Molecular Screening of Caveolin 1 (*CAV1*) Gene Polymorphisms in Patients with Type II Diabetes Mellitus

SUBMITTED BY

Aaraz Ahmad (00000356711)

Ayesha Sana (00000357763)

Iqra Karamat (00000357512)

Shanza Ayub (00000357973)

Supervisor: Prof. Dr. Attya Bhatti

Co-Supervisor: Prof. Dr. Peter John

Atta-Ur-Rahman School of Applied Biosciences, National University of Sciences and Technology, Islamabad

2024

Molecular Screening of Caveolin 1 (*CAV1*) Gene Polymorphisms in Patients with Type II Diabetes Mellitus

SUBMITTED BY

Aaraz Ahmad (00000356711)

Ayesha Sana (00000357763)

Iqra Karamat (00000357512)

Shanza Ayub (00000357973)

A dissertation submitted by the partial fulfillment of requirement

for the degree of Bachelors of Applied Biosciences from

Atta-Ur-Rahman School of Applied Biosciences

Supervisor: Prof. Dr. Attya Bhatti

Co-Supervisor: Prof. Dr. Peter John

Atta-Ur-Rahman School of Applied Biosciences, National University of Sciences and Technology, Islamabad

2024

THESIS ACCEPTANCE CERTIFICATE

This document certifies that this is the final copy of BS FYP Thesis written by **Ms. Aaraz Ahmad** (Reg No. 356711), **Ms. Ayesha Sana** (Reg No. 357763), **Ms. Iqra Karamat** (Reg No. 357512), and **Ms. Shanza Ayub** (357973) of **Atta-Ur-Rahman School of Applied Biosciences (ASAB)**. The undersigned has thoroughly reviewed and found that the document complies with all regulations set by the National University of Sciences and Technology (NUST). It is confirmed to be free of plagiarism, errors, and mistakes, and is accepted as a partial fulfillment for the Bachelor of Sciences degree in Applied Biosciences. Additionally, all necessary amendments suggested during the final presentation have been included in the thesis.

Signature: _____

Name of Supervisor: **Prof. Dr. Attya Bhatti**

Date: 24-06-2024

Signature (HOD): _____

Head of Department Biomedicine: **Dr. Hussain Mustatab Wahedi**

Date: 24-06-2024

Signature (Dean/Principal): _____

Principal ASAB: **Dr. Muhammad Asghar**

Date: 24-06-2024

DECLARATION

We affirm that the research work titled **" Molecular Screening of Caveolin 1 (CAV1) Gene Polymorphisms in Patients with Type II Diabetes Mellitus"** authored by **Aaraz Ahmad, Ayesha Sana, Iqra Karamat** and **Shanza Ayub**, is our original creation. This work has not been presented elsewhere for evaluation. The research was conducted during our undergraduate studies at Atta-Ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), under the guidance of **Prof. Dr. Attya Bhatti** and **Prof. Dr. Peter John**. Any material borrowed from external sources has been appropriately cited and acknowledged.

Signature of Student _____

Aaraz Ahmad

Reg No. 356711

Signature of Student _____

Ayesha Sana

Reg No. 357763

Signature of Student _____

Iqra Karamat

Reg No. 357512

Signature of Student _____

Shanza Ayub

Reg No. 357973

CERTIFICATE FOR PLAGIARISM

It is certified that this Undergraduate Thesis Titled "**Molecular Screening of Caveolin 1 (CAV1) Gene Polymorphisms in Patients with Type II Diabetes Mellitus**" written by by **Ms. Aaraz Ahmad** (Reg No. 356711), **Ms. Ayesha Sana** (Reg No. 357763), **Ms. Iqra Karamat** (Reg No. 357512), and **Ms. Shanza Ayub** (357973) of **Atta-Ur-Rahman School of Applied Biosciences (ASAB)**, has been examined by me.

I undertake that:

a. The thesis contains original contributions and novel insights not found in previously published works or those currently being reviewed for publication elsewhere. No exact sentences, equations, diagrams, tables, paragraphs, or sections have been directly copied from prior sources without proper quotation marks and citations.

b. The work presented is original and own work of the authors i.e., there is no plagiarism.

c. There is no falsification of data or misrepresentation of research findings in the records. The thesis has undergone inspection using TURNITIN (an originality report is attached), confirming compliance with the HEC plagiarism policy and relevant guidelines.

Signature of Supervisor

Prof. Dr. Attya Bhatti

Department of Biomedicine

Atta-Ur-Rahman School of Applied Biosciences (ASAB)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

ACKNOWLEDGEMENT

We would like to express our deepest gratitude to our supervisor, **Prof. Dr. Attya Bhatti** and co-supervisor, **Prof. Dr. Peter John**, for their unwavering support, insightful guidance, and invaluable expertise throughout the duration of this research. Their encouragement, patience, and constructive feedback have been instrumental in shaping the direction and success of this thesis. We are sincerely thankful for their willingness to invest their time and effort in our academic growth, and for providing us with the freedom to explore our ideas while ensuring that we remain focused and on track.

We are also profoundly grateful to the lab technician, Mam Fouzia, whose technical assistance and meticulous attention to detail have been indispensable. Their dedication to maintaining a productive and efficient lab environment has allowed us to conduct our experiments with confidence. Their problem-solving skills and readiness to help with any technical issues that arose have been a tremendous support, and we deeply appreciate their commitment to excellence.

Additionally, we would like to extend our heartfelt thanks to our seniors, Ms. Ilhaam Durrani and Ms. Kashaf Rasool, for their mentorship, camaraderie, and invaluable advice. Their willingness to share their experiences and knowledge has been a source of inspiration and guidance throughout our academic journey. We are especially thankful for their encouragement during challenging times and for their constructive criticism, which has helped us refine our work.

Moreover, we would like to acknowledge the support and camaraderie of our fellow colleagues. Their friendship, collaboration, and intellectual exchange have enriched our research experience. Special thanks to Ayesha Afzal and Harooma Jamshed for their collaborative spirit and for always being there to lend a hand or offer a listening ear. We are also indebted to the administrative staff, who have been incredibly helpful in managing the logistical aspects of our research. Their efficiency and support have ensured that we could focus on our studies without unnecessary distractions.

Last but not least, we would like to thank our families and friends for their unwavering support and encouragement throughout this journey. Their understanding and patience, especially during the more demanding phases of our research, have been a pillar of strength for us. Thank you all for your contributions and support. This project would not have been possible without each and every one of them.

Table of Contents

List of Figures

List of Tables

List of Acronyms

ADA	American Diabetes Association
ADP	Adenosine diPhosphate
AKT	Ak strain transforming
ANT2	Adenine Nucleotide Translocase 2
ATP	Adenosine triPhosphate
BC	Breast Cancer
BMI	Body Mass Index
CD36	Cluster Of Differentiation
CDC	Centers for Disease Control and Prevention
CDC42	Cell Division Control Protein
CKI	Casein Kinase I
CAD	Coronary Artery Disease
CHD	Coronary Heart Disease
CVD	Cardiovascular Disease
DM	Diabetes Mellitus
DPP	Diabetes Prevention Programme
EDTA	Ethylenediamine Tetra acetic Acid
EGP	Endogenous Glucose Production
ERK	Extracellular Signal-Regulated Kinase
EASD	European Association for Study of Diabetes
FFA	Free Fatty Acids
FOXO	Forkhead Box Transcription Factors
FPG	Fasting Plasma Glucose
GDP	Guanosine diphosphate
GIP	Gastric Inhibitory Polypeptide
GLP1	Glucagon-like peptide 1
GLUT	Glucose Transporter
GM3	Monosialodihexosylganglioside

GTP	Guanosine triphosphate
HCC	Hepatocellular Carcinoma
HDL	High Density Lipoprotein
HF	Heart Failure
HGP	Hepatic Glucose Production
IEC	International Expert Committee
IGT	Insulin Glycated Retinopathy
IL	Interleukin
IR	Insulin Resistance
IRS	Insulin Receptor Substrate
MS	Metabolic Syndrome
MβCD	Methyl-β-cyclodextrin
NPH	Neutral Protamine Hagedorn
OGTT	Oral Glucose Tolerance Test
PKC	Protein Kinase C
PMDI	Precision Medicine in Diabetes Initiative
PTM	Post-Translational Modifications
RMSD	Root-Mean Square Deviation
ROS	Reactive Oxygen Sepcies
SAH	Systemic Arterial Hypertension
SNAP 25	Synaptosomal-Associated Protein 25
SNP	Single Nucleotide Polymorphism
SVM	Support Vector Machines
T2DM	Type II Diabetes Mellitus
TAE	Tris Acetate EDTA
TLR	Toll-like Receptor
TNF	Tumor Necrosis Factor

ABSTRACT

Type II Diabetes Mellitus (T2DM) is one of the most frequently occurring heterogenous, multi-system disorders characterized by hyperglycemia, responsible for 90% of the 463 million diabetes cases around the world. Complex etiology, painful clinical manifestations, and associated comorbidities makes T2DM a disease of great concern. T2DM has a strong genetic basis and studies have elucidated the association of SNPs of different genes with its progression. SNPs form the basis for precision medicine as the analysis of SNPs can aid the tailoring of medical treatments to individual genetic profiles, enhancing the efficacy and safety of therapies. This personalized approach helps in predicting disease risk, optimizing drug dosages, and minimizing adverse drug reactions, ultimately leading to more precise and effective healthcare. The current study aims to investigate the association of rs369262127 and rs201966419 with the risk of T2DM using Gradient PCR. In-silico analysis was performed to identify pathogenic exonic missense SNPs of *CAV1* gene. In order to identify damaging nsSNPs in the *CAV1* gene, computational tools such as SIFT, SNPs&GO, SNAP2, and MutPred2 were used. MUPro and ConSurf predicted protein stability and conservation profiles, respectively. Phyre2 and I-TASSER designed the 3D structure of *CAV1* mutants. Post-translational modifications were determined by GPS-MSP, NetPhos3.1, NetOglyc4.0, and NetNGlyc 1.0. GENEMania and STRING predicted the protein-protein interaction. We found six nsSNPs in the coding region to be highly damaging: rs369884333, rs372416448, rs200052661, rs369262127, rs151024969, and rs201966419. Most alterations in structure were caused by rs201966419. As a result, rs369262127 and rs201966419 were selected for invitro analysis as they were found to be clinically significant in ClinVar. Invitro analysis including gradient PCR and gel electrophoresis revealed no significant difference between the genotype and allele frequency of patients and healthy controls. So, it can be concluded that the risk alleles of the two SNPs were not associated with susceptibility to T2DM in our sample pool, but replication of this study with larger sample size is required for further validation of results.

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

1.1.1 Type II Diabetes Mellitus

Type II Diabetes Mellitus (T2DM) is one of the most frequently occurring heterogenous, multi-system disorders 47 characterized by hyperglycemia. It occurs by two main reason, one is when the pancreas produce less insulin and the other is the tissues carrying insulin receptors become unresponsive to high insulin levels in the blood leading to insulin resistance 47 (see

Figure *1*). Insulin hormone is required by the body for the uptake of glucose into cells, thereby maintaining normal blood glucose levels and ensuring that the cells get energy 47. In former case, the beta cell of pancreas produces insulin in response to the elevated level of glucose in the blood hence maintain glucose level. But in later case occurs when insulin binds to the insulin receptor and causes the activation of downstream signaling PI3K/AKT pathway leading to the translocation of GLUT4 to the plasma membrane for the uptake of glucose, hence maintaining the glucose level in the blood. But in case of diabetes, the downstream PI3K/AKT signaling pathway is unable to be activated and GLUT4 is unable to be translocated to the membrane, leading to the elevated glucose level in the blood termed hyperglycemia.

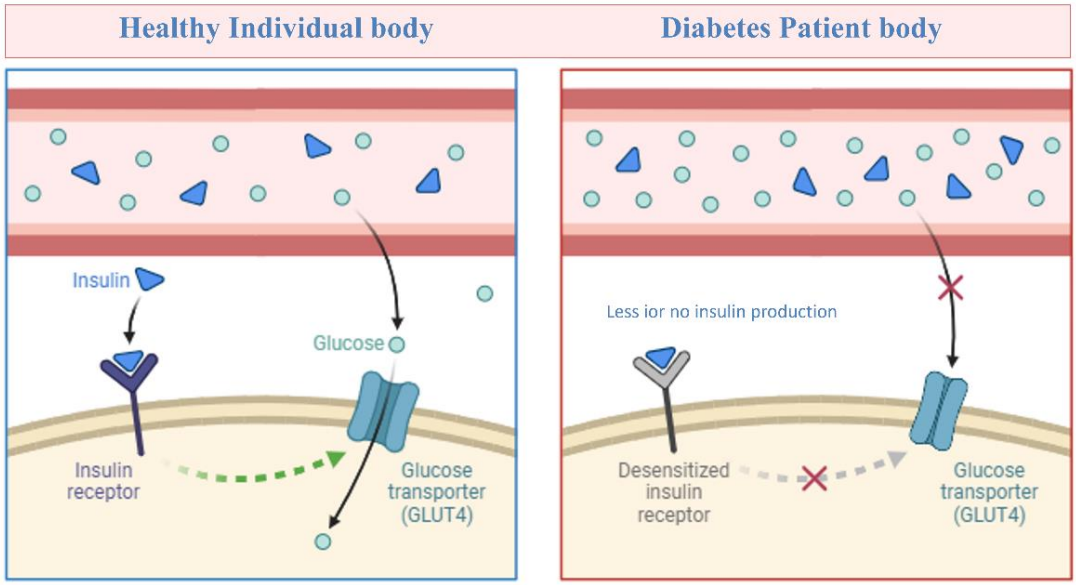


Figure 1: The figure shows the body of a healthy individual and a type II diabetic patient. A healthy body displays normal insulin-glucose metabolism, whereas a type II diabetic has desensitized insulin receptors leading to hyperglycemia. (Created with BioRender.com).

1.2 EPIDEMIOLOGY

Diabetes is most commonly found in adults over the age of 45 47. T2DM is responsible for 90% of the 463 million diabetes cases around the world 47 with 10.5% of the world adult population (20-79 years) affected by it 47. According to the International Diabetes Federation, 3 in 4 adults in low- and middle-income countries are diabetic, making Pakistan highly vulnerable 47. In 2021 alone, 32,964,000 people in Pakistan were diabetic and 396,625,000 lost their lives from it. The number of diabetes patients are expected to nearly double (62,018,000) by 2045 47.

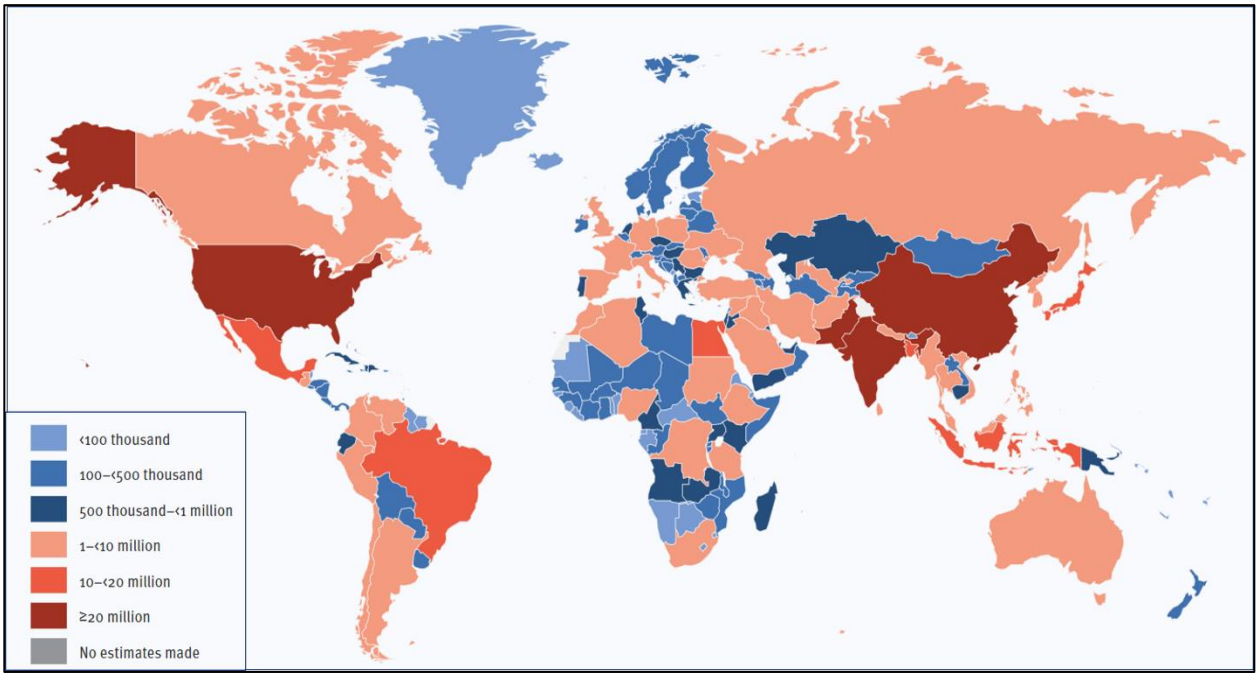


Figure 2: The global epidemiology of diabetes in 2021 according to the IDF. Pakistan had >20 million diabetes cases in 202147.

1.3 CLINICAL MANIFESTATIONS

The symptoms include frequent urination, thirst, hunger, weight loss, tiredness and vision impairment 47.

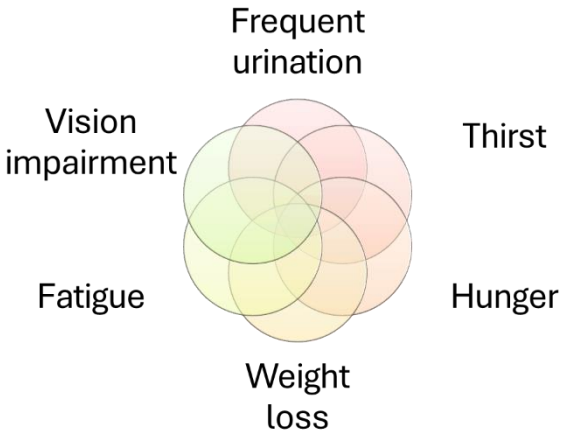


Figure 3: Symptoms of type II diabetes.

