spread across several genes that may all be understudied novel candidates for the associated trait. In this set of research studies, we have modeled one way to identify the functional variant(s) and gene(s) and we begin to elucidate the

causal molecular and biological mechanisms that explain the trait associations. Integration of the available statistical, clinical, and bioinformatic research data with a thorough understanding of the available literature is critical to increasing the quality of hypotheses that can be experimentally tested. Unlike many important statistical and bioinformatic research efforts that help identify new genetic 'leads' and contribute to building good hypotheses, well-done functional studies that identify and investigate the biological role of these candidate genes require iterative hypothesis generation and experimental testing. This hypothesis-driven research is highly time-consuming and expensive. To increase our collective scientific capacity to translate the investment we have made in GWAS into developing therapies and interventions, more attention, resources, and funding need to be applied to this crucial but rate-limiting step.

REFERENCES

1. Voight, B.F., Scott, L.J., Steinthorsdottir, V., Morris, A.P., Dina, C., Welch, R.P., Zeggini, E., Huth, C., Aulchenko, Y.S., Thorleifsson, G., et al. (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat. Genet. *42*, 579–589.

2. Strawbridge, R.J., Dupuis, J., Prokopenko, I., Barker, A., Ahlqvist, E., Rybin, D., Petrie, J.R., Travers, M.E., Bouatia-Naji, N., Dimas, A.S., et al. (2011). Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes. Diabetes *60*, 2624–2634.

3. Nielsen, T., Sparsø, T., Grarup, N., Jørgensen, T., Pisinger, C., Witte, D.R., Hansen, T., and Pedersen, O. (2011). Type 2 diabetes risk allele near CENTD2 is associated with decreased glucose-stimulated insulin release. Diabetologia *54*, 1052–1056.

4. Ohshige, T., Iwata, M., Omori, S., Tanaka, Y., Hirose, H., Kaku, K., Maegawa, H., Watada, H., Kashiwagi, A., Kawamori, R., et al. (2011). Association of New Loci Identified in European Genome-Wide Association Studies with Susceptibility to Type 2 Diabetes in the Japanese. PLoS ONE *6*, e26911.

5. Centers for Disease Control and Prevention (2011). National diabetes fact sheet: national estimates and general information on diabetes and prediabtes in the United States, 2011.

6. Orci, L., Ravazzola, M., Storch, M.-J., Anderson, R.G.W., Vassalli, J.-D., and Perrelet, A. (1987). Proteolytic maturation of insulin is a post-Golgi event which occurs in acidifying clathrin-coated secretory vesicles. Cell *49*, 865–868.

7. Christopher J. Rhodes (2004). Processing of the Insulin Molecule. In Diabetes Mellitus: A Fundamental and Clinical Text, (Lippincott Williams & Wilkins),.

8. Rhodes, C.J., and Halban, P.A. (1987). Newly synthesized proinsulin/insulin and stored insulin are released from pancreatic B cells predominantly via a regulated, rather than a constitutive, pathway. J. Cell Biol. *105*, 145–153.

9. Juutilainen, A., Lehto, S., Rönnemaa, T., Pyörälä, K., and Laakso, M. (2005). Type 2 Diabetes as a "Coronary Heart Disease Equivalent." Diabetes Care *28*, 2901–2907.

10. (2006). Atlas of Diabetes (Philadelphia, PA: Current Medicine LLC).

11. Newman, B., Selby, J.V., King, M.C., Slemenda, C., Fabsitz, R., and Friedman, G.D. (1987). Concordance for type 2 (non-insulin-dependent) diabetes mellitus in male twins. Diabetologia *30*, 763–768.

12. Kaprio, J., Tuomilehto, J., Koskenvuo, M., Romanov, K., Reunanen, A., Eriksson, J., Stengård, J., and Kesäniemi, Y.A. (1992). Concordance for Type 1 (insulin-dependent) and Type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. Diabetologia *35*, 1060–1067.

13. Sladek, R., Rocheleau, G., Rung, J., Dina, C., Shen, L., Serre, D., Boutin, P., Vincent, D., Belisle, A., Hadjadj, S., et al. (2007). A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature *445*, 881–885.