1.4 COMPLICATIONS OF T2DM

T2DM affects multiple organ systems. According to WHO, diabetes can cause harm to the heart, eyes, kidneys, nerves, vasculature 47 gums, feet and bladder 47. According to CDC, diabetic individuals are twice as likely to develop CVDs than normal individuals 47. Hyperglycemia damages blood vessels in the eyes, kidneys and nerves leading to diabetic retinopathy 47, chronic kidney disease 47 and neuropathies 47 respectively. **Figure 4** shows the complications associated with T2DM.

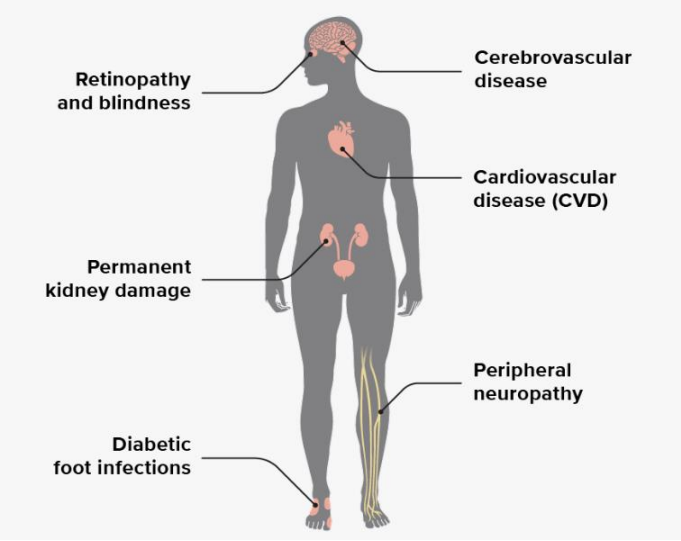


Figure 4: The complications of type II diabetes 47.

1.5 RISK FACTORS OF TYPE 2 DIABETES MELLITUS

The failure of β cells is an important hallmark of T2DM. A combination of genetic, metabolic and environmental factors leads to the development of diabetes. Some risk factors are non-modifiable, but working on the modifiable factors can delay the onset of diabetes.

1.6.1 Non-Modifiable Risk Factors

1.5.1.1 Genetic Predisposition

It is dependent upon the race and ethnicity of an individual. Black populations are the most prone to diabetes 47. Japanese, Hispanics and Native Americans also have a high chance of developing diabetes whereas White American population has a low likelihood in comparison 47. A research study showed that Finnish families displayed inheritance for defective insulin secretion 47. T2DM is a polygenic disease that involves multiple loci, some impairing insulin secretion while others weakening insulin action 47.

The genes *ABCC8* (sulphonyl urea receptor), *KCNJ11* (KIR6.2), *SLC2A2* (GLUT2), *HNF4A* (HNF4α), and *INS* affect the activity of pancreatic β-cells, which in turn causes insulin resistance. Three more genes, *INSR*, *PIK3R1*, and *SOS1*, are associated with insulin. Other genes responsible for T2DM include *SLC27A5*, *FABP6*, *Thr79Met*, *DBI*, *PTGES2*, and *CLPS*, *RAPGEF1*, HNF1 homeobox A (*HNF1A*) and HNF1 homeobox B (*HNF1B*) 48. The interaction of the environment with the susceptible loci could explain the heritable nature of the disease, however, much is still not understood regarding T2DM loci.

1.5.1.2 Age

Insulin levels decline with age due to a decrease in β cell function 48.

1.6.2 Modifiable Risk Factors

1.5.2.1 Obesity

Increase in BMI results in earlier diagnosis of T2DM 48. this is because obesity leads to some metabolic abnormalities including both inter-organ connections and cell autonomous processes, resulting in IR 47.

1.5.2.2 Low physical activity

Three aspects of an active lifestyle reduce the incidence of diabetes: decreased intra-abdominal fat, enhanced glucose uptake, and improved blood flow and glucose uptake caused by skeletal muscle contraction 48. Weight loss has proven to improve insulin sensitivity and decrease chances of IR 48.

1.5.2.3 Diet

The risk of type 2 diabetes is increased by Western diets, but it is decreased by traditional foods like fish and marine mammals. The use of vegetarian, vegan, Mediterranean, and dietary approaches can lower blood pressure, enhance insulin sensitivity, and reduce glycated hemoglobin 48.

1.5.2.4 Persistent Organic Pollutants (POPs)

Adipose tissues accumulate POPs and enter the human body through consumption of meat, fish and dairy. This causes obesity, dyslipidemia and insulin resistance 48.

1.6 PATHOPHYSIOLOGY

1.7.1 β Cell Dysfunction

Insulin is stored in granules of β cells until secretion. β cells secrete insulin during high glucose levels, but other factors can be responsible for it. During high glucose levels, glucose enters β cells through GLUT2 and increases ATP/ADP ratio. ATP-dependent potassium channels close and voltage-gated calcium channels open and cause insulin release. This is a highly regulated process and any damage to islets can

interfere with this process. β cell dysfunction can occur from high nutritional states (like obesity) causing hyperglycemia and hyperlipidemia, toxic pressures (inflammatory/oxidative stress) leading to loss of islets. Damage to the islets causes insulin insensitivity and T2DM 47.

1.7.2 Pathological Conditions Propagating T2DM

1.6.2.1 Nutritional Factors: Diet containing **elevated levels** of fats and carbohydrates cause the production of ROS species and lead to inflammatory stress 47. Glucotoxicity and lipotoxicity damage β cells and decrease insulin secretion (48; 48). Both mechanisms **participate in** the pathogenesis of diabetes.

1.6.2.2 Physical Activity: Sedentary lifestyle causes obesity which leads to a constant state of inflammation. Weight loss is an effective way to reduce inflammation and increase insulin sensitivity 47.

1.6.2.3 Gut Dysbiosis: Gut microbiota participates in many biological processes. Changes in diet and lifestyle can disturb gut microbiota and lead to pathogenesis of T2DM 48. In mice, high-fat diets can cause the synthesis of lipopolysaccharides, which can cause insulin resistance and inflammation 48. Intestinal dysbiosis can result in reduced fatty acid synthesis, disturbed glucose homeostasis, and the development of T2DM 48.

1.6.2.4 Metabolic Memory: The ongoing presence of diabetes problems even after glycemic management is known as metabolic memory. Four pathways participate in metabolic memory: oxidative stress, non-enzymatic protein glycation, chronic inflammation, and epigenetics. Imbalance in any one of these areas can lead to diabetes pathogenesis 47.

1.6.2.5 Mitochondrial Dysfunction: Mitochondrial malfunction results in higher O_2 production, a decrease in nutrition oxidation efficiency 48, and an accumulation of ROS within the mitochondria 48. One of the main processes regulating the quantity and quality of mitochondria is mitophagy 48. Reduced glucose homeostasis and hepatic insulin sensitivity are caused by impaired mitophagy 49.

1.7 INSULIN RESISTANCE

The pathophysiological abnormality that commonly initiates T2DM is Insulin Resistance (IR) in the tissue, muscle and liver. This eventually leads to the decline in the function of pancreatic β cell. IR is most commonly an inherited factor, though its inheritance is unclear 49. Insulin resistance is the insensitivity of insulin-responding tissues to high insulin levels 47. Muscle, liver and adipose tissue are the central organs in glucose metabolism and provide insight into the mechanism of insulin resistance.

1.8.1 Selective Insulin Resistance:

Some tissues become more sensitive to insulin during IR. This is known as selective insulinresistance. E.g., Insulin is required by the liver for lipid synthesis. During insulin resistance, hepatic lipogenesis increases, suggesting that hepatic lipid metabolism is independent of the hepatic glucose metabolism pathway 49.

1.8.2 Liver Insulin Resistance:

Liver is responsible for endogenous glucose production (EGP) during fasting 49. In patients with T2DM, the hepatic glucose production (HGP) rises, even under the influence of high plasma insulin. This shows resistance to insulin by hepatocytes 49. Insulin insensitivity leads to a decreased synthesis of glycogen in the liver and impaired glucose metabolism 47.

1.8.3 Skeletal Muscle Insulin Resistance:

Insulin causes the transport of glucose in skeletal muscles by GLUT4 translocation to the membrane. Research has shown that knocking out of the GLUT4 gene in mice muscle cells decreases GLUT4 translocation to the cell surface, reducing glucose uptake in the and consequently diabetes-like symptoms 49. Mutations in glycogen synthase, hexokinase, and GLUT4 can also cause hyperglycemia 47. Muscle IR accounts for 85-90% of the impairment in body glucose disposal 48.

1.8.4 Adipose Insulin Resistance:

Adipose tissue plays a significant role in glucose and lipid synthesis and body weight homeostasis. Insulin causes uptake of glucose, fat and glycerol into adipocytes. IR disables this system and causes free fatty acid (FFA) release into bloodstream. This stimulates adenine nucleotide translocase 2 (ANT2) and adipose tissue dysfunction 47.

1.8 DISEASE PROGRESSION:

Insulin resistance worsening and β-cell function decline are the main hallmarks of type II diabetes progression. Clinically, there is a decline in multiple markers, including postprandial glucose levels, fasting plasma glucose (FPG), and A1C 49.

1.9.1 Pre-Diabetes

The phrase "pre-diabetes" has gained widespread usage and denotes a chance of developing into overt diabetes. However, according to data from the Diabetes Prevention Programme (DPP), those with insulin-glycated retinopathy (IGT) are more likely than people with normal A1C levels to develop diabetic retinopathy. These people have a 10% incidence of diabetic retinopathy, 80% loss of β-cell function, and maximal or near-maximum insulin resistance. It is appropriate to classify these pre-diabetic people as having type 2 diabetes 49.

1.9.2 Pre-Diabetes to Overt Diabetes:

A lower HDL cholesterol level, higher BMI, blood pressure, and triglycerides and a steeper rate of increase in fasting glucose indicate development of diabetes 49.

1.9.3 Acute Insulin Response

Following an increase in blood glucose, the insulin levels rise rapidly known as acute insulin response. In type II diabetes, this response progressively declines, indicating damage to β cells 49.

1.9.4 Progression to Medication

The chance of starting treatment and the risk of advancement declines with age. While elderly patients with stable weight may be monitored without pharmaceutical therapy, younger individuals should be administered medication earlier 49.

1.9.5 Loss of Glycemic Control on Medication

Glycemic control in type 2 diabetes is declining over time. The U.K. Prospective Diabetes Study (UKPDS) showed that when compared to standard care, treatment with metformin, sulfonylurea, or insulin dramatically lowers A1C and FPG (49; 49).

1.9.6 Decline in β Cell Function.

β cell function declines progressively with hyperglycemia despite treatment. This failure of β cells amplifies in old age when other complications aggravate diabetes (49).

1.9 SCREENING AND DIAGNOSIS

The test for screening and diagnosis of T2DM is common and follows the guidelines set either by ADA in 1997 or WHO National diabetic group criteria of 2006 50. These guidelines state that for a person to have diabetes, they must show one raised blood glucose reading with diabetes symptoms (e.g. polyuria, polydipsia, polyphagia and weight loss) or raised values on two instances of either fasting plasma glucose (FPG) (126 mg/dL) or oral glucose tolerance test (OGTT) 2 hours after taking oral glucose drink (200 mg/dL) 50. While the WHO emphasizes the OGTT, the 1997 ADA criteria for DM diagnosis place more emphasis on the FPG.

The International Expert Committee (IEC) suggested in July 2009 that DM patients should have an extra diagnostic criterion of a HbA1c value 6.5%. IEC said that instead of the term 'pre-diabetes,' HbA1c values of 6% to 6.5% be referred to as having high risk of developing diabetes 50.

1.10 PREVENTION

The Diabetes Prevention Program (DPP) led by National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) is a pioneer in cost-effective approaches to prevent diabetes. Their research shows that moderate weight loss or the drug metformin significantly delay or prevent diabetes 50. DPP focuses on lifestyle changes using real-life well maintained diabetic fitness trainers as educators 50. This study is in use by CDC in their National Diabetes Prevention Program 50.

1.11 MANAGEMENT

Modern research has allowed us to individualize diabetes treatment and management according to complications, comorbidities and the duration of diabetes. For this purpose, education and awareness on diabetes self-management is essential to preserve β cell function 50.

1.12 TREATMENT

In 1921, Banting, Best and colleagues in Toronto, Canada discovered insulin, the first effective treatment to diabetes. Insulin went through numerous changes to become NPH insulin in 1946 that we still use today 50.

The first pharmacological drug to treat diabetes (sulfonylurea (SU) carbutamide) in 1955 was closely followed by others still being used today 50.

- Sulfonylureas: Stimulates insulin secretion 50.
- Biguanides: Reduces glucose absorption, improves insulin sensitivity, glucose uptake and fatty acid oxidation 50.
- Meglitinides: Stimulates insulin release from pancreatic β cells 50.
- Alpha-Glucosidase Inhibitors: Treat postprandial hyperglycemia 50.
- Thiazolidinediones: Insulin sensitizer 50.
- Incretin-Based Therapies: Long-term gains in body weight management and glycemic control 50.
- Insulin Analogs: Mimics normal physiological insulin secretion 50.

- Dipeptidyl-Peptidase IV Inhibitors: Inhibits DPP-4, a ubiquitous enzyme that inactivates both GLP-1 and GIP which improves glycemic control 50.
- Bromocriptine: Reduces HbA1c levels after 24 weeks of treatment 51

1.13 CURRENT TRENDS

Modern technologies in the field of medical biology have allowed advancements in research of diabetes. The American Diabetes Association (ADA) collaborated with the European Association for the Study of Diabetes (EASD) on the Precision Medicine in Diabetes Initiative (PMDI). The aim of this initiative is to provide tailored care to diabetes patients by assessing the unique need of each patient and customizing treatment to fit their needs 51.

1.14 GAP OF STUDY

Type II diabetes is a complex metabolic disorder that demands a thorough study of its underlying genetic factors. One of these is mutations in *CAV1*. However, a crucial knowledge gap surrounds *CAV1* gene variants and their role in IR and T2DM, especially in the Pakistani population. Further research will understand their specific impact, potentially providing valuable insights into insulin resistance mechanisms and aiding in the development of tailored treatment approaches.

1.15 OBJECTIVES OF STUDY

The main objectives of the study are as follow:

1. To check the pathogenicity of the SNPs of *CAV1* through in-silico analysis.
2. To investigate the association of selected SNPs with T2DM through in-vitro analysis.

1.16 IMPACT OF STUDY

Research has shown that diabetes may begin as early as in vitro, progress slowly over a lifetime and develop into a chronic illness that may lead to an assortment of conditions such as Alzheimer disease, cancer, liver failure, depression, and hearing loss 50. These comorbidities combined with the millions of dollars spent every year can be avoided by in-depth understanding of the metabolic pathways involved in diabetes and research on new and improved methods to understand the special needs of diabetic patients.

CHAPTER 2

LITERATURE REVIEW

2.1 DIABETES TYPE 2

Type II diabetes mellitus ranks among the most prevalent metabolic disorders with 90% of diabetes cases. It is characterized by insufficient insulin secretion from pancreatic islet β-cells, tissue resistance to insulin, and an insufficient compensatory insulin secretion in response 51. Targeted therapy is the only efficient way to minimize symptoms and avoid the onset of comorbidities associated with T2DM.

2.2 CAVEOLAE

Caveolae, tiny invaginations of the cell membrane, are integral to essential physiological processes like cell signaling, endocytosis, and the movement of cholesterol within cells 51. Caveolin-1, -2, and -3 are unique membrane proteins known to be the major components of caveolae 51. In mammals, three caveolin genes are identified as *CAV1*, *-2*, and *-3,* respectively. These genes encode five distinct isoforms of the caveolin protein 51.

2.3 CAVEOLIN-1

Caveolin-1, encoded by the gene *CAV1* 51 is the principal component of flask-shaped caveolae in membranes 51. It is the primary isoform of the caveolin family outside of muscle tissues, abundant in both caveolae and vesicles targeted to the apical surface of polarized epithelial cells. Also known as VIP21, it was first recognized as a protein undergoing tyrosine phosphorylation in cells transformed by *Rous sarcoma* virus. Caveolae were initially observed between 1953 and 1955 as plasmalemmal vesicles engaged in endocytosis, facilitating the movement of molecules across endothelial cells 51.

2.4 LOCATION OF *CAV1*

CAV1 is located on chromosome 07: 116,524,994-116,561,179 forward strand as shown in **Figure 5** (https://asia.ensembl.org/; Accessed May 5th 2024).

Figure 5: Location of *CAV1* on chromosome 7.

2.5 STRUCTURE

The Caveolin-1 (*CAV1*) monomer weighs approximately 22 kDa and combines with 14-16 monomers in the endoplasmic reticulum to form a 350 kDa oligomer. The oligomers interact with cholesterol molecules in the golgi body and transported to the cellular membrane where additional cytosolic adaptor proteins named Cavins contribute to caveolar formation. *CAV1* mediates the recruitment of Cavin proteins (CAVIN1/2/3/4) to the caveolae 51. **Figure 6** displays the 3D protein model of Caveolin 1 (https://www.rcsb.org/; Accessed on May 16th, 2024).

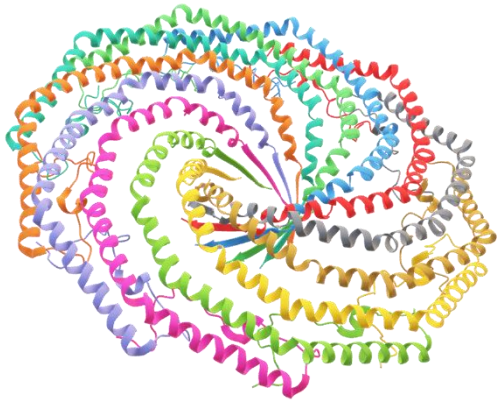


Figure 6: Caveolin 1 model. The figure displays 3D Caveolin 1 model obtained from RCSB PDB.

Caveolin proteins exhibit three clearly defined regions 51.

1. A cytoplasmic N-terminal domain
2. A unique 33-amino acid hydrophobic membrane-spanning segment and,
3. A cytoplasmic C-terminal domain

The protein adopts a hairpin-loop structure within the lipid bilayer, allowing both the N and C termini to be exposed on the cytoplasmic surface. Scaffolding domain engages with numerous signaling proteins 51. **Figure 7** shows the structure of caveolae and the caveolin 1 protein with its domains.

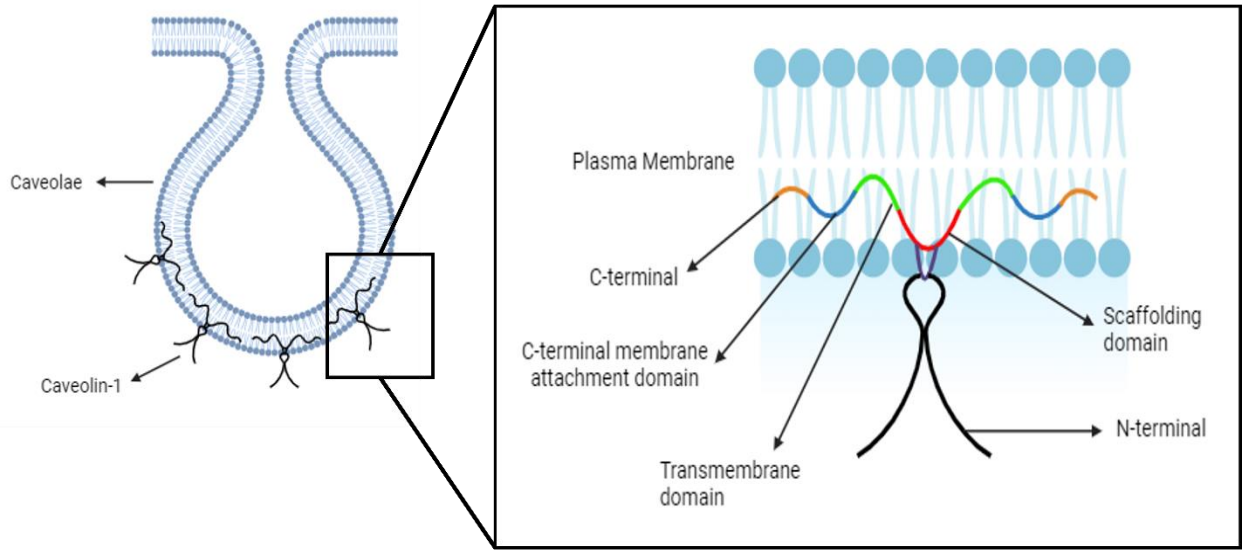


Figure 7: Structure of different domains of caveolin 1 protein in caveolae. Caveolin 1 has four main domains namely, the N-terminal, C-terminal, transmembrane domain and the scaffolding domain.

2.6 TRANSCRIPTS OF *CAV1*

CAV1 consists of 9 transcripts. Among them 6 are protein coding, 1 retained intron and 1 is nonsense mediated decay (https://asia.ensembl.org/; Accessed May 5th 2024) as shown in **Table 1**.

Table 1: Transcripts of *CAV1* with their base pair and biotype as obtained from Ensembl.

Transcript ID	Name	Base pairs	Protein	Translation ID	Biotype	Uni Prot Match
ENST00000341049.7	*CAV1*-201	2456	178aa	ENSP00000339191.2	Protein coding	Q03135-1
ENST00000405348.6	*CAV1*-205	2652	147aa	ENSP00000384348.1	Protein coding	Q03135-2
ENST00000393467.1	*CAV1*-202	2628	147aa	ENSP00000377110.1	Protein coding	Q03135-2
ENST00000614113.5	*CAV1*-209	1130	115aa	ENSP00000479447.2	Protein coding	A0A7P0YWJ6
ENST00000393468.1	*CAV1*-203	845	147aa	ENSP00000377111.1	Protein coding	Q03135-2
ENST00000393470.1	*CAV1*-204	598	167aa	ENSP00000377113.1	Protein coding	E9PCT5
ENST00000456473.5	*CAV1*-207	557	138aa	ENSP00000389033.1	Protein coding	C9JKI3
ENST00000451122.5	*CAV1*-206	1653	86aa	ENSP00000409541.1	Nonsense mediated decay	F8WDM7
ENST00000489856.1	*CAV1*-208	575	No protein	-	Retained intron	-

2.7 PROTEIN CODING TRANSCRIPTS

CAV1-201, *CAV1*-209, *CAV1*-204, *CAV1*-205, *CAV1*-207, *CAV1*-203 and *CAV1*-202 are protein coding transcripts i.e. they contain an open reading frame (https://asia.ensembl.org/; Accessed May 5th 2024).

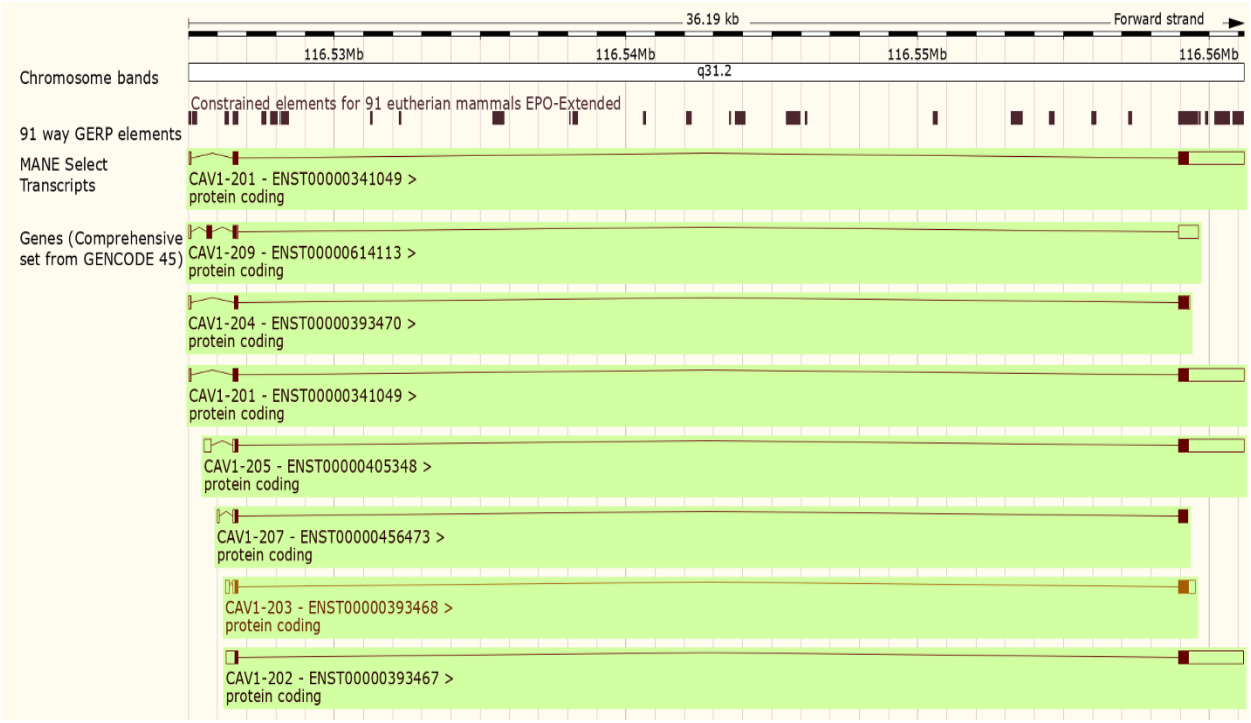


Figure 8: Protein coding transcripts of *CAV1* gene. The 7 protein coding transcripts out of 9 total transcripts of *CAV1* gene.

2.8 ISOFORMS OF CAVEOLIN 1

Caveolin 1 has two isoforms: Q03135-1 and Q03135-2. Q03135-1 is the canonical isoform reviewed in Swiss-Prot, with 178 amino acids. By depleting the 1-31 amino acids from isoform 1, Q03135-2 varies from Q03135-1. Q03135-2 has a total of 147 amino acids. However, there are 4 unreviewed computationally mapped potential isoforms of caveolin 1(https://www.uniprot.org/; Accessed on May 5th 2024) as shown in **Table *2***.

Table 2: The computationally mapped potential isoforms of *CAV1*. The table displays the entry name and length of the isoform.

ENTRY	Review status in Swiss Prot	Protein name	Gene Name	Organism	Length (Amino Acids)
Q03135	Reviewed	Caveolin 1	*CAV1*	Homo Sapiens	178
C9JKI3	Unreviewed	Caveolin	*CAV1*	Homo Sapiens	138
E9PCT5	Unreviewed	Caveolin	*CAV1*	Homo Sapiens	167
F8WDM7	Unreviewed	Caveolin 1	*CAV1*	Homo Sapiens	86
A0A7P0YWJ6	Unreviewed	Caveolin 1	*CAV1*	Homo Sapiens	115

2.9 LARGELY STUDIED PATHOGENIC VARIANTS OF *CAV1* GENE

There are 188 variants of *CAV1*, out of which 9 variants are reported in ClinVar as pathogenic. These variants include rs121434501, rs2116116365, rs879255578, rs797045176, rs587780295, rs879255566, rs587777017, rs797044871. Most of these are single-nucleotide alterations either through deletion, insertion, or substitution making them a part of large-scale studies on certain diseases (https://www.ncbi.nlm.nih.gov/clinvar/;Accessed on May 5th 2024). Further details are displayed in **Table 3**.

Table 3: Largely studied pathogenic variants of *CAV1*. The table shows the nucleotide substitution, position of substitution and the diseases caused by them.

Variation ID	Position number	Nucleotide change	Disease Caused	Pathogenicity
rs121434501	chr7:116526606	G>T	Partial lipodystrophy, congenital cataracts, and neurodegeneration syndrome 51 Congenital generalized lipodystrophy type 3 51	Pathogenic
rs2116116365	chr7:116559113-116559119[i]	delCT	Partial lipodystrophy, congenital cataracts, and neurodegeneration syndrome	Pathogenic
rs879255578	chr7:116559150	delA	Partial lipodystrophy 51, congenital cataracts, and neurodegeneration syndrome 52	Pathogenic
rs797045176	chr7:116559174	C>T	Partial lipodystrophy, congenital cataracts, and neurodegeneration syndrome, Congenital generalized lipodystrophy type 3 52	Pathogenic
rs587780295	chr7:116559151-116559152	delTTinsA	Pulmonary hypertension, primary, 3 52	Pathogenic
rs879255566	chr7:116559221-116559223	delC	Pulmonary hypertension 52	Pathogenic
rs587777017	chr7:116559224	delA	Pulmonary hypertension 52	Pathogenic
rs797044871	chh7:116559228-116559229	delTT	Partial lipodystrophy, congenital cataracts, and neurodegeneration syndrome 52, Congenital	Pathogenic

			generalized lipodystrophy type 3, Inborn genetic diseases 52	
rs2116117201	chr7:116559196-116559197	insT	Pulmonary hypertension, primary, 3	Likely Pathogenic

2.10 CAVEOLIN IN T2DM

As mentioned earlier, diabetes is caused either by failure of β cells to secrete insulin, or insulin resistant adipocytes, causing hyperglycemia. Mutation in *CAV1* may disrupt the insulin signaling and secretion, as Caveolin-1 knockout mice develop remarkable insulin resistance (52 51), while upregulation of Caveolin-1 could enhance insulin signal transduction and therefore improve glucose uptake after insulin stimulation 52. Moreover, *CAV1* acts as a molecular chaperone that is necessary for the proper stabilization of the insulin receptor in adipocytes. *CAV1* depletion *in vivo* enhances insulin secretion, moderates lipotoxicity, down regulates the expression of cell cycle arrest proteins and upregulates the expression of cell cycle activators.

2.10.1 Role of Caveolin-1 in Insulin Secretion

Caveolin-1 is known to be involved in insulin secretion from β cells of pancreas. Under physiological low glucose conditions, *CAV1* forms a complex with insulin granule proteins including the Rho GTPase cell division cycle 42 (cdc42), vesicle associated membrane protein 2 (VAMP2), and the guanine nucleotide exchange factor 7 (βPIX). Hence present in an inactivated form 51.

At high glucose conditions, the increased ATP/ADP ratio results in the closure of the ATP-sensitive K^+ channel Kv2.1, which in turn prompts the opening of the voltage-dependent Ca^{2+} channel *CAV1*.2. The increase in cytoplasmic Ca^{2+} ion concentration triggers the activation of exocytotic machinery.

The process is initiated by the dissociation of the *CAV1*-cdc42-GDP complex through active Src kinase mediated *CAV1* P-Y14 phosphorylation. The released inactive cdc42-GDP binds to βPIX resulting in conformational changes that activate cdc42-GTP which interacts with insulin secretory granule-bound VAMP2 molecules, which are then targeted to fusion with the plasma membrane through the indirect interaction between cdc42, VAMP2, F-actin (filamentous actin), Syntaxin 1A, and SNAP-25 (synaptosomal-associated protein 25) modulations and aid in secretion of insulin from vesicle (Haddad et al., 2020). *CAV1* facilitates insulin secretion through at least interacting with cdc42 and being an integral part of the insulin secretion vesicles 51 as demonstrated in **Figure 9**.

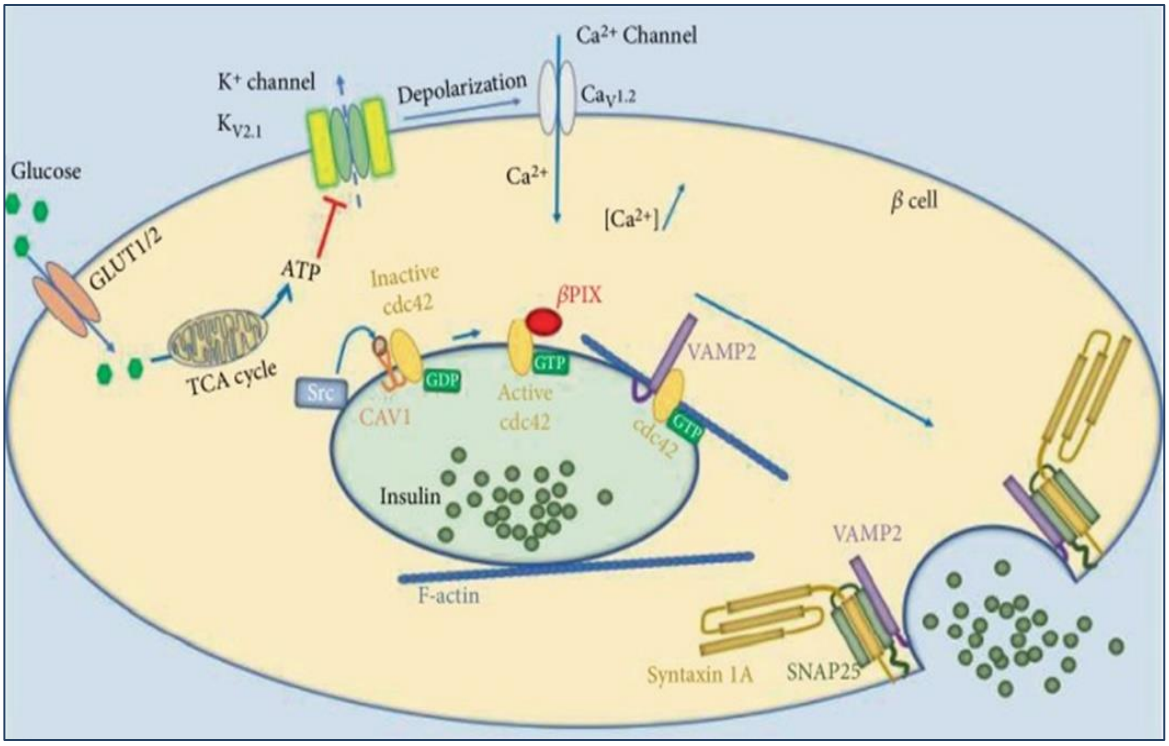


Figure 9: Role of *CAV1* in Insulin Secretion from β cells of pancreas. Abbreviations: β cell—pancreatic β cell line; GLUT—glucose transporter; TCA cycle—tricarboxylic acid cycle or Krebs cycle; cdc42—cell division cycle 42; βPIX—guanine nucleotide exchange factor 7; VAMP2—vesicle-associated membrane protein 2; F-actin—filamentous actin; SNAP-25—synaptosomal-associated protein 25.

In pancreatic cell lines INS-1 and MIN6, *CAV1* siRNA knockdown results in a significant increase in insulin secretion under physiological glucose levels 52. Hence, depletion of Caveolin-1 leads to hyperinsulinemia as seen in *CAV1* knock out mice.

2.10.2 Role of Caveolin-1 in Insulin Signaling

Caveolin-1 is a crucial molecular chaperone for stabilizing insulin receptors in adipocytes. In mice lacking caveolin-1, insulin signaling is impaired, particularly within adipocytes. Caveolin-1 is involved in Akt-mediated insulin signaling in vascular smooth muscle cells. Depletion of caveolin-1 impairs insulin sensitivity, decreases PI3K p110 expression, and inhibits Akt activation which can lead to insulin resistance and T2DM 52. Targeting caveolin-1 could be a therapeutic approach for managing T2DM, obesity, and other metabolic disorders characterized by insulin resistance. It is important for insulin sensitivity on insulin receptors as well as in translocation of GLUT4 to the plasma membrane for uptake of glucose inside the cell. During normal insulin sensitivity, insulin receptor binds insulin and activates *CAV1* by phosphorylating it at tyrosine 14 position, and then undergoes a series of autophosphorylations and activates the IRS/PI3K/AKT signaling pathway, which eventually leads to GLUT4 translocation to the surface in *CAV1*-mediated vesicles 51.

During insulin resistance, while *CAV1* and insulin receptor expression is increased, insulin receptor dissociates from *CAV1* and moves into GM3-rich lipid rafts leading to impaired insulin signaling and

decreased GLUT4 translocation to the membrane and causing a reduction in glucose uptake by the cell. As seen in

Figure *10*.