14. Scott, L.J., Mohlke, K.L., Bonnycastle, L.L., Willer, C.J., Li, Y., Duren, W.L., Erdos, M.R., Stringham, H.M., Chines, P.S., Jackson, A.U., et al. (2007). A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science *316*, 1341–1345.

15. Zeggini, E., Scott, L.J., Saxena, R., Voight, B.F., Marchini, J.L., Hu, T., de Bakker, P.I.W., Abecasis, G.R., Almgren, P., Andersen, G., et al. (2008). Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat. Genet. *40*, 638–645.

16. Wu, C., Orozco, C., Boyer, J., Leglise, M., Goodale, J., Batalov, S., Hodge, C.L., Haase, J., Janes, J., Huss, J.W., 3rd, et al. (2009). BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome Biol. *10*, R130.

17. Su, A.I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K.A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., et al. (2004). A gene atlas of the mouse and human protein-encoding transcriptomes. Proc. Natl. Acad. Sci. U. S. A. *101*, 6062–6067.

18. Miura, K., Jacques, K.M., Stauffer, S., Kubosaki, A., Zhu, K., Hirsch, D.S., Resau, J., Zheng, Y., and Randazzo, P.A. (2002). ARAP1: a point of convergence for Arf and Rho signaling. Mol. Cell *9*, 109–119.

19. Olayioye, M.A., Vehring, S., Müller, P., Herrmann, A., Schiller, J., Thiele, C., Lindeman, G.J., Visvader, J.E., and Pomorski, T. (2005). StarD10, a START domain protein overexpressed in breast cancer, functions as a phospholipid transfer protein. J. Biol. Chem. *280*, 27436–27442.

20. Rosman, G.J., Martins, T.J., Sonnenburg, W.K., Beavo, J.A., Ferguson, K., and Loughney, K. (1997). Isolation and characterization of human cDNAs encoding a cGMP-stimulated 3',5'-cyclic nucleotide phosphodiesterase. Gene *191*, 89–95.

21. Matsushita, M., Suzuki, N.N., Obara, K., Fujioka, Y., Ohsumi, Y., and Inagaki, F. (2007). Structure of Atg5.Atg16, a complex essential for autophagy. J. Biol. Chem. *282*, 6763–6772.

22. Katoh, M., and Katoh, M. (2004). Identification and characterization of human FCHSD1 and FCHSD2 genes in silico. Int. J. Mol. Med. *13*, 749–754.

23. Cao, H., Yin, X., Cao, Y., Jin, Y., Wang, S., Kong, Y., Chen, Y., Gao, J., Heller, S., and Xu, Z. (2013). FCHSD1 and FCHSD2 are expressed in hair cell stereocilia and cuticular plate and regulate actin polymerization in vitro. PloS One *8*, e56516.

24. Hasseine, L.K., Hinault, C., Lebrun, P., Gautier, N., Paul-Bellon, R., and Van Obberghen, E. (2009). miR-139 impacts FoxO1 action by decreasing FoxO1 protein in mouse hepatocytes. Biochem. Biophys. Res. Commun. *390*, 1278–1282.

25. Greenawalt, D.M., Dobrin, R., Chudin, E., Hatoum, I.J., Suver, C., Beaulaurier, J., Zhang, B., Castro, V., Zhu, J., Sieberts, S.K., et al. (2011). A survey of the genetics of stomach, liver, and adipose gene expression from a morbidly obese cohort. Genome Res. *21*, 1008–1016.

26. Schultz, J., Milpetz, F., Bork, P., and Ponting, C.P. (1998). SMART, a simple modular architecture research tool: Identification of signaling domains. Proc. Natl. Acad. Sci. *95*, 5857–5864.

27. Letunic, I., Doerks, T., and Bork, P. (2011). SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Res. *40*, D302–D305.

28. Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., et al. (2011). The Pfam protein families database. Nucleic Acids Res. *40*, D290–D301.

29. Zdobnov, E.M., and Apweiler, R. (2001). InterProScan--an integration platform for the signature-recognition methods in InterPro. Bioinforma. Oxf. Engl. *17*, 847–848.

30. Lemmon, M.A. (2004). Pleckstrin homology domains: not just for phosphoinositides. Biochem. Soc. Trans. *32*, 707–711.