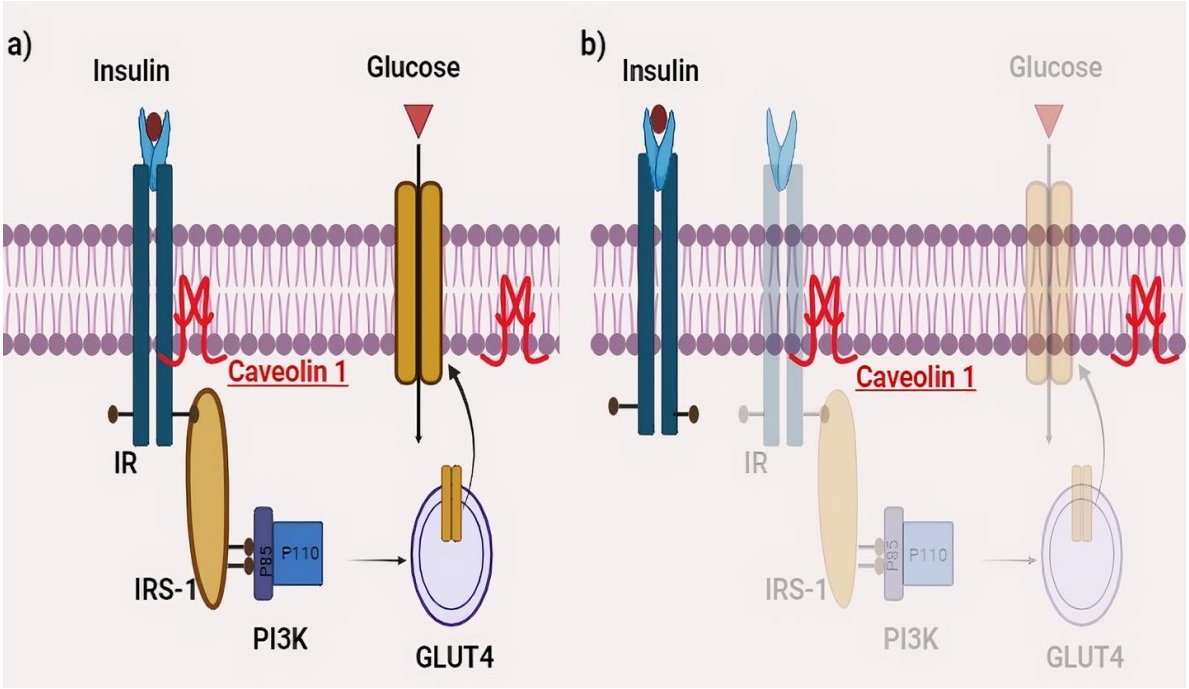


Figure 10: Role of *CAV1* in insulin signaling. a) Role of normal *CAV1* in Insulin Signaling. b) Insulin Resistance due to *CAV1* structural or functional disruption. Abbreviations: IR—insulin receptor; PI3K—phosphatidylinositol-3-kinase; IRS1—insulin receptor substrate 1; GLUT4—glucose transporter 4.

Insulin receptor is localized and highly enriched in caveolae. In pancreatic β cells, insulin receptor signaling was impaired by cholesterol depletion or by expression of a dominant-negative *CAV1* mutant. *CAV1* silencing showed a reduction in insulin induced GLUT4 recruitment to the cellular membrane and insulin receptor activation in adipocytes 51.

2.11 ASSOCIATION OF CAVEOLIN 1 WITH OTHER PROTEINS IN T2DM

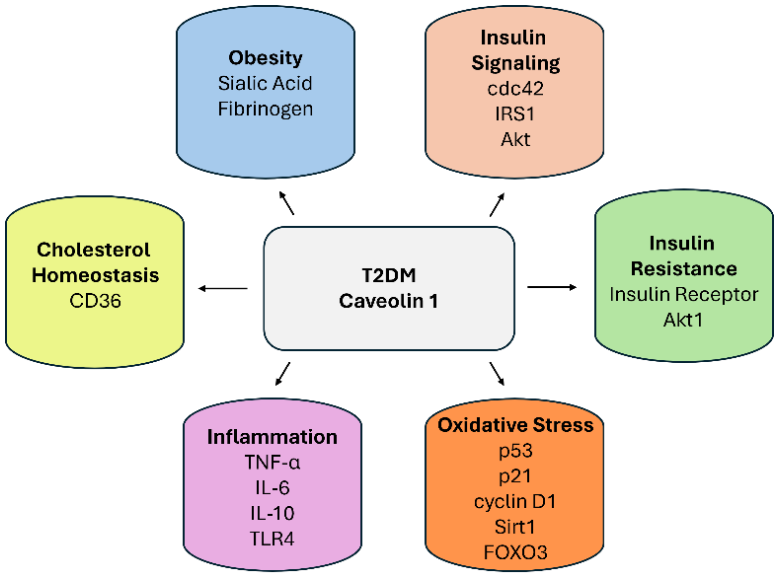


Figure 11: Shows a web diagram of the interaction of *CAV1* with other proteins in pathophysiology of T2DM (for explanation see **page 18**)

2.11.1 Obesity

CAV1 participates in the chronic inflammatory response linked to obesity and insulin resistance, characterized by cytokine production, acute-phase reactant synthesis, and proinflammatory signaling pathways. 53found that *CAV1* mRNA levels are directly related to the presence of circulating inflammatory markers like sialic acid and fibrinogen. This indicates that *CAV1* plays a role in inflammation stemming from obesity.

2.11.2 Insulin Signaling

CAV1 knockout mice show a significant decrease in insulin receptor-β tyrosine phosphorylation, because of downregulation of insulin receptor protein in *CAV1* null adipose tissue, resulting in a 90% decrease in insulin receptor 52.

(51summarize caveolin-1 to interact with cdc42 to increase insulin secretion. (53found that cholesterol decline disturbs caveolae, leading to compromised insulin signaling in a pathway that involves IRS1(insulin receptor) and Akt.

2.11.3 Insulin Resistance

(51note that high *CAV1* expression increase the quantity of insulin receptor and causes dephosphorylation of IR/IRS1/PI3K, leading to insulin insensitivity. Alternatively, insulin resistance causes the separation of insulin receptor with *CAV1*. Collectively, the disassociation of insulin receptor from *CAV1* disrupts the insulin signaling pathway.

In another study, *CAV1* knockdown not only downregulated insulin receptor expression but also decreased Akt1 phosphorylation 53.

2.11.4 Oxidative Stress

The downregulation of cyclin D1 and the elevation of p53 and p21, resulting from *CAV1* negative regulation of the ERK1/2 and PI3K/AKT signaling pathways 51.

Senescent cells, oxidative stress, and diabetes increase *CAV1* expression, suppressing Sirtuin 1, a histone deacetylase that regulates senescence. Sirt1 deacetylates FOXO3 to prevent cell cycle arrest and apoptosis 51.

2.11.5 Inflammation

Overexpression of *CAV1* decreases TNF-α and IL-6 production and increases IL-10 production in vitro. In in vivo models, decline in monocyte *CAV1* production caused TNF-α and TLR4 to increase, implying that *CAV1* takes part in TLR4-mediated inflammatory cascade in T2DM 51.

2.11.6 Cholesterol Homeostasis

CAV1 null mice show a decrease in CD36 (cell surface fatty acid binding protein). This causes a reduction in lipid droplet size in adipocytes 51.

2.12 *CAV1* GENE POLYMORPHISM IN METABOLIC SYNDROME

Some of the clinical features that make up metabolic syndrome are systemic arterial hypertension (SAH), insulin resistance, hyperglycemia, dyslipidemia, obesity, and large abdominal circumference. It has been found that elevated levels of *CAV1* mRNA have been observed in patients with metabolic syndrome (MS) 53. And the *CAV1* variants, reported in literature associated with MS clinical features are rs926198, rs11773845 and rs3807989 53. Their mRNA levels were found to be higher in fasting insulin 53. Moreover, a study done on Hispanics and Caucasian cohort proved that the *CAV1* gene variant rs926918 is associated with metabolic syndrome 53. The rs3807992 polymorphism located in the intronic region of *CAV1* gene has also been found to be associated with the metabolic syndrome risk 53.

The caveolin-1 gene variant, rs1997623, has been linked to metabolic syndrome in Arab adults and youth. However, while this association was significant in Arab populations, it showed weaker and insignificant correlations in South Asian and Southeast Asian cohorts, indicating ethnic variability in the relationship between this genetic variant and metabolic syndrome 53; 51).

2.13 *CAV1* GENE POLYMORPHISM IN CARDIOVASCULAR COMPLICATIONS

Major cardiovascular diseases (CVDs) linked with type 2 diabetes mellitus (T2DM) encompass ischemic heart disease, heart failure (HF), stroke, coronary artery disease (CAD), and peripheral artery disease. These complications can lead to mortality in over 50% of individuals diagnosed with T2DM 53. Reduced insulin sensitivity significantly affects endothelial function and plays a key role in the progression of diabetic macrovascular complications. Endothelial dysfunction, characterized by impaired vasodilation in both conduit and resistance arteries, contributes to the development of hypertension, atherosclerosis, and coronary artery disease. Studies using *CAV1*-deficient mice have demonstrated that *CAV1* downregulates endothelial nitric oxide synthase (eNOS), the primary enzyme responsible for nitric oxide (NO) production in vascular endothelial cells, thereby increasing paracellular permeability 51. The results

of a case control study done on coronary heart disease (CHD) patients in Türkiye suggested that *CAV1* rs3807990 contributes to CHD development by increasing the risk of hypercholesterolemia 53. Similarly, 54found significant association of *CAV1* SNP rs3807989 (minor allele A) with Myocardial Infarction (MI) and Coronary arterial disease (CAD) in Chinese population and the major allele G for this SNP was found to be associated with Arterial fibrillation.

2.14 *CAV1* GENE POLYMORPHISM AND CANCER RISK

CAV1 is pivotal in cancer development, being often overexpressed or mutated across various solid human tumors. Increased expression and phosphorylation of *CAV1* appear to facilitate invasion and enhance the survival of multidrug-resistant tumor cells. In multiple cancer types, elevated *CAV1* levels contribute to cancer cell survival and proliferation, promoting tumor advancement 54. The presence of the A allele in the *CAV1* gene rs3807987 polymorphism may elevate the risk of breast cancer, particularly evident in both the general population and specifically among individuals of Chinese descent. 54 did a study on Taiwanese population which showed that this variant is implicated in susceptibility to upper urothelial tract cancer among individuals. Additionally, this variant is linked to an increased susceptibility to T2DM, suggesting a potential shared genetic predisposition between breast cancer and T2DM 54; 54). The *CAV1* polymorphism shows a significant association with the risk of HCC and HBV-related hepatocellular carcinoma (HCC) 54; 54). Moreover, this variant increases the risk of urinary cancer as uncovered by 54.

The presence of rs7804372 in *CAV1*; located in intronic region, could potentially serve as a prognostic indicator for breast cancer (BC) outcomes, while also being linked to an increased likelihood of developing hypertriglyceridemia and digestive cancer (54; 54; 54). Haplotype analysis suggests that the A allele of *CAV1* rs3807987 and the T allele of *CAV1* rs7804372 could serve as promising biomarkers for early screening and risk prediction of upper tract urothelial cancer (UTUC) 54. rs7804372 may be used as a novel genomic marker for screening of nasopharyngeal carcinoma and prostate cancer as suggested by a study conducted in Taiwan (55; 55).

2.15 *CAV1* INHIBITORS AS THERAPEUTIC AGENTS

Caveolin-1, the integral protein of membrane microdomains caveolae, are essential for modulating signal transduction pathways and therefore targeting *CAV1* could be a persuasive intervention to treat diabetic complications. MβCD (Methyl-β-cyclodextrin) and daidzein are inhibitors of *CAV1* that reduces the phosphorylation and downstream signaling pathway. Research suggests that utilizing MβCD to partially remove cholesterol from muscle fibers could represent a novel therapeutic approach for treating insulin resistance in T2DM. Whereas daidzein exhibits beneficial effects on hyperglycemia, insulin resistance, dyslipidemia, obesity, and inflammation 55.

2.16 SELECTED INTRONIC SNPS

The SNPs selected for present study are **rs369262127** (A120T) and **rs201966419** (F167S)**.** These are nonsynonymous SNPs that have been found harmful and contribute to T2DM. Moreover, they have uncertain significance in ClinVar database.

- **SNP rs369262127**

This SNP is in exon 3 of *CAV1* gene.

- **SNP rs201966419**

This SNP is in exon 3 of *CAV1* gene, is reported in ClinVar and has an association with monogenic diabetes.

Previous researches have been conducted on mutations in intronic region of CAV1 gene and limited research is done on mutation in exonic region, and the association of intronic mutations has been studied with cancer, metabolic syndromes, and cardiovascular diseases, but not with T2DM. And limited research is conducted in Pakistani population. Therefore, close association of CAV1 gene with T2DM shows the urgent need to look for genetic mutations in exonic regions of *CAV1* gene particularly SNPs in Pakistani Population, which will help us to predict on individual's risk of developing disease especially T2DM and development of personalized therapeutic approaches

CHAPTER 3

METHODOLOGY

3.1 INSILICO ANALYSIS

3.1.1 SNP Data Mining

The Ensembl Genetic Variation Database was used to retrieve all gene variations of the human *CAV1* protein (http://asia.ensembl.org/; accessed June 10th 2023). Only the nsSNPs were selected for further study. The protein sequence of *CAV1* gene in FASTA format was obtained from NCBI (http://ncbi.nlm.nih.gov/; accessed June 10th 2023).

3.1.2 Identification of High-Risk SNPs

Different tools were employed to predict the deleterious and functionally damaging nsSNPs. These tools included SIFT (Sorting Tolerant From Intolerant) (Sim, N.-L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., & Ng, P. C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Research, 40(W1), W452–W457. https://doi.org/10.1093/nar/gks539 , SNPs&GO (Calabrese, R., Capriotti, E., Fariselli, P., Martelli, P. L., & Casadio, R. (2009). Functional annotations improve the predictive score of human disease-related mutations in proteins. Human Mutation, 30(8), 1237–1244. https://doi.org/10.1002/humu.21047 , and SNAP2 (Johnson, A. D., Handsaker, R. E., Pulit, S. L., Nizzari, M. M., O'Donnell, C. J., & I.W, P. (2008). SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics, 24(24), 2938–2939. https://doi.org/10.1093/bioinformatics/btn564 . nsSNPs identified as deleterious by all three tools were selected for further study.

3.1.3 Prediction of Structural and Functional Effects of nsSNPs

MutPred2 was used to predict the structural and functional effects the identified nsSNPs had on the *CAV1* protein (**Error! Reference source not found.** . Through MutPred2, we can analyze the extent of pathogenicity of a mutation. We submitted the *CAV1* protein sequence to MutPred2 in FASTA format along with the amino acid substitutions. A P-value of 0.05 indicates pathogenicity.

3.1.4 Prediction of Stability

The protein stability of the selected deleterious nsSNPs was predicted using the MUpro tool, which works using two methods: Support Vector Machines (SVM) and Neural Networks. MUpro predicts the sign and change of energy using SVM and sequence information. A confidence score was also computed, ranging from -1 to +1, using both of the two methods (Pejaver, V., Urresti, J., Lugo-Martinez, J., Pagel, K. A., Lin, G. N., Nam, H. J., Mort, M., Cooper, D. N., Sebat, J., Iakoucheva, L. M., Mooney, S. D., & Radivojac, P.

(2020). Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. Nature communications, 11(1), 5918. https://doi.org/10.1038/s41467-020-19669-x

Cheng, J., Randall, A., & Baldi, P. (2005). Prediction of protein stability changes for single-site mutations using support vector machines. Proteins: Structure, Function, and Bioinformatics, 62(4), 1125–1132. https://doi.org/10.1002/prot.20810 . Negative score suggests a decrease in the stability of protein structure.

3.1.5 Prediction of Conservation Profiles

The effects of selected nsSNPs on the evolutionary conserved sequences were predicted through ConSurf server, which analyzes the phylogenetic relation among the identified homologous sequences using an empirical Bayesian inference. Based on the phylogenetic distance, ConSurf server gives conservation scores from 1 to 9, with 9 being highly conserved (Ben Chorin, A., Masrati, G., Kessel, A., Narunsky, A., Sprinzak, J., Lahav, S., Ashkenazy, H., & Ben-Tal, N. (2019). ConSurf-DB: An accessible repository for the evolutionary conservation patterns of the majority of PDB proteins. Protein Science, 29(1), 258–267. https://doi.org/10.1002/pro.3779

3.1.6 3D Structure Prediction and Validation

The 3D structure of both the wild type and mutated *CAV1* protein was predicted through Phyre2. Phyre2 is an online tool that predicts 3D structure of the protein through homology modeling (

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols, 10(6), 845–858. https://doi.org/10.1038/nprot.2015.053. I-TASSER and GalaxyWEB were then used for 3D structure prediction and refinement respectively. I-TASSER uses threading alignments and repetitive simulations for the assembly of structures. Functional interpretation is done by comparison with 3D structures of similar proteins (120; 121; **Error! Reference source not found.** GalaxyWEB server refines loop areas using ab initio modelling and predicts protein

structure using template-based modelling.

Based on the "Seok-server" approach, it rebuilds unstable loops and constructs core structures from multiple templates (Ko et al., 2012). Further analyzing was done through Chimera 1.17.3 (Pettersen et al., 2004). The best structures obtained through I-TASSER and GalaxyWEB were submitted to TM-align which structurally aligns two proteins to calculate their RMSD (root-mean-square deviation) values and TM-score (threshold 0.5) 125. Structures with high RMSD values indicate high variation between wild type and mutant type. The models were then subjected to structural validation using two softwares; ERRAT that calculates a Quality factor using atomic interactions126 and ProSAWeb displays a Z-score127128

3.1.7 Prediction of Potential PTM Sites

Post-translational modifications (PTMs) are extremely important to a protein's function and folding. By altering the amino acid sequence, nsSNPs can cause a protein to gain or lose potential PTM site. To predict lysine and arginine methylation sites in *CAV1* protein, GPS-MSP (Deng, W., Wang, Y., Ma, L., Zhang, Y., Ullah, S., & Xue, Y. (2016). Computational prediction of methylation types of covalently modified lysine and arginine residues in proteins. Briefings in Bioinformatics, bbw041. https://doi.org/10.1093/bib/bbw041 was used, while phosphorylation at serine, threonine, and tyrosine sites was predicted through NetPhos3.1 (Blom, N., Gammeltoft, S., & Brunak, S. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. Journal of Molecular Biology, 294(5), 1351–1362. https://doi.org/10.1006/jmbi.1999.3310. We report amino acids above the threshold cycle of 0.5 to be potentially phosphorylated in NetPhos3.1. The potential glycosylation sites of *CAV1* protein were predicted through NetOglyc4.0 (Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T-B. G., Lavrsen, K., Dabelsteen, S., Pedersen, N. B., Marcos-Silva, L., Gupta, R., Paul Bennett, E., Mandel, U., Brunak, S., Wandall, H. H., Levery, S. B., & Clausen, H.

(2013). Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. The EMBO Journal, 32(10), 1478–1488. https://doi.org/10.1038/emboj.2013.79 and NetNGlyc 1.0 (Gupta, R., & Brunak, S. (2002). Prediction of glycosylation across the human proteome and the correlation to protein function. Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing, 310–322. https://pubmed.ncbi.nlm.nih.gov/11928486/. A location having glycosylation potential of <0.5 (below threshold) in NetNGlyc1.0 predicts a negative site. O-Glycosylation is thought to occur at sites with a potential of >0.5 (above threshold) in NetOglyc4.0.

3.1.8 Protein-Protein Interactions

The function and regulation of a protein depends upon its interaction with other proteins. We used GeneMANIA and STRING to predict the functional interaction of *CAV1* with other proteins. Through proteomic and genomic data, GeneMANIA creates a list for the gene in question with additional genes that are functionally related to it. Additionally, it illustrates the connection between the genes and the query (Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C.,

Kazi, F., Lopes, C. T., Maitland, A., Mostafavi, S., Montojo, J., Shao, Q., Wright, G., Bader, G. D., & Morris, Q. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Research, 38(suppl_2), W214–W220. https://doi.org/10.1093/nar/gkq537 . STRING has a database of proteins from 5,090 organisms and predicts protein-protein interaction networks if the proteins have a direct or indirect association (Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & Mering, C. von. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Research, 47(Database issue), D607–D613. https://doi.org/10.1093/nar/gky1131 . We used 'homo sapiens' and '*CAV1*' as input options for both tools.

3.1.9 ClinVar Prediction

ClinVar is a publicly accessible database that categorizes human variations by diseases and treatment reactions, providing a platform for discussing hypotheses and supporting data, following the HGVS standard for reporting and mapping variations to reference sequences (Landrum et al., 2014). The rsIDs of the six SNPs were submitted in ClinVar and their results were noted.

3.2 IN-VITRO ANALYSIS

3.2.1 Study Subjects and Sample Collection

The underlying case control study was conducted to analyze the role of *CAV1* (rs369262127 and rs201966419) gene polymorphisms in patients with T2DM in comparison with normal control.

Informed written consent was taken from all the patients and normal controls who participated in this study. A total of 36 human subjects were included in this study among which 30 were clinically diagnosed T2DM patients and 6 were normal controls. Samples were collected from PIMS hospital Islamabad, Polyclinic

hospital and Research and Diagnostic laboratory ASAB over a period of 2 months.

3.2.2 Criteria of Inclusion and Exclusion

The following inclusion and exclusion criteria was observed.

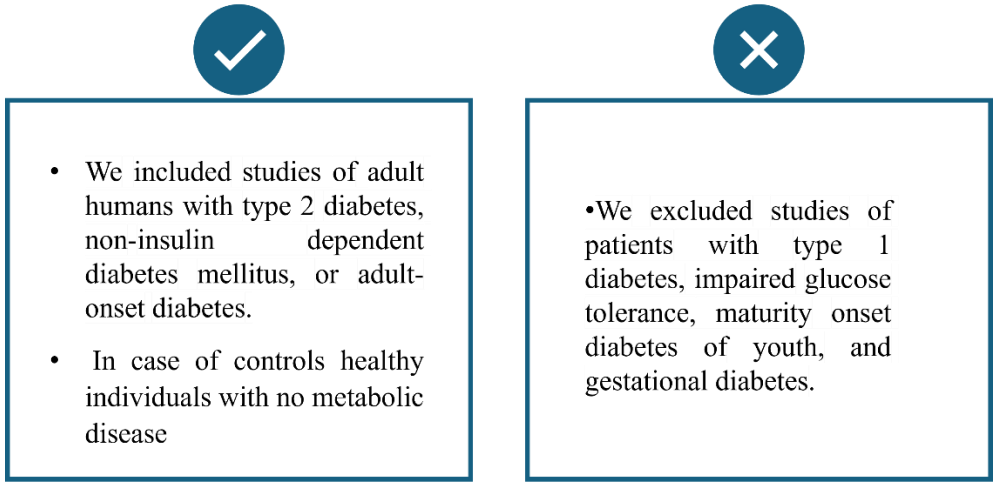


Figure 12: Inclusion and exclusion criteria.

3.2.3 Collection and Storage of Blood Samples

Blood samples from the subjects were collected in 3ml Ethylenediaminetetraacetic acid (EDTA) blood collecting tubes by hospital staff. EDTA tubes were properly labelled with name and age of patients and controls as well as date of collection of blood sample. EDTA tubes containing blood samples were transported to the laboratory by placing them in icebox and were placed at 4°C in refrigerator.

3.2.4 Extraction of Genomic DNA by Phenol-Chloroform Method

DNA was extracted from the whole blood samples by using Phenol-Chloroform method of DNA extraction. It is a two-day protocol which is cost effective as well as a reliable method of DNA extraction giving satisfactory results. All the glassware and plastic wares used in DNA extraction were properly washed and autoclaved before use.

3.2.4.1 Preparation of Solutions for DNA Extraction

Composition and function of all the solutions used in Phenol-Chloroform method of DNA extraction are as follow:

Table 4: Solutions required for DNA extraction and their functions.

Component	Molarity	Quantity	Function
Solution A			

Sucrose	0.32M	109.55 g	Facilitates the release of cell components including DNA by causing cell lysis and breakage of cell membrane.
Tris (pH 7.5)	10mM	12.114 g	
Magnesium Chloride	5mM	0.476 g	
Autoclaved Distilled water	-	Up to 1000 ml	

Solution B

Tris (pH 7.5)	10mM	12.114 g	Results in DNA precipitation and separation of proteins
Sodium chloride	400mM	23.37 g	
Ethylene Diamine Tetra Acetic Acid (EDTA)	2mM	0.58 g	
Autoclaved Distilled water	-	Up to 1000 ml	

Solution C

Phenol	-	250 µl	Helps in the separation of DNA from protein and other cell debris by forming an aqueous layer

Solution D

Chloroform	-	48 µl	Involved in DNA purification by stabilizing the coagulated proteins and reducing foaming
Iso-amyl Alcohol	-	2 µl	

20% SDS Solution

Sodium Dodecyle Sulphate (SDS)	-	20 g	Involved in denaturation of

Autoclaved Distilled Water		Up to 1000 ml	proteins and lipids and separates them from DNA

3M Sodium Acetate Solution

Sodium Acetate	3M	12.3 g	Involved in precipitation of DNA by neutralizing the negative charges on phosphate backbone
Autoclaved Distilled water	-	Up to 50 ml	

3.2.4.2 Reagents for DNA Extraction

Other chemicals needed in Phenol-Chloroform method of DNA Extraction and their functions are as follow:

Table 5: Reagents required for DNA extraction and their function.

Component	Function
Proteinase K	Involved in digestion of proteins
Isopropanol	Involved in precipitation of DNA by removing the solvation cell that surrounds the DNA
98% Ethanol	Involved In precipitation of DNA
PCR water	Used for storage of DNA as it prevents degradation of DNA

3.2.4.3 Protocol of DNA Extraction

Day 1:

1. Blood in EDTA tube was kept at room temperature for some time and mixed well by inverting the tube several times. 750 µl of blood was then added in a 1.5 ml microcentrifuge tube and same amount of solution A i.e., 750 µl was added in it.
2. The microcentrifuge tube was inverted 4-6 times and then kept at room temperature for 10 minutes.
3. The mixture in microcentrifuge tube was then centrifuged at 13,000 revolutions per minute (rpm) in a microcentrifuge for 10 minutes.
4. After centrifugation, the supernatant was carefully discarded, and the nuclear pallet was resuspended in 400 µl of solution B.
5. The nuclear pallet was dissolved completely in solution B by continuous tapping or slight vertexing.

6. The mixture was again centrifuged at 13,000 revolutions per minute (rpm) for 10 minutes.
7. After centrifugation the supernatant was again discarded and 400 µl of solution B, 15 µl of 20% SDS solution and 8 µl of Proteinase K was added in microcentrifuge tube.
8. The mixture was then incubated at 37°C overnight.

Day 2

1. A fresh mixture of 250 µl each of solution C and solution D was prepared in a microcentrifuge tube and then was added to the microcentrifuge tube incubated overnight.
2. The resulting mixture was centrifuged at 13,000 revolutions per minute (rpm) for 10 minutes.
3. After centrifugation the aqueous layer formed, that contains the DNA, was transferred into a new microcentrifuge tube.
4. 500 µl of solution D was added in the separated aqueous layer and it was then centrifuged at 13000 revolutions per minute (rpm) for 10 minutes.
5. The resulting aqueous layer was again transferred to a new microcentrifuge tube and then 55 µl of 3M sodium acetate solution and 500 µl of isopropanol was added in it.
6. DNA was allowed to precipitate by inverting the tube several times and then mixture was again centrifuged for 10 minutes at 13000 revolutions per minute (rpm).
7. After centrifugation the supernatant was discarded and 200 µl of chilled 98% ethanol was added in the microcentrifuge tube containing DNA pellet.
8. Centrifugation was done for 8 minutes at 13000 revolutions per minute (rpm).
9. After centrifugation the ethanol was evaporated by air drying and DNA pellet was submerged in 100 µl of PCR water and stored at -20°C.

3.2.5 Gel Electrophoresis of DNA

To analyze the quality of extracted DNA 2% (w/v) agarose gel electrophoresis was performed. The details of preparation of buffer solutions and reagents required for gel electrophoresis are as follow:

3.2.5.1 0.5 M EDTA Solution

0.5 M EDTA solution was prepared by dissolving 186.12 grams of EDTA in distilled water to a final volume of 1000 ml which was then used in preparation of TAE buffer for gel electrophoresis. pH of 0.5M EDTA solution should be 8.

3.2.5.2 50X Tris Acetate Ethylene-Diamine-Tetra-Acetic Acid (TAE) Buffer

50X stock solution of Tris Acetate Ethylene-diamine-tetra-acetic acid (TAE) Buffer was prepared as follow:

Table 6: Components required for preparation of TAE buffer.

Component	Quantity
Tris Base	242 g
0.5M EDTA Solution	100 ml

Glacial Acetic Acid	57.1 ml
Deionized water	Up to 1000 ml

pH of 50 X TAE buffer was adjusted between 8.2 to 8.4. After preparation 50X TAE

buffer was autoclaved and stored at room temperature.

3.18.6 1X Tris Acetate Ethylene-Diamine-Tetra-Acetic Acid (TAE) Buffer

1X TAE buffer was prepared from 50X TAE buffer by taking 20 ml of 50X TAE buffer and adding 980 mL of distilled water in it.

3.2.5.3 Protocol

1. 2g of agarose was weighed by using electronic weighing balance and was added into 100 ml of 1X TAE buffer.
2. The mixture was dissolved by heating it in microwave oven for approximately 2 minutes.
3. After slight cooling, 4 µl of ethidium bromide was added for staining of DNA.
4. The gel solution was then poured into gel casting tray and was allowed to solidify at room temperature. Wells for sample loading were made with the help of comb.
5. When solidified, gel was placed in electrophoresis tank which was filled with 1X TAE buffer.
6. 2 µl of loading dye was used to stain 6 µl of extracted DNA sample each, which was then loaded into gel wells.
7. 1 kbp DNA ladder was also loaded in gel well for the purpose of determining the length of DNA sample.
8. Gel electrophoresis was performed at 80 volts for 30 minutes.
9. After completion of gel electrophoresis, gel was placed in UV Transilluminator as well as ChemiDoc and DNA bands were analyzed and compared with DNA ladder.

3.2.6 Quantification of DNA

ThermoScientific Nanodrop 2000 UV-Vis Spectrophotometer and NanoDrop 2000™ software was used to quantify the extracted DNA samples. 1 µl of PCR water was first placed on pedestal which acts as blank and then 1 µl of DNA sample was loaded to analyze the absorbance ratio at 260 nm wavelength which was kept as standard. The purity of DNA was determined based upon the absorbance ratio of 260/280nm and a ratio of ~1.8 is considered as in case of DNA.

3.2.7 Primer Designing

Primers were designed manually for each gene polymorphism to carry out Gradient PCR. FASTA sequence

of each SNP was retrieved from Ensemble. For each SNP three primers were designed as shown in **Table 7**.

Table 7:Primers designed for Gradient PCR. The SNP specific primer was the reverse primer. The melting temperature, GC content and product sizes are also mentioned.

SNP	Variation	Primer Sequence (5'-3')	Tm	GC Content	Product Size
rs369262127	G/A	**Wildtype Forward** AAGGTCCAAGGGGAACCCAA	61.00	55	695
		Wildtype Reverse CGGTGTGGACGTAGATGGAA	59.47	55	
		Mutant Reverse TGCAGGAAAGAGAGAATGGTG	57.90	47.6	
rs201966419	T/C	**Wildtype Forward** AAGGTCCAAGGGGAACCCAA	61.00	55	826
		Wildtype Reverse CGGTGTGGACGTAGATGGAA	59.47	55	
		Mutant Reverse CATTGCTGGATATTTTCCCAAC	56.5	40.9	

A common wildtype forward primer and two reverse primers (one for each SNP) were designed. Eurofins software was used to determine the GC content, melting temperatures (Tm), hair-pin formation and self-complementarity of primers. Primer specificity was determined by using Primer-Blast as well as UCSC in-silico PCR software which gives information about amplicon size as well as shows the target size where primers bind.

3.2.7.1 Preparation Of Working Dilutions of Primers

Primers were received in a lyophilized state, so first 100 µM stock was prepared for each primer. The amount of nuclease free water added into the lyophilized primers was determined by multiplying the nmol of primer by 10. After preparation of stock working dilutions of primers were made to avoid contamination and reduce freeze/thaw cycles of stock. Working dilutions were made by adding 10 µl of primer from stock and 90 µl of nuclease free water in a sterile microcentrifuge tube.

3.2.8 Gradient PCR

Gradient PCR was carried out on extracted DNA samples of subjects to validate the presence of SNP. For each SNP, two separate PCR reactions were carried out. Common primer was added in both the reactions whereas one reaction was carried forward using wildtype primer and other by using mutant primer. If wild

type SNP is present in sample only that PCR reaction will amplify DNA that contains wildtype primer and same will be in case of mutant SNP. In this way the presence of SNP will be indicated in samples. A reaction mixture of 20 µl was made in each PCR tube. Reagents of each PCR tube shown in **Table 8**.

Table 8: Components of PCR reaction mixture.

Components	Quantity
PCR Water	12 µl
Forward Primer	1 µl
Reverse Primer	1 µl
DNA Template	2 µl
PCR Master Mix (contains DNA Polymerase, MgCl2, PCR Buffer, dNTPs)	4 µl
Total Volume	20 µl

The reaction mixture was prepared in Biosafety cabinet to avoid contamination and on ice to prevent activation of DNA polymerase enzyme. After preparation, reaction mixture was short spun for few seconds for removal of air bubbles and homogenous mixing. The PCR tubes were then placed in a Thermocycler for amplification. First the primers were optimized using gradient PCR at annealing temperatures ranging from 57°C to 63°C and then optimized primers were used to amplify sample DNA and detect SNP.

Three vials were prepared for each sample. Vial 1 contained the forward and reverse primers of the wildtype CAV1 gene. The second vial was meant to amplify the first mutant (rs201966419) and the third vial amplified the second mutant (rs369262127) if present. This is displayed diagrammatically in **Figure 13**.

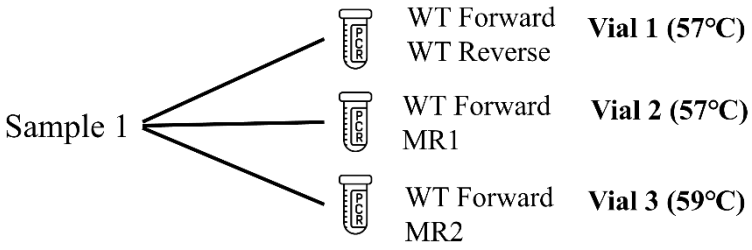
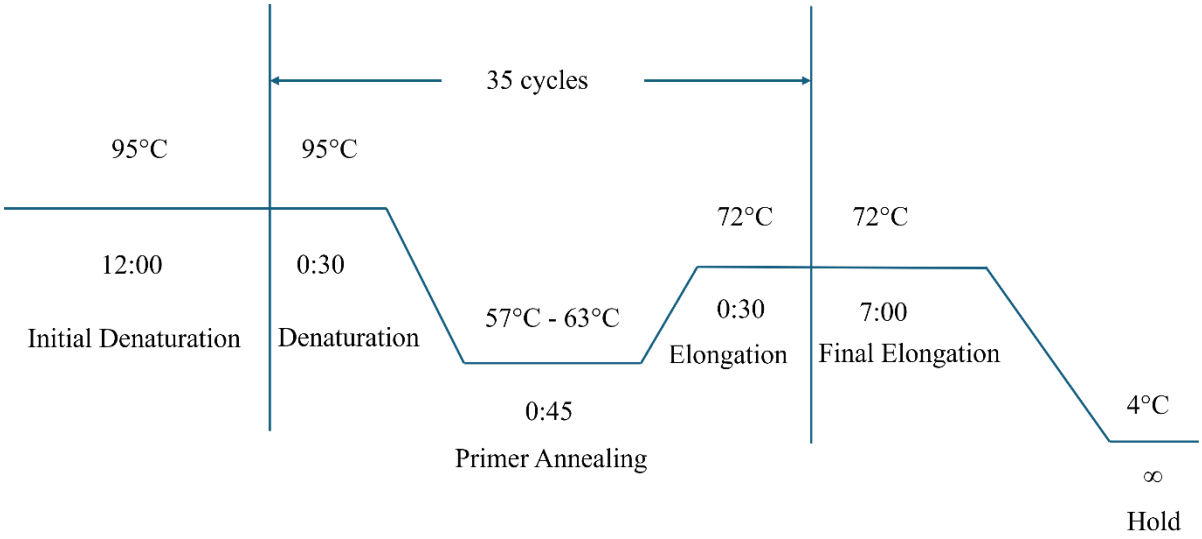


Figure 13: The PCR vials prepared for each sample. The first vial contained forward and reverse primers of the wildtype CAV1 gene. The second vial was meant to amplify the first mutant (rs201966419) and the third vial amplified the second mutant (rs369262127) if present.