31. Kim, C.A., and Bowie, J.U. (2003). SAM domains: uniform structure, diversity of function. Trends Biochem. Sci. *28*, 625–628.

32. Cuthbert, E.J., Davis, K.K., and Casanova, J.E. (2008). Substrate specificities and activities of AZAP family Arf GAPs in vivo. Am. J. Physiol. Cell Physiol. *294*, C263–C270.

33. Yoon, H.-Y., Miura, K., Cuthbert, E.J., Davis, K.K., Ahvazi, B., Casanova, J.E., and Randazzo, P.A. (2006). ARAP2 effects on the actin cytoskeleton are dependent on Arf6-specific GTPase-activating-protein activity and binding to RhoA-GTP. J. Cell Sci. *119*, 4650–4666.

34. Yoon, H.-Y., Lee, J.-S., and Randazzo, P.A. (2008). ARAP1 regulates endocytosis of EGFR. Traffic Cph. Den. *9*, 2236–2252.

35. Daniele, T., Di Tullio, G., Santoro, M., Turacchio, G., and De Matteis, M.A. (2008). ARAP1 regulates EGF receptor trafficking and signalling. Traffic Cph. Den. *9*, 2221–2235.

36. Símová, S., Klíma, M., Cermak, L., Sourková, V., and Andera, L. (2008). Arf and Rho GAP adapter protein ARAP1 participates in the mobilization of TRAIL-R1/DR4 to the plasma membrane. Apoptosis Int. J. Program. Cell Death *13*, 423–436.

37. D'Souza-Schorey, C., and Chavrier, P. (2006). ARF proteins: roles in membrane traffic and beyond. Nat. Rev. Mol. Cell Biol. *7*, 347–358.

38. Donaldson, J.G., and Jackson, C.L. (2011). ARF family G proteins and their regulators: roles in membrane transport, development and disease. Nat. Rev. Mol. Cell Biol. *12*, 362–375.

39. Sadakata, T., Shinoda, Y., Sekine, Y., Saruta, C., Itakura, M., Takahashi, M., and Furuichi, T. (2010). Interaction of Calcium-Dependent Activator Protein for Secretion 1 (CAPS1) with the Class II ADP-Ribosylation Factor Small GTPases Is Required for Dense-Core Vesicle Trafficking in the Trans-Golgi Network. J. Biol. Chem. *285*, 38710–38719.

40. Jayaram, B., Syed, I., Kyathanahalli, C.N., Rhodes, C.J., and Kowluru, A. (2011). Arf nucleotide binding site opener [ARNO] promotes sequential activation of Arf6, Cdc42 and Rac1 and insulin secretion in INS 832/13 β-cells and rat islets. Biochem. Pharmacol. *81*, 1016–1027.

41. Lawrence, J.T.R., and Birnbaum, M.J. (2003). ADP-ribosylation factor 6 regulates insulin secretion through plasma membrane phosphatidylinositol 4,5-bisphosphate. Proc. Natl. Acad. Sci. U. S. A. *100*, 13320–13325.

42. Wang, Z., Oh, E., and Thurmond, D.C. (2007). Glucose-stimulated Cdc42 signaling is essential for the second phase of insulin secretion. J. Biol. Chem. *282*, 9536–9546.

43. Usui, I., Imamura, T., Huang, J., Satoh, H., and Olefsky, J.M. (2003). Cdc42 is a Rho GTPase family member that can mediate insulin signaling to glucose transport in 3T3-L1 adipocytes. J. Biol. Chem. *278*, 13765–13774.

44. Ridley, A.J. (2006). Rho GTPases and actin dynamics in membrane protrusions and vesicle trafficking. Trends Cell Biol. *16*, 522–529.

45. Cho, Y.S., Chen, C.-H., Hu, C., Long, J., Ong, R.T.H., Sim, X., Takeuchi, F., Wu, Y., Go, M.J., Yamauchi, T., et al. (2012). Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in East Asians. Nat. Genet. *44*, 67–72.

46. Cho, Y.S., Lee, J.-Y., Park, K.S., and Nho, C.W. (2012). Genetics of type 2 diabetes in East Asian populations. Curr. Diab. Rep. *12*, 686–696.