The optimized conditions for PCR reaction for rs369262127 and rs201966419 are shown in **Figure 14**

Figure 14: Schematic representation of Gradient PCR Profile. The first step is the denaturation of double-stranded DNA at 95°C, followed by annealing of primer at optimized annealing temperature (57°C to 63°C) and finally elongation of the DNA strand at 72°C.



3.2.9 Gel Electrophoresis for Analysis of PCR Products

For analysis of Gradient PCR products 2 % gel electrophoresis was performed again with the same protocol mentioned before but with a change that 50 bp ladder was used and gel was run for 45 minutes at 80 volts. After that gel was visualized using UV Transilluminator and ChemiDoc and PCR products were analyzed and compared with ladder.

3.2.10 Statistical Analysis

Statistical analysis was performed using GraphPad Prism version 10. Fisher's exact test was performed to check the frequency of occurrence of variation and the subsequent probability value. The test also gave the value of relative risk. P*<0.05 is considered significant and risk value >1 is considered significant.

CHAPTER 4

RESULTS

4.1 IN-SILICO ANALYSIS OF MISSENSE SNPS OF *CAV1*

4.1.1 SNP Retrieval

The gene variant table for *CAV1* protein on Ensemble returned a total of 77,626 SNPs. Out of these, 72,812 were intronic region SNPs. The exonic SNPs included 747 missense SNPs and 356 synonymous SNPs. 487 located in 5'UTR, while 2,463 were present in 3'UTR region. The remaining 761 SNPs included miscellaneous SNPs like stop site SNPs and undefined SNPs. This study only concerned itself with missense or nsSNPs for further in silico analysis. **Figure 15** shows the graphical representation of all the SNPs.

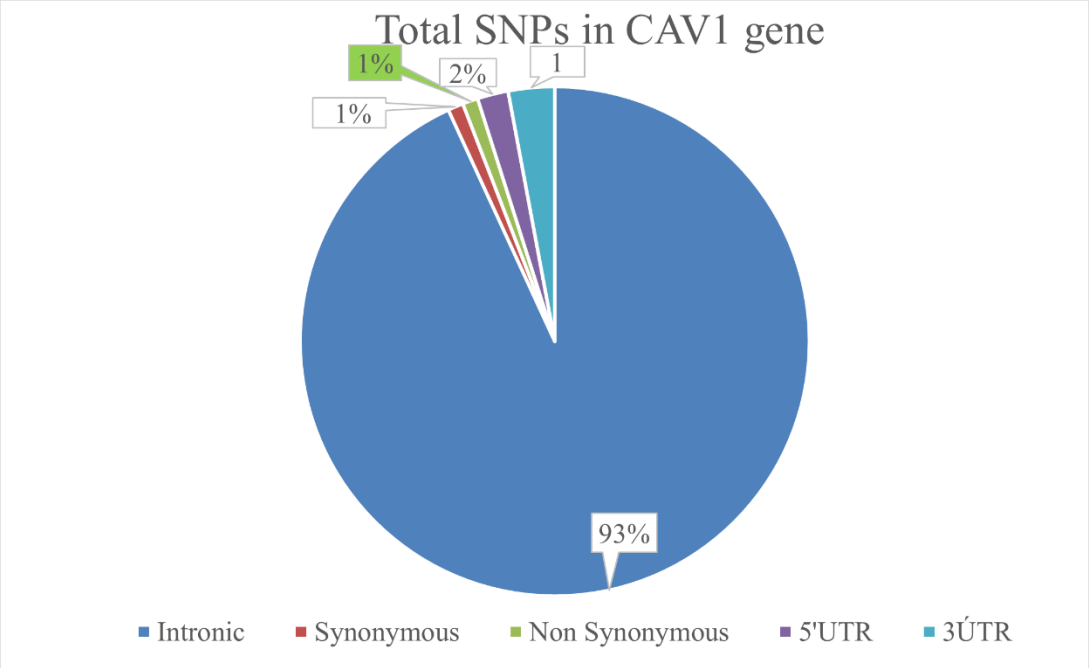


Figure 15: Pie chart of *CAV1* gene variants. A pie chart displaying the results for *CAV1* gene variants obtained using Ensembl. 72,812 of the total SNPs were intronic region SNPs. The exonic SNPs included 747 missense SNPs and 356 synonymous SNPs. 487 located in 5'UTR, while 2,463 were present in 3'UTR region. The remaining 761 SNPs included miscellaneous SNPs like stop site SNPs and undefined SNPs.

4.1.2 Identification of Damaging SNPs

The obtained nsSNPs were run through SIFT to identify the deleterious SNPs. A SIFT scorethreshold of 0.05 was taken and nsSNPs that had a SIFT score less than this threshold value were deemed to be affected. SIFT predicted 44 nsSNPs to be deleterious. The *CAV1* protein sequence was also given to SNAP2 as input which predicted the effect of all possible 3560 mutations in the sequence. By taking variations that had a

score of higher than 50 as deleterious, we identified 1454 such variations. Variations common in SIFT and SNAP2 were selected and were analyzed to SNPs&GO, which highlighted six nsSNPs to be deleterious (see **Table 9**). We continued our computational analysis with these 6 nsSNPs. gives the nsSNPs detected by SIFT, SNAP2, and SNPs&GO as deleterious.

Table 9: The table below shows the rsID of the SNPs, along with the name and position change in the residue from the wild type CAV1 sequence.

SNP ID	Residual Change	Position	Wildtype Amino Acid	Mutated Amino Acid
rs369884333	S80G	80	S	G
rs372416448	T95M	95	T	M
rs200052661	R101H	101	R	H
rs369262127	A120T	120	A	T
rs201966419	F167S	167	F	S

4.1.3 Prediction of Structural and Functional Effects of nsSNPs:

MutPred2 was used to predict the effect of the six shortlisted nsSNPs. The predicted effects include altered transmembrane protein, altered ordered interface, gain of relative solvent accessibility, altered DNA binding, loss of allosteric site and gain of GPI-anchor amidation. MutPred2 predicted an alteration of transmembrane protein for all nsSNPs and predicted altered ordered interface in all nsSNPs except S80G. The mutation of F167S was of particular interest as it was predicted to cause allosteric site loss and altered DNA binding among other effects. The results of MutPred2 are given in

Table *10*.

Table 10: Shows the results of Mutpred2.The Mutpred score, predicted effects in the protein structure and the p-values of the predicted effects are shown.

SNP ID	Residual Change	MutPred Score	Predicted Effects	*P*-Values of Predicted Effects
rs369884333	S80G	0.853	Altered Transmembrane protein	2.4e-03
rs372416448	T95M	0.827	Altered Transmembrane protein Altered Ordered interface	0.02 0.02

rs200052661	R101H	0.668	Altered Transmembrane protein Altered Ordered interface	0.04 0.02
rs369262127	A120T	0.825	Altered Transmembrane protein Altered Ordered interface	1.0e-03 0.02
rs201966419	F167S	0.871	Altered Transmembrane protein Altered Ordered interface Gain of Relative solvent accessibility Altered DNA binding Loss of Allosteric site at F167 Gain of GPI-anchor amidation at N16	0.04 0.03 0.02 8.9e-03 0.03 0.03

4.1.4 Prediction of Stability

A protein's function is related to its stability; hence it's crucial to determine whether nsSNPs have altered a protein's stability. The MUpro tool was used to assess the effect on stability of protein structure because of damaging nsSNPs. The six nsSNPs were submitted into the software and ΔG (difference of free energy between native and denatured states of protein) and confidence scores were obtained. All the nsSNPs decreased the structural stability of Caveolin 1 protein (negative confidence scores). T95M, and F167S were the most damaging mutations. The results are displayed in **.**

Table *11*.

Table 11: The table below shows the results of protein stability in all the *CAV1* variants assessed using MUpro. ΔG score, SVM confidence score and Neural network confidence score are displayed.

SNP ID	Residual Change	Stability	ΔG	SVM Confidence Score	Neural Network Confidence Score
rs369884333	S80G	Decrease	-1.20	-0.34	-0.58
rs372416448	T95M	Decrease	-0.01	-0.78	-0.58
rs200052661	R101H	Decrease	-1.15	-0.49	-0.69
rs369292127	A120T	Decrease	-1.38	-0.03	-0.61
rs201966419	F167S	Decrease	-1.40	-1	-1

4.1.5 Conservation of Amino Acids

The highly conserved regions in a protein structure usually have a high functional importance. We checked for the conservation profile of amino acids of *CAV1* through ConSurf. S80G and A120T were predicted to be highly conserved and buried; T95M and F167S were predicted to be buried, while R101H was predicted to be exposed. The amino acids and their respective conservation scores are given in **Table 12Table *12*.**

Table 12: The table below shows the result of ConSurf conservation profiles. A conservation score (range: 1-9) is displayed, with 9 being highly conserved.

SNP ID	Residual Change	Conservation Score	Prediction
rs369884333	S80G	9	Highly conserved and buried (s)
rs372416448	T95M	5	Buried
rs200052661	R101H	7	Exposed
rs369262127	A120T	9	Highly conserved and buried (s)
rs201966419	F167S	7	Buried

4.1.6 3D Structure Prediction

Phyre2 was used to initially predict the structures of the wild and mutant types of *CAV1* protein. However, the Phyre2 results were incomplete as it only predicted the structure for 128 amino acids (72% coverage) out of 178 amino acids. We then used I-TASSER to generate 3D models for the wild type *CAV1* protein as well as the six mutants as it is more accurate. I-TASSER produced five models, but only those that demonstrated the highest confidence score were analyzed using Chimera 1.17.3. The results are shown in **Figure 16** and **Figure 17.** The chosen models were submitted in TM-Align to calculate RMSD values and TM-score. All models differed slightly from the wildtype. T95M and R101H displayed the highest RMSD. The results are given in **Table 13.**

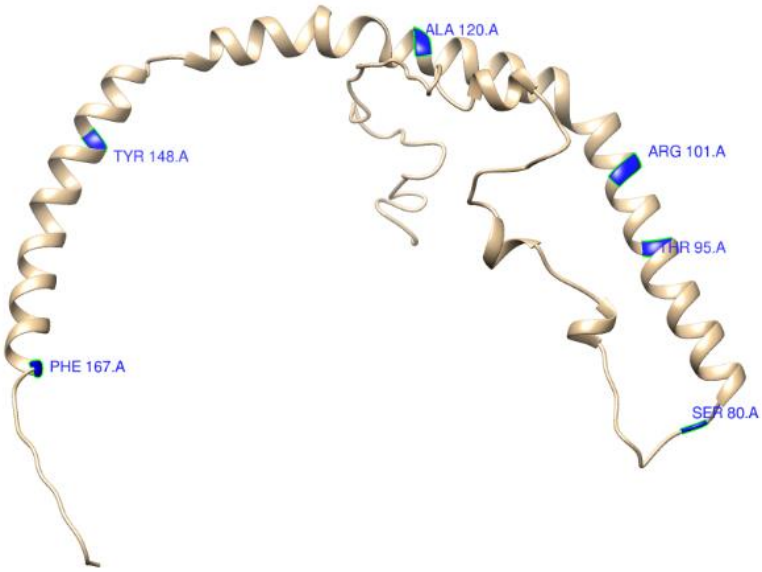


Figure 16: Shows the 3D structure of wildtype Caveolin-1 protein. The annotations in blue represent all six SNP positions and their wildtype residues; Serine at position 80, Threonine at position 95, Arginine at position 101, Alanine at position 120, Tyrosine at position 148 and Phenylalanine at position 167.

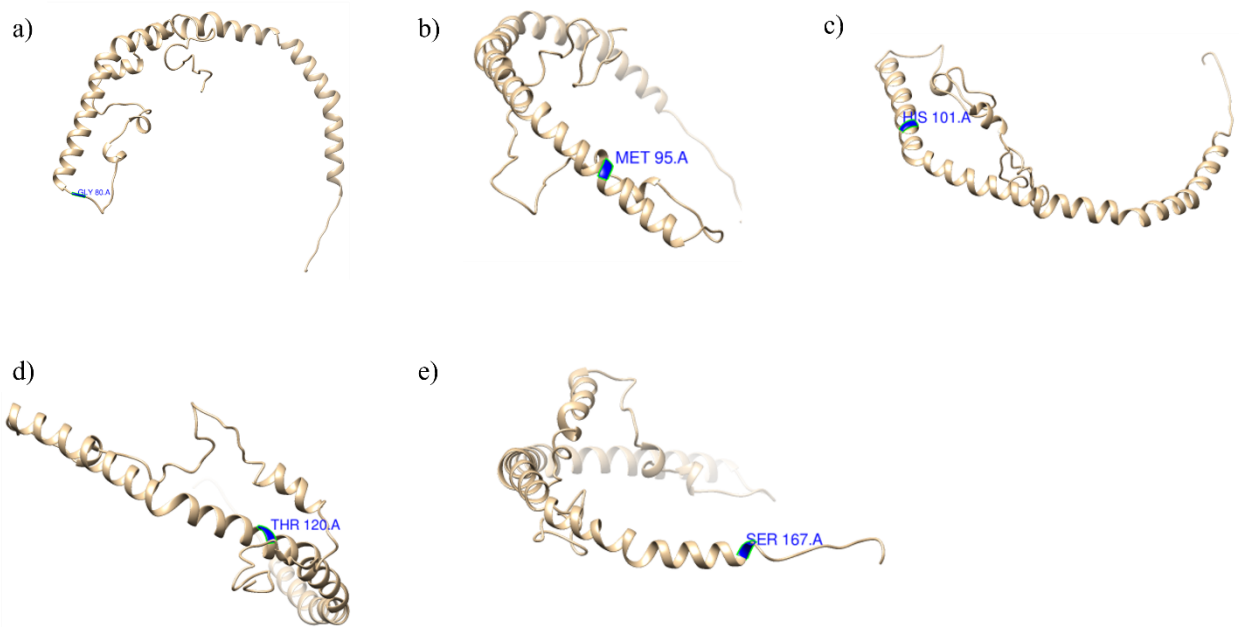


Figure 17: The figure shows 3D models of mutant Caveolin-1 proteins generated by I-TASSER and GalaxyWEB. The SNP positions are annotated in blue. a) Glycine at position 80, b) Methionine at position 95, c) Histidine at position 101, d) Threonine at position 120, e) Serine at position 167.

Table 13: The table shows the TM-score and RMSD values of mutant models obtained through TM-Align.

SNP ID	Residual Change	TM-score	RMSD/Å
rs369884333	S80G	0.85957	2.57
rs372416448	T95M	0.81837	2.91
rs200052661	R101H	0.83237	2.70
rs369262127	A120T	0.69153	2.55
rs201966419	F167S	0.86743	2.18

For the sake of accuracy, it is imperative that flaws in the protein structures be pointed out. Thus, two distinct protein structural validation tools were used to validate the 3D models: ProSAWeb and ERRAT. Quality Factor >50 and Z-score in the range -10 to 10 are considered acceptable. For the wildtype Caveolin 1, ERRAT displayed a Quality factor of 85.119 and ProSAWeb showed a Z-score of 1.39. Both of these values indicate a high-quality model. Structure validation was also performed for the entire mutant models obtaining Quality factor, Z-score. All the models were of a high quality. The results are given in **Table 14**.

Table 14: Shows the results of structural validation. ERRAT produced the Quality factor and ProSAWeb provided the Z-score.

Residual Change	Quality Factor	Z-score
Wildtype	85.119	1.39
S80G	84.049	0.68
T95M	75.610	1.40
R101H	86.391	1.23
A120T	72.455	0.72
F167S	67.059	1.03

4.1.7 Prediction of PTM Sites

The post-translational modification sites for both wild type and mutated *CAV1* protein were predicted by different tools. The results are given below.

4.1.7.1 Methylation: GPS-MSP 3.0 predicted no methylation sites in Caveolin 1.

4.1.7.2 Phosphorylation: NetPhos3.1. determined the serine at position 80 to be phosphorylated by Casein Kinase I (CKI), and the threonine at position 95 was predicted to be phosphorylated by Protein Kinase C (PKC). The substitutions S80G and T95M caused a loss of these PTM sites. A score of >0.5 indicates a positive prediction. The results of NetPhos3.1 are given in Table 15.

Table 15: Shows the results of NetPhos3.1. The phosphorylation sites, enzymes and scores are mentioned. A score of >0.5 indicates a positive prediction.

SNP ID	Residual Change	Enzyme	Score
rs369884333	S80G	CKI	0.526
rs372416448	T95M	PKC	0.639

4.1.7.3 Glycosylation: NetNGlyc 1.0 and NetOGlyc4.0 found no N-glycosylation or O-glycosylation sites in *CAV1* protein.

4.1.8 Gene-Gene Interactions

GeneMANIA and String were used to predict the gene-gene interactions of *CAV1* protein. GeneMANIA predicted a physical interaction of *CAV1* gene with *CAV2, CAV3, NEU3, NOS3, WASL, NOSTRIN, KCNH2, EGFR, HDAC6, GRB7, CAVIN1, PLIN1, RAC1, TXNRD1, PTGIS, PTGS2, LRP6, BSG* and *F2R*. *CAV1* had a co-expression with *CAV2, CAV3, NEU3, NOS3, EGFR, HDAC6, GRB7, CAVIN1, PLIN1, RAC1, TXNRD1, PTGIS, PTGS2, STRN* and *F2R*. *CAV1* was thought to be co-localized with *CAV2, PTGIS* and *PTGS2*. Genetic interaction was predicted to be with *GRB7, STRN, TXNRD1* and *PTGIS*. *CAV1* also shared a pathway with *NOS3, WASL NOSTRIN, EGFR, RAC1, GRB7, PLIN1, LRP6* and *BSG*. *CAV1* only shared protein domains with *CAV2* and *CAV3*. The results are given in **Figure 18.**

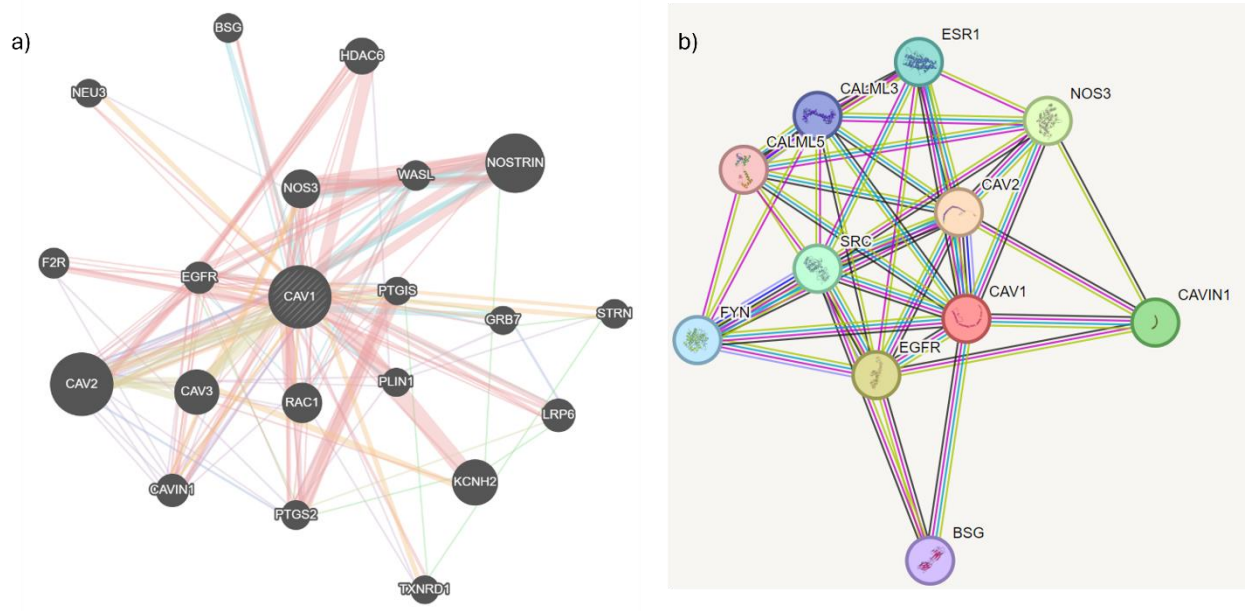


Figure 18: Gene to gene interactions. a) The figure shows gene to gene interactions of *CAV1* with other genes using GeneMANIA. Each thread color represents a different type of interaction between genes. Pink (physical interaction); purple (co-expression); orange (predicted); dark blue (co-localization); green (genetic interactions); light blue (pathway); and yellow (shared protein domains). b) The result of gene-to-gene interaction using STRING.

4.1.9 ClinVar Prediction

ClinVar was used to predict clinical significance of the nsSNPs. The results are displayed in **Table *16***.

Table 16: ClinVar prediction results. The table displays the variation ID, accession number and its prediction in ClinVar.

SNP	Variation ID	Accession	ClinVar Prediction
rs369884333	548689	VCV000548689.2	Uncertain Significance
rs372416448	-	-	Not Reported in ClinVar
rs200052661	1469176	VCV001469176.5	Uncertain significance
rs369262127	2894539	VCV002894539.1	Uncertain significance (However, it is expected to disrupt *CAV1* protein function)
rs201966419	541324	VCV000541324.10	Conflicting classifications of pathogenicity Uncertain significance (1); Likely benign (1)

The results show only rs201966419 to have some clinical significance as reported in research. rs369262127 although has uncertain significance, it still was mentioned to disrupt *CAV1* function according to a 2024 study. Hence these two SNPs were shortlisted for further invitro analysis.

4.2 IN-VITRO ANALYSIS

A case control study was conducted to check the association of rs369262127 and rs201966419 both missense SNPs of *CAV1* with the risk of T2DM in Pakistani population. A total of 36 whole blood samples (30 diseases samples and 6 healthy controls) were collected from Polyclinic hospital, PIMS hospital Islamabad and Research and Diagnostic laboratory ASAB. DNA was extracted from the whole blood samples using Phenol-Chloroform method of DNA extraction. The quality and quantity of extracted DNA was determined through 2% agarose gel electrophoresis and ThermoScientific Nanodrop 2000 UV-Vis Spectrophotometer and NanoDrop 2000™ Software respectively. The pure DNA samples having absorbance ratio ~1.8 were further analyzed by performing gradient PCR to screen the samples for the presence of polymorphism. 2% agarose gel electrophoresis was performed to analyze the PCR products. Statistical analysis was performed on the data obtained from Gradient PCR to check any significant association of the two polymorphisms with T2DM within Pakistani population.

4.2.1 Association Analysis of *CAV1* rs369262127 and rs201966419 Polymorphism

CAV1 rs369262127 is a missense SNP (G>A), where G is an ancestral allele and A is the risk allele. It is present in exon 3 of the gene on chromosome no 7 at position chr7:116559108 and results in a single amino acid change from adenine to threonine (A120T). *CAV1* rs201966419 is a missense SNP (T>C), where T is an ancestral allele and C is the risk allele. It is present in exon 3 of the gene on chromosome no 7 at position 7:116559250 and results in a single amino acid change from phenylalanine to serine (F167S). For these SNPs, 2% agarose gel electrophoresis results showing PCR products of amplicon size=697 bp(A120T) and 826bp (F167S) of diseased samples are shown in **Figure 19**. Fishers exact test values as well as the relative risk is mentioned in

Table *17*. Moreover, a graph for the comparison of frequency of both the mutations in control and diseased group is shown in **Figure 20**. There was no significant association observed between these two SNPs and T2DM patients according to statistical analysis.

Fishers exact test was performed to check the association of rs369262127 and rs201966419 with T2DM in using GraphPad Prism version 10. The p value thus obtained is 0.1662 and 0.1907 respectively for these two SNPs. The Relative risk for rs201966419 is 3 whereas for rs369262127 is 0.2955.

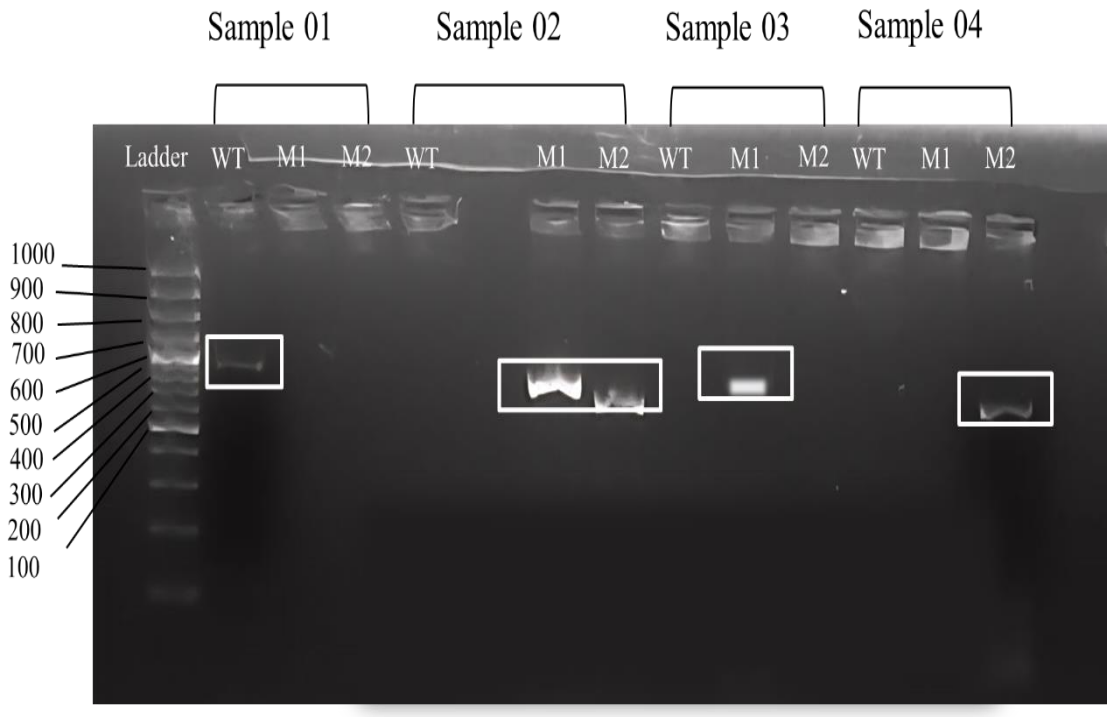


Figure 19: Gel electrophoresis results. 2% agarose gel electrophoresis showing PCR products for diseased samples for rs369262127 and rs201966419 of CAV1 gene. A ladder of 1000bp was used for comparison.

Table 17: Results of the Fisher's exact test. The table shows p-value, statistical significance and relative risk. There was no significant association observed between these two SNPs and T2DM patients according to statistical analysis.

Genotype	Test Performed	P value		Statistical significance	Relative risk
Variation 1 (F167S)	Fisher’s exact test	0.1907		ns	3
Variation 2 (A120T)	Fisher’s exact test	0.1662		ns	0.2955

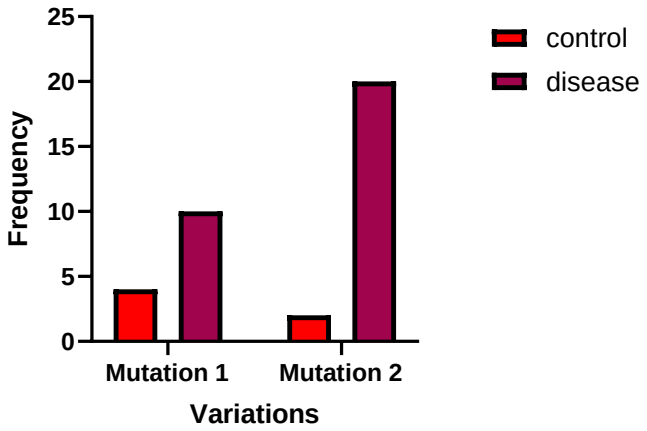


Figure 20: Bar graph of mutations. Graph showing comparison of frequency of occurrence of SNPs in individuals in control group and in disease group. Mutation 1 and 2 occurred in greater frequency in diseased individuals than in individuals in control group.

CHAPTER 5

DISCUSSION

Understanding protein structures is crucial for comprehending their functions and mechanisms. It helps identify drug-binding sites and predict interactions with compounds of interest. Bioinformatics tools analyze large genomic and proteomic datasets, aiding in discovering novel proteins and their roles, bridging genomics and functional biology. To ensure the accuracy of results, we employed multiple bioinformatics tools and based our analysis on their compiled results.

Studies on Caveolin 1 have revealed its role in regulating cellular signaling, cholesterol homeostasis, and vesicle trafficking. Importantly, mutations or dysregulation of *CAV1* are associated with a range of diseases, including cancer, cardiovascular disorders, and pulmonary diseases **Error! Bookmark not defined.**. Our task involved identifying missense SNPs within the *CAV1* gene that could contribute to T2DM.

The gene variant table for *CAV1* protein on Ensemble returned 77,626 SNPs, out of which 747 were missense SNPs. We began our analysis by subjecting the obtained non-synonymous single nucleotide polymorphisms (nsSNPs) to SIFT, SNAP2, and SNPs&GO analysis, revealing six nsSNPs that were consistently classified as deleterious by all three software. MutPred2 predicted changes in transmembrane protein structure for all nsSNPs and altered ordered interfaces for all except S80G. The F167S mutation stood out as it resulted in several significant effects, including the loss of an allosteric site and altered DNA binding. MUpro, a protein stability assessment tool, revealed all nsSNPs to reduce the structural stability of Caveolin 1. T95M and F167S were the most destabilizing mutations. Using ConSurf, we analyzed the conservation profile of *CAV1* amino acids. The more conserved regions indicate a greater role in protein function 56. Buried residues are involved in maintaining protein conformation, whereas exposed residues participate in ligand binding and protein interactions 57. ConSurf showed that S80G and A120T were highly conserved and buried within the protein structure. Conversely, R101H was found to be exposed, indicating greater accessibility in the protein's structure.

We utilized I-TASSER and Phyre2 to predict the 3D protein structure of wildtype and mutant Caveolin 1 and the RMSD values for the top structures determined through TM-Align. RMSD is a reliable method to observe differences between similar structures. The higher the RMSD value, the greater the variability between wildtype and mutant structure 57. The results indicate some deviation of structure from the wildtype.

Among post-translational modifications, phosphorylation was the only one identified. Phosphorylation is crucial in protein activation, signal transduction, and the interaction between proteins. Serine, threonine, and tyrosine are primary residues in protein phosphorylation 57. NetPhos3.1 pinpointed potential phosphorylation sites, emphasizing serine at position 80 as a target for CKI and threonine at position 95 as a target for PKC. The S80G and T95M substitutions resulted in the loss of these phosphorylation sites.

GeneMANIA and STRING analyses have identified a set of shared proteins associated with the *CAV1* gene, which include *CAV2, NOS3, EGFR, CAVIN1,* and *BSG*. Notably, *CAV2* has been demonstrated to form hetero oligomers with *CAV1* within lipid membranes and exhibits a robust physical interaction with *CAV1* 57 Mutations, T95M, A120T, R101H, and F167S, which disrupt the interface order of the *CAV1* protein, could interfere with its interaction with other proteins, potentially leading to diabetes. These SNPs have not been investigated in insilico or clinical studies therefore, it is essential to include these in further research on *CAV1*.

However, biological systems are incredibly complex. Bioinformatics tools often make simplifications and assumptions that may not fully capture the intricacies of biological processes. Predictive models generated by bioinformatics tools may have limitations in terms of accuracy, particularly for multifactorial and dynamic biological phenomena. Hence bioinformatics predictions and findings often require experimental validation to ensure reliability and accuracy.

For invitro analysis, 2 SNPs out of six were shortlisted based on the results that were received from ClinVar database. All the SNPs had uncertain significance. However, for rs369262127 the frequency came out to be 0.04% in population databases and has an allele count higher than expected for a pathogenic variant and it also has some association with pulmonary hypertension. The variant rs201966419 had conflicting classification of pathogenicity and was reported to be likely benign. So, primers were designed for these two SNPs using Primer 3 plus. The validation was done with Eurofin and UCSC genome browser since both are reliable tools. The amplicon size for variation 1 came out to be 695bp and for variation 2 it was 826bp.

In in-vitro analysis, a case control study was carried out to check the role of two selected SNPs i.e., rs369262127 and rs201966419 in T2DM in Pakistani population. Gradient PCR was performed to screen the samples, collected from various hospitals of Islamabad in comparison with healthy controls for the presence of polymorphism. 2% agarose gel electrophoresis was performed to analyze amplified products. The results were then statistically analyzed using GraphPad Prism version 10 for any significant association with T2DM. For this Fisher's exact test was used because the sample size was small, and this test gives accurate probability value rather than an estimation. The p value came out to be greater than 0.05 for both the SNPs because no significant difference was observed in frequency of occurrence of these SNPs in controls and in diseased T2DM individuals in Pakistani population, suggesting lack of evidence to reject the null hypothesis. For variation 1 the relative risk came out to be 3, which means that there is threefold increase in risk of getting the disease (T2DM) however with probability value being greater than 0.05 implies that despite the increase in risk, the variation is still statistically non-significant. And for variation 2 the relative risk value less than 1 suggests that there is decrease in risk of getting T2DM in individuals with this particular mutation.

Thus, the statistical analysis showed that there was no significant association of the two selected SNPs with risk of T2DM within Pakistani population. These results were contrary to what was predicted in insilico studies as all the 6 SNPs were predicted to be deleterious. One reason for no association might be the small sample size so there is a need for determination of these identified SNPs in large sample size to obtain large data size in order to get a better statistical analysis. Another reason could be that these SNPs though had no association with T2DM might be associated with metabolic syndrome, CVDs or even cancer but this of course needs to be validated through experimentation and large-scale studies.

CHAPTER 6

CONCLUSION

In 2021, Pakistan saw the tragic loss of nearly 396,625 lives to Type II Diabetes Mellitus (T2DM), a complex disorder that affects multiple systems in the body. Among the various factors contributing to T2DM, genetics plays a significant role. Out of the 79 genes identified in association with T2DM, one gene, Caveolin 1 (CAV1), stands out due to its involvement in insulin secretion and signaling. Despite this, research specifically exploring genetic mutations in the CAV1 gene in the Pakistani population remains sparse. To bridge this gap, a study was undertaken to investigate the association of Caveolin 1 with T2DM using both in-silico and in-vitro methods. The research focused on six specific single nucleotide polymorphisms (SNPs) within the CAV1 gene, chosen for their predicted harmful effects on the protein's function, as verified by SNPs and GO databases. The study revealed that certain mutations (T95M and F167S) significantly destabilized the structure and function of the Caveolin 1 protein. Detailed analysis showed that the A120T mutation led to structural changes, while the S80G and T95M mutations resulted in the loss of crucial phosphorylation sites, potentially disrupting insulin signaling pathways. Furthermore, these mutations could interfere with the interaction between CAV1 and other proteins such as Caveolin 2 (CAV2). Among the SNPs studied, A120T and F167S were particularly noteworthy; they were not only predicted to be deleterious but also showed uncertain significance in the ClinVar database. Hence, these nsSNPs can be helpful in understanding the main cause of insulin signaling and secretion disruption in T2DM. These deleterious SNPs can be involved in the risk of T2DM and pathogenesis of other diseases as well and can be used as potential molecular markers for the diagnosis and personalized treatment of T2DM. In order to validate this, the primers of these two mutations were run in gradient PCR with the DNA extracted from patients of diabetes type 2, and in some samples one out of two mutations was observed and, in some samples, both these mutations were shown. This study demonstrates no significant association of rs369262127 (a missense SNP ofCAV1) and rs201966419 with T2DM in our studied samples. This result was not consistent with our in-silico findings which were predicted to be deleterious. The present results need to be further validated by replication of this research in larger sample population. Moreover, the in-silico analysis predicted additional polymorphisms located in exonic region of CAV1 having disease causing potential as they can affect the normal structure and function of the caveolin 1 protein.