47. Imamura, M., Maeda, S., Yamauchi, T., Hara, K., Yasuda, K., Morizono, T., Takahashi, A., Horikoshi, M., Nakamura, M., Fujita, H., et al. (2012). A single-nucleotide polymorphism in ANK1 is associated with susceptibility to type 2 diabetes in Japanese populations. Hum. Mol. Genet. *21*, 3042–3049.

48. Perry, J.R.B., Voight, B.F., Yengo, L., Amin, N., Dupuis, J., Ganser, M., Grallert, H., Navarro, P., Li, M., Qi, L., et al. (2012). Stratifying type 2 diabetes cases by BMI identifies genetic risk variants in LAMA1 and enrichment for risk variants in lean compared to obese cases. PLoS Genet. *8*, e1002741.

49. Morris, A.P., Voight, B.F., Teslovich, T.M., Ferreira, T., Segrè, A.V., Steinthorsdottir, V., Strawbridge, R.J., Khan, H., Grallert, H., Mahajan, A., et al. (2012). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat. Genet. *44*, 981–990.

50. Li, H., Gan, W., Lu, L., Dong, X., Han, X., Hu, C., Yang, Z., Sun, L., Bao, W., Li, P., et al. (2013). A genome-wide association study identifies GRK5 and RASGRP1 as type 2 diabetes loci in Chinese Hans. Diabetes *62*, 291–298.

51. Tabassum, R., Chauhan, G., Dwivedi, O.P., Mahajan, A., Jaiswal, A., Kaur, I., Bandesh, K., Singh, T., Mathai, B.J., Pandey, Y., et al. (2013). Genome-wide association study for type 2 diabetes in Indians identifies a new susceptibility locus at 2q21. Diabetes *62*, 977–986.

52. Saxena, R., Saleheen, D., Been, L.F., Garavito, M.L., Braun, T., Bjonnes, A., Young, R., Ho, W.K., Rasheed, A., Frossard, P., et al. (2013). Genome-wide association study identifies a novel locus contributing to type 2 diabetes susceptibility in Sikhs of Punjabi origin from India. Diabetes *62*, 1746–1755.

53. Huyghe, J.R., Jackson, A.U., Fogarty, M.P., Buchkovich, M.L., Stančáková, A., Stringham, H.M., Sim, X., Yang, L., Fuchsberger, C., Cederberg, H., et al. (2013). Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion. Nat. Genet. *45*, 197–201.

54. Parker, S.C.J., Stitzel, M.L., Taylor, D.L., Orozco, J.M., Erdos, M.R., Akiyama, J.A., van Bueren, K.L., Chines, P.S., Narisu, N., NISC Comparative Sequencing Program, et al. (2013). Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proc. Natl. Acad. Sci. U. S. A. *110*, 17921–17926.

55. Stancáková, A., Javorský, M., Kuulasmaa, T., Haffner, S.M., Kuusisto, J., and Laakso, M. (2009). Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men. Diabetes *58*, 1212–1221.

56. Delaneau, O., Zagury, J.-F., and Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. Nat. Methods *10*, 5–6.

57. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J., and Abecasis, G.R. (2012). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat. Genet. *44*, 955–959.

58. Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S.-Y., Freimer, N.B., Sabatti, C., and Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. Nat. Genet. *42*, 348–354.

59. Thorisson, G.A., Smith, A.V., Krishnan, L., and Stein, L.D. (2005). The International HapMap Project Web site. Genome Res. *15*, 1592–1593.

60. Fogarty, M.P., Xiao, R., Prokunina-Olsson, L., Scott, L.J., and Mohlke, K.L. (2010). Allelic expression imbalance at high-density lipoprotein cholesterol locus MMAB-MVK. Hum. Mol. Genet. *19*, 1921–1929.

61. 1000 Genomes Project Consortium, Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56–65.

62. Gaulton, K.J., Nammo, T., Pasquali, L., Simon, J.M., Giresi, P.G., Fogarty, M.P., Panhuis, T.M., Mieczkowski, P., Secchi, A., Bosco, D., et al. (2010). A map of open chromatin in human pancreatic islets. Nat. Genet. *42*, 255–259.

63. Stitzel, M.L., Sethupathy, P., Pearson, D.S., Chines, P.S., Song, L., Erdos, M.R., Welch, R., Parker, S.C.J., Boyle, A.P., Scott, L.J., et al. (2010). Global epigenomic analysis of primary human pancreatic islets provides insights into type 2 diabetes susceptibility loci. Cell Metab. *12*, 443–455.

64. Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M.A., Beaudet, A.L., Ecker, J.R., et al. (2010). The NIH Roadmap Epigenomics Mapping Consortium. Nat Biotech *28*, 1045–1048.

65. (2011). A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol. *9*, e1001046.

66. Miyazaki, J., Araki, K., Yamato, E., Ikegami, H., Asano, T., Shibasaki, Y., Oka, Y., and Yamamura, K. (1990). Establishment of a pancreatic beta cell line that retains glucose-inducible insulin secretion: special reference to expression of glucose transporter isoforms. Endocrinology *127*, 126–132.

67. Hohmeier, H.E., Mulder, H., Chen, G., Henkel-Rieger, R., Prentki, M., and Newgard, C.B. (2000). Isolation of INS-1-derived cell lines with robust ATP-sensitive K+ channel-dependent and -independent glucose-stimulated insulin secretion. Diabetes *49*, 424–430.

68. Portales-Casamar, E., Thongjuea, S., Kwon, A.T., Arenillas, D., Zhao, X., Valen, E., Yusuf, D., Lenhard, B., Wasserman, W.W., and Sandelin, A. (2010). JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. Nucleic Acids Res. *38*, D105–D110.

69. Matys, V., Fricke, E., Geffers, R., Gössling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A.E., Kel-Margoulis, O.V., et al. (2003). TRANSFAC: transcriptional regulation, from patterns to profiles. Nucleic Acids Res. *31*, 374–378.

70. Sandelin, A., Wasserman, W.W., and Lenhard, B. (2004). ConSite: web-based prediction of regulatory elements using cross-species comparison. Nucleic Acids Res. *32*, W249–W252.

71. Levy, S., and Hannenhalli, S. (2002). Identification of transcription factor binding sites in the human genome sequence. Mamm. Genome Off. J. Int. Mamm. Genome Soc. *13*, 510–514.

72. Ghosh, D. (2000). Object-oriented Transcription Factors Database (ooTFD). Nucleic Acids Res. *28*, 308.

73. Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R., and Willer, C.J. (2010). LocusZoom: regional visualization of genome-wide association scan results. Bioinforma. Oxf. Engl. *26*, 2336–2337.

74. Pastinen, T. (2010). Genome-wide allele-specific analysis: insights into regulatory variation. Nat. Rev. Genet. *11*, 533–538.

75. Pisania, A., Weir, G.C., O'Neil, J.J., Omer, A., Tchipashvili, V., Lei, J., Colton, C.K., and Bonner-Weir, S. (2010). Quantitative analysis of cell composition and purity of human pancreatic islet preparations. Lab. Investig. J. Tech. Methods Pathol. *90*, 1661–1675.

76. Ichii, H., Inverardi, L., Pileggi, A., Molano, R.D., Cabrera, O., Caicedo, A., Messinger, S., Kuroda, Y., Berggren, P.-O., and Ricordi, C. (2005). A Novel Method for the Assessment of Cellular Composition and Beta-Cell Viability in Human Islet Preparations. Am. J. Transplant. *5*, 1635–1645.

77. Dimas, A.S., Deutsch, S., Stranger, B.E., Montgomery, S.B., Borel, C., Attar-Cohen, H., Ingle, C., Beazley, C., Gutierrez Arcelus, M., Sekowska, M., et al. (2009). Common regulatory variation impacts gene expression in a cell type-dependent manner. Science *325*, 1246–1250.

78. Yang, T.-P., Beazley, C., Montgomery, S.B., Dimas, A.S., Gutierrez-Arcelus, M., Stranger, B.E., Deloukas, P., and Dermitzakis, E.T. (2010). Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. Bioinforma. Oxf. Engl. *26*, 2474–2476.

79. Smith, S.B., Ee, H.C., Conners, J.R., and German, M.S. (1999). Paired-homeodomain transcription factor PAX4 acts as a transcriptional repressor in early pancreatic development. Mol. Cell. Biol. *19*, 8272–8280.

80. Fujitani, Y., Kajimoto, Y., Yasuda, T., Matsuoka, T.A., Kaneto, H., Umayahara, Y., Fujita, N., Watada, H., Miyazaki, J.I., Yamasaki, Y., et al. (1999). Identification of a portable repression domain and an E1A-responsive activation domain in Pax4: a possible role of Pax4 as a transcriptional repressor in the pancreas. Mol. Cell. Biol. *19*, 8281–8291.

81. Ritz-Laser, B., Estreicher, A., Gauthier, B.R., Mamin, A., Edlund, H., and Philippe, J. (2002). The pancreatic beta-cell-specific transcription factor Pax-4 inhibits glucagon gene expression through Pax-6. Diabetologia *45*, 97–107.

82. Petersen, H.V., Jørgensen, M.C., Andersen, F.G., Jensen, J., F-Nielsen, T., Jørgensen, R., Madsen, O.D., and Serup, P. (2000). Pax4 represses pancreatic glucagon gene expression. Mol. Cell Biol. Res. Commun. MCBRC *3*, 249–254.

83. Duncan, M.K., Haynes, J.I., II, Cvekl, A., and Piatigorsky, J. (1998). Dual Roles for Pax-6: a Transcriptional Repressor of Lens Fiber Cell-Specific β-Crystallin Genes. Mol. Cell. Biol. *18*, 5579.

84. Shyr, C.-R., Tsai, M.-Y., Yeh, S., Kang, H.-Y., Chang, Y.-C., Wong, P.-L., Huang, C.-C., Huang, K.-E., and Chang, C. (2010). Tumor suppressor PAX6 functions as androgen receptor co-repressor to inhibit prostate cancer growth. The Prostate *70*, 190–199.

85. Liu, T., Zhao, Y., Tang, N., Feng, R., Yang, X., Lu, N., Wen, J., and Li, L. (2012). Pax6 directly down-regulates Pcsk1n expression thereby regulating PC1/3 dependent proinsulin processing. PloS One *7*, e46934.

86. Shimajiri, Y., Sanke, T., Furuta, H., Hanabusa, T., Nakagawa, T., Fujitani, Y., Kajimoto, Y., Takasu, N., and Nanjo, K. (2001). A missense mutation of Pax4 gene (R121W) is associated with type 2 diabetes in Japanese. Diabetes *50*, 2864–2869.

87. Yasuda, T., Kajimoto, Y., Fujitani, Y., Watada, H., Yamamoto, S., Watarai, T., Umayahara, Y., Matsuhisa, M., Gorogawa, S., Kuwayama, Y., et al. (2002). PAX6 mutation as a genetic factor common to aniridia and glucose intolerance. Diabetes *51*, 224–230.

88. Plengvidhya, N., Kooptiwut, S., Songtawee, N., Doi, A., Furuta, H., Nishi, M., Nanjo, K., Tantibhedhyangkul, W., Boonyasrisawat, W., Yenchitsomanus, P., et al. (2007). PAX4 mutations in Thais with maturity onset diabetes of the young. J. Clin. Endocrinol. Metab. *92*, 2821–2826.

89. Mauvais-Jarvis, F., Smith, S.B., Le May, C., Leal, S.M., Gautier, J.-F., Molokhia, M., Riveline, J.-P., Rajan, A.S., Kevorkian, J.-P., Zhang, S., et al. (2004). PAX4 gene variations predispose to ketosis-prone diabetes. Hum. Mol. Genet. *13*, 3151–3159.

90. Wen, J.H., Chen, Y.Y., Song, S.J., Ding, J., Gao, Y., Hu, Q.K., Feng, R.P., Liu, Y.Z., Ren, G.C., Zhang, C.Y., et al. (2008). Paired box 6 (PAX6) regulates glucose metabolism via proinsulin processing mediated by prohormone convertase 1/3 (PC1/3). Diabetologia *52*, 504–513.

91. Ma, R.C.W., Hu, C., Tam, C.H., Zhang, R., Kwan, P., Leung, T.F., Thomas, G.N., Go, M.J., Hara, K., Sim, X., et al. (2013). Genome-wide association study in a Chinese population identifies a susceptibility locus for type 2 diabetes at 7q32 near PAX4. Diabetologia *56*, 1291–1305.

92. Wethmar, K., Smink, J.J., and Leutz, A. (2010). Upstream open reading frames: molecular switches in (patho)physiology. BioEssays News Rev. Mol. Cell. Dev. Biol. *32*, 885–893.

93. Kulzer, J.R., Stitzel, M.L., Morken, M.A., Huyghe, J.R., Fuchsberger, C., Kuusisto, J., Laakso, M., Boehnke, M., Collins, F.S., and Mohlke, K.L. (2014). A common functional regulatory variant at a type 2 diabetes locus upregulates ARAP1 expression in the pancreatic beta cell. Am. J. Hum. Genet. *94*, 186–197.

94. Kowluru, A. (2010). Small G Proteins in Islet β-Cell Function. Endocr. Rev. *31*, 52–78.

95. Dion, L.D., Fang, J., and Garver, R.I. (1996). Supernatant rescue assay vs. polymerase chain reaction for detection of wild type adenovirus-contaminating recombinant adenovirus stocks. J. Virol. Methods *56*, 99–107.

96. Miyata, M., Rikitake, Y., Takahashi, M., Nagamatsu, Y., Yamauchi, Y., Ogita, H., Hirata, K., and Takai, Y. (2009). Regulation by afadin of cyclical activation and inactivation of Rap1, Rac1, and RhoA small G proteins at leading edges of moving NIH3T3 cells. J. Biol. Chem. *284*, 24595–24609.

97. Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., et al. (2012). GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. *22*, 1760–1774.

98. McCarthy, M.I., Rorsman, P., and Gloyn, A.L. (2013). TCF7L2 and diabetes: a tale of two tissues, and of two species. Cell Metab. *17*, 157–159.

99. Mykkänen, L., Haffner, S.M., Kuusisto, J., Pyörälä, K., Hales, C.N., and Laakso, M. (1995). Serum proinsulin levels are disproportionately increased in elderly prediabetic subjects. Diabetologia *38*, 1176–1182.

100. Vitale, N., Caumont, A.-S., Chasserot-Golaz, S., Du, G., Wu, S., Sciorra, V.A., Morris, A.J., Frohman, M.A., and Bader, M.-F. (2001). Phospholipase D1: a key factor for the exocytotic machinery in neuroendocrine cells. EMBO J. *20*, 2424–2434.

101. Vitale, N., Chasserot-Golaz, S., Bailly, Y., Morinaga, N., Frohman, M.A., and Bader, M.-F. (2002). Calcium-regulated exocytosis of dense-core vesicles requires the activation of ADP-ribosylation factor (ARF)6 by ARF nucleotide binding site opener at the plasma membrane. J. Cell Biol. *159*, 79–89.

102. Béglé, A., Tryoen-Tóth, P., de Barry, J., Bader, M.-F., and Vitale, N. (2009). ARF6 regulates the synthesis of fusogenic lipids for calcium-regulated exocytosis in neuroendocrine cells. J. Biol. Chem. *284*, 4836–4845.

103. Zeniou-Meyer, M., Zabari, N., Ashery, U., Chasserot-Golaz, S., Haeberlé, A.-M., Demais, V., Bailly, Y., Gottfried, I., Nakanishi, H., Neiman, A.M., et al. (2007). Phospholipase D1 production of phosphatidic acid at the plasma membrane promotes exocytosis of large dense-core granules at a late stage. J. Biol. Chem. *282*, 21746–21757.

104. Hughes, W.E., Elgundi, Z., Huang, P., Frohman, M.A., and Biden, T.J. (2004). Phospholipase D1 Regulates Secretagogue-stimulated Insulin Release in Pancreatic β-Cells. J. Biol. Chem. *279*, 27534–27541.

105. Morris, A.J., Frohman, M.A., and Engebrecht, J. (1997). Measurement of phospholipase D activity. Anal. Biochem. *252*, 1–9.

106. Caumont, A.S., Galas, M.C., Vitale, N., Aunis, D., and Bader, M.F. (1998). Regulated exocytosis in chromaffin cells. Translocation of ARF6 stimulates a plasma membrane-associated phospholipase D. J. Biol. Chem. *273*, 1373–1379.

107. Halban, P.A. (1990). Proinsulin trafficking and processing in the pancreatic B cell. Trends Endocrinol. Metab. *1*, 261–265.