CHAPTER 7

FUTURE RECOMMENDATIONS

Performing a comprehensive analysis of single nucleotide polymorphisms (SNPs) in a large sample size is crucial for obtaining robust data sets that can yield meaningful statistical insights. By having a sizable cohort, researchers can uncover subtle genetic variations that may influence disease susceptibility, treatment response, and overall health outcomes. This abundance of data not only enhances the reliability of statistical analyses but also provides a broader understanding of the genetic landscape associated with the condition under investigation.

Moreover, large-scale studies offer invaluable resources for the development of in-silico drug targets. By identifying SNPs that correlate with disease progression or treatment efficacy, researchers can pinpoint potential therapeutic targets within the genome. This approach facilitates the design of tailored interventions that address the specific genetic factors contributing to the disease phenotype. Through computational modeling and simulation, in-silico drug discovery can expedite the identification of promising candidates for further experimental validation, ultimately accelerating the drug development process.

In parallel, integrating expressional analysis with SNP data enables a comprehensive assessment of gene expression patterns and their relationship to disease states. By comparing the expression levels of the CAV1 gene between experimental and control groups, researchers can elucidate the impact of genetic variations on molecular pathways underlying the disease. This comparative analysis provides insights into the functional consequences of SNP-associated changes in gene expression, shedding light on potential mechanisms driving disease pathology.

Overall, the integration of large-scale SNP analysis, in-silico drug target identification, and expressional analysis offers a multifaceted approach to understanding the genetic basis of disease and designing effective therapeutic strategies. By harnessing the power of big data and computational techniques, researchers can uncover novel insights that pave the way for precision medicine and personalized healthcare interventions.

REFERENCES

1. Brunton, S. (2016). Pathophysiology of Type 2 Diabetes: The Evolution of Our Understanding. The Journal of Family Practice, 65(4 Suppl). https://pubmed.ncbi.nlm.nih.gov/27262256/
2. Lee, S.-H., Park, S.-Y., & Choi, C. S. (2021). Insulin Resistance: from Mechanisms to Therapeutic Strategies. Diabetes & Metabolism Journal, 46(1). https://doi.org/10.4093/dmj.2021.0280
3. Basu, R. (2017, May). Type 2 Diabetes. National Institute of Diabetes and Digestive and Kidney Diseases. https://www.niddk.nih.gov/health-information/diabetes/overview/what-is-diabetes/type-2-diabetes
4. Azeem, S., Khan, U., & Liaquat, A. (2022). The increasing rate of diabetes in Pakistan: A silent killer. Annals of Medicine and Surgery, 79, 103901. https://doi.org/10.1016/j.amsu.2022.103901
5. International Diabetes Federation. (2023). Diabetes Facts & figures. International Diabetes Federation; International Diabetes Federation. https://idf.org/about-diabetes/diabetes-facts-figures/
6. International Diabetes Federation. (2021). Pakistan Diabetes report 2000 — 2045. In IDF Diabetes Atlas. https://diabetesatlas.org/data/en/country/150/pk.html
7. World Health Organization. (2023). Diabetes. World Health Organization; World Health Organization. https://www.who.int/health-topics/diabetes#tab=tab_1
8. Raza, A. (2017). Diabetes symptoms. In Pacific Medical Training. https://pacificmedicalacls.com/acls-library-guide-to-diabetes
9. CDC. (2022, June 20). Diabetes and Your Heart. Centers for Disease Control and Prevention. https://www.cdc.gov/diabetes/library/features/diabetes-and-heart.html
10. National Institute of Diabetes and Digestive and Kidney Diseases. (2019, March 25). Diabetic Eye Disease | NIDDK. National Institute of Diabetes and Digestive and Kidney Diseases. https://www.niddk.nih.gov/health-information/diabetes/overview/preventing-problems/diabetic-eye-disease
11. Centers for Disease Control and Prevention. (2021). Diabetes and Chronic Kidney Disease. Centers for Disease Control and Prevention. https://www.cdc.gov/diabetes/managing/diabetes-kidney-disease.html
12. Brazier, Y. (2019, March 19). Diabetic neuropathy: Types, symptoms, and causes. Www.medicalnewstoday.com. https://www.medicalnewstoday.com/articles/245310
13. Slater, R. (2022). Potential complications of type 2 diabetes. In www.medicalnewstoday.com. https://www.medicalnewstoday.com/articles/diabetes-mellitus-type-2-complications#kidney-disease
14. Haines, L., Wan, K. C., Lynn, R., Barrett, T. G., & Shield, J. P. H. (2007). Rising Incidence of Type 2 Diabetes in Children in the U.K. Diabetes Care, 30(5), 1097–1101. https://doi.org/10.2337/dc06-1813
15. Galicia-Garcia, U. (2020). Pathophysiology of Type 2 Diabetes Mellitus. International Journal of Molecular Sciences, 21(17), 1–34. https://doi.org/10.3390/ijms21176275
16. Watanabe, R. M., Valle, T., Hauser, E. R., Ghosh, S., Eriksson, J., Kohtamäki, K., Ehnholm, C., Tuomilehto, J., Collins, F. S., Bergman, R. N., & Boehnke, M. (1999). Familiality of Quantitative Metabolic Traits in Finnish Families with Non-Insulin-Dependent Diabetes mellitus. Human Heredity, 49(3), 159–168. https://doi.org/10.1159/000022865
17. McCarthy, M. I. (2010). Genomics, Type 2 Diabetes, and Obesity. New England Journal of Medicine, 363(24), 2339–2350. https://doi.org/10.1056/nejmra0906948

18. Mambiya, M., Shang, M., Wang, Y., Li, Q., Liu, S., Yang, L., Zhang, Q., Zhang, K., Liu, M., Nie, F., Zeng, F., & Liu, W. (2019). The Play of Genes and Non-genetic Factors on Type 2 Diabetes. Frontiers in Public Health, 7(7). https://doi.org/10.3389/fpubh.2019.00349
19. Muller, D. C., Elahi, D., Tobin, J. D., & Andres, R. (1996). Insulin response during the oral glucose tolerance test: The role of age, sex, body fat and the pattern of fat distribution. Aging Clinical and Experimental Research, 8(1), 13–21. https://doi.org/10.1007/bf03340110
20. Hillier, T. A., & Pedula, K. L. (2003). Complications in Young Adults With Early-Onset Type 2 Diabetes: Losing the relative protection of youth. Diabetes Care, 26(11), 2999–3005. https://doi.org/10.2337/diacare.26.11.2999
21. Mishra, S. (2013). Effect of Physical activity on Insulin Resistance, Inflammation and Oxidative Stress in Diabetes Mellitus. JOURNAL of CLINICAL and DIAGNOSTIC RESEARCH, 7(8). https://doi.org/10.7860/jcdr/2013/6518.3306
22. American Diabetes Association. (2018). 3. Prevention or Delay of Type 2 Diabetes: Standards of Medical Care in Diabetes—2019. Diabetes Care, 42(Supplement 1), S29–S33. https://doi.org/10.2337/dc19-s003
23. Rossetti, L., Shulman, G. I., W. Zawalich, & Defronzo, R. A. (1987). Effect of chronic hyperglycemia on in vivo insulin secretion in partially pancreatectomized rats. Journal of Clinical Investigation, 80(4), 1037–1044. https://doi.org/10.1172/jci113157
24. DeFronzo, R. A. (2009). From the Triumvirate to the Ominous Octet: A New Paradigm for the Treatment of Type 2 Diabetes Mellitus. Diabetes, 58(4), 773–795. https://doi.org/10.2337/db09-9028
25. Sircana, A., Framarin, L., Leone, N., Berrutti, M., Castellino, F., Parente, R., De Michieli, F., Paschetta, E., & Musso, G. (2018). Altered Gut Microbiota in Type 2 Diabetes: Just a Coincidence? Current Diabetes Reports, 18(10). https://doi.org/10.1007/s11892-018-1057-6
26. Cani, P. D., Amar, J., Iglesias, M. A., Poggi, M., Knauf, C., Bastelica, D., Neyrinck, A. M., Fava, F., Tuohy, K. M., Chabo, C., Waget, A., Delmee, E., Cousin, B., Sulpice, T., Chamontin, B., Ferrieres, J., Tanti, J.-F. ., Gibson, G. R., Casteilla, L., & Delzenne, N. M. (2007). Metabolic Endotoxemia Initiates Obesity and Insulin Resistance. Diabetes, 56(7), 1761–1772. https://doi.org/10.2337/db06-1491
27. Shan, Z., Sun, T., Huang, H., Chen, S., Chen, L., Luo, C., Yang, W., Yang, X., Yao, P., Cheng, J., Hu, F. B., & Liu, L. (2017). Association between microbiota-dependent metabolite
28. trimethylamine-N-oxide and type 2 diabetes. The American Journal of Clinical Nutrition, 106(3), 888–894. https://doi.org/10.3945/ajcn.117.157107
29. Shigenaga, M. K., Hagen, T. M., & Ames, B. N. (1994). Oxidative damage and mitochondrial decay in aging. Proceedings of the National Academy of Sciences, 91(23), 10771–10778. https://doi.org/10.1073/pnas.91.23.10771
30. Sergi, D., Naumovski, N., Heilbronn, L. K., Abeywardena, M., O'Callaghan, N., Lionetti, L., & Luscombe-Marsh, N. (2019). Mitochondrial (Dys)function and Insulin Resistance: From Pathophysiological Molecular Mechanisms to the Impact of Diet. Frontiers in Physiology, 10. https://doi.org/10.3389/fphys.2019.00532
31. Wei, H., Liu, L., & Chen, Q. (2015). Selective removal of mitochondria via mitophagy: distinct pathways for different mitochondrial stresses. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 1853(10), 2784–2790. https://doi.org/10.1016/j.bbamcr.2015.03.013

32. Kim, J., Wei, Y., & Sowers, J. R. (2008). Role of Mitochondrial Dysfunction in Insulin Resistance. Circulation Research, 102(4), 401–414. https://doi.org/10.1161/circresaha.107.165472
33. Valaiyapathi, B., Gower, B., & Ashraf, A. P. (2018). Pathophysiology of Type 2 Diabetes in Children and Adolescents. Current Diabetes Reviews, 14. https://doi.org/10.2174/1573399814666180608074510
34. F. Otero, Y., M. Stafford, J., & P. McGuinness, O. (2014). Pathway-selective Insulin Resistance and Metabolic Disease: The Importance of Nutrient Flux. Journal of Biological Chemistry, 289(30), P20462-20469. https://doi.org/10.1074/jbc.R114.576355
35. Krssak, M., Brehm, A., Bernroider, E., Anderwald, C., Nowotny, P., Man, C. D., Cobelli, C., Cline, G. W., Shulman, G. I., Waldhausl, W., & Roden, M. (2004). Alterations in Postprandial Hepatic Glycogen Metabolism in Type 2 Diabetes. Diabetes, 53(12), 3048–3056. https://doi.org/10.2337/diabetes.53.12.3048
36. Jeng, C.-Y., Sheu, W. H-H., Fuh, M. M-T., Chen, Y-D. I., & Reaven, G. M. (1994). Relationship Between Hepatic Glucose Production and Fasting Plasma Glucose Concentration in Patients With NIDDM. Diabetes, 43(12), 1440–1444. https://doi.org/10.2337/diab.43.12.1440
37. Kim, J. K., Zisman, A., Fillmore, J. J., Peroni, O. D., Kotani, K., Perret, P., Zong, H., Dong, J., Kahn, C. R., Kahn, B. B., & Shulman, G. I. (2001). Glucose toxicity and the development of diabetes in mice with muscle-specific inactivation of GLUT4. Journal of Clinical Investigation, 108(1), 153–160. https://doi.org/10.1172/jci10294
38. Fonseca, V. A. (2009). Defining and Characterizing the Progression of Type 2 Diabetes. Diabetes Care, 32(suppl_2), S151–S156. https://doi.org/10.2337/dc09-s301
39. Diabetes Prevention Program Research Group. (2007). The prevalence of retinopathy in impaired glucose tolerance and recent-onset diabetes in the Diabetes Prevention Program. Diabetic Medicine, 24(2), 137–144. https://doi.org/10.1111/j.1464-5491.2007.02043.x
40. Nichols, G. A., Hillier, T. A., & Brown, J. B. (2007). Progression From Newly Acquired Impaired Fasting Glucose to Type 2 Diabetes. Diabetes Care, 30(2), 228–233. https://doi.org/10.2337/dc06-1392
41. Weyer, C., Tataranni, P. A., Bogardus, C., & Pratley, R. E. (2001). Insulin Resistance and Insulin Secretory Dysfunction Are Independent Predictors of Worsening of Glucose Tolerance During Each Stage of Type 2 Diabetes Development. Diabetes Care, 24(1), 89–94. https://doi.org/10.2337/diacare.24.1.89
42. Pani, L. N., Nathan, D. M., & Grant, R. W. (2007). Clinical Predictors of Disease Progression and Medication Initiation in Untreated Patients With Type 2 Diabetes and A1C Less Than 7%. Diabetes Care, 31(3), 386–390. https://doi.org/10.2337/dc07-1934
43. UK Prospective Diabetes Study (UKPDS) Group. (1998a). Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). The Lancet, 352(9131), 854–865. https://doi.org/10.1016/S0140-6736(98)07037-8
44. UK Prospective Diabetes Study (UKPDS) Group. (1998b). Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). The Lancet, 352(9131), 837–853. https://doi.org/10.1016/S0140-6736(98)07019-6
45. U.K. prospective diabetes study 16. Overview of 6 years' therapy of type II diabetes: a progressive disease. U.K. Prospective Diabetes Study Group. (1995). Diabetes, 44(11), 1249–1258. https://pubmed.ncbi.nlm.nih.gov/7589820/

46. Olokoba, A. B., Obateru, O. A., & Olokoba, L. B. (2012). Type 2 Diabetes Mellitus: A Review of Current Trends. Oman Medical Journal, 27(4), 269–273. https://doi.org/10.5001/omj.2012.68
47. Cox, M. E., & Edelman, D. (2009). Tests for Screening and Diagnosis of Type 2 Diabetes. Clinical Diabetes, 27(4), 132–138. https://doi.org/10.2337/diaclin.27.4.132
48. International Expert Committee. (2009). International Expert Committee Report on the Role of the A1C Assay in the Diagnosis of Diabetes. Diabetes Care, 32(7), 1327–1334. https://doi.org/10.2337/dc09-9033
49. Knowler, W. C., Barrett-Connor, E., Fowler, S. E., Hamman, R. F., Lachin, J. M., Walker, E. A., & Nathan, D. M. (2002). Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. The New England Journal of Medicine, 346(6), 393–403. https://pubmed.ncbi.nlm.nih.gov/11832527/
50. Katula, J. A., Vitolins, M. Z., Rosenberger, E. L., Blackwell, C. S., Morgan, T. M., Lawlor, M. S., & Goff, D. C. (2011). One-Year Results of a Community-Based Translation of the Diabetes Prevention Program: Healthy-Living Partnerships to Prevent Diabetes (HELP PD) Project. Diabetes Care, 34(7), 1451–1457. https://doi.org/10.2337/dc10-2115
51. Fradkin, J. E., & Rodgers, G. P. (2013). Diabetes Research: A Perspective From the National Institute of Diabetes and Digestive and Kidney Diseases. Diabetes, 62(2), 320–326. https://doi.org/10.2337/db12-0269
52. Rydén, L., & Lindsten, J. (2021). The history of the Nobel prize for the discovery of insulin. Diabetes Research and Clinical Practice, 175, 108819. https://doi.org/10.1016/j.diabres.2021.108819
53. Tahrani, A. A., Barnett, A. H., & Bailey, C. J. (2016). Pharmacology and therapeutic implications of current drugs for type 2 diabetes mellitus. Nature Reviews Endocrinology, 12(10), 566–592. https://doi.org/10.1038/nrendo.2016.86
54. Chiniwala, N., & Jabbour, S. (2011). Management of diabetes mellitus in the elderly. Current Opinion in Endocrinology, Diabetes and Obesity, 18(2), 148–152. https://doi.org/10.1097/med.0b013e3283444ba0
55. Collier, C. A., Bruce, C. R., Smith, A. C., Lopaschuk, G., & Dyck, D. J. (2006). Metformin counters the insulin-induced suppression of fatty acid oxidation and stimulation of triacylglycerol storage in rodent skeletal muscle. American Journal of Physiology-Endocrinology and Metabolism, 291(1), E182–E189. https://doi.org/10.1152/ajpendo.00272.2005
56. Fuhlendorff, J., Rorsman, P., Kofod, H., Brand, C. L., Rolin, B., MacKay, P., Shymko, R., & Carr, R. D. (1998). Stimulation of insulin release by repaglinide and glibenclamide involves both common and distinct processes. Diabetes, 47(3), 345–351. https://doi.org/10.2337/diabetes.47.3.345
57. Yki-Järvinen, H. (2004). Thiazolidinediones. New England Journal of Medicine, 351(11), 1106–1118. https://doi.org/10.1056/nejmra041001
58. Stonehouse, A. H., Darsow, T., & Maggs, D. G. (2012). Incretin-based therapies. Journal of Diabetes, 4(1), 55–67. https://doi.org/10.1111/j.1753-0407.2011.00143.x
59. Burge, M. R., & Schade, D. S. (1997). INSULINS. Endocrinology and Metabolism Clinics of North America, 26(3), 575–598. https://doi.org/10.1016/S0889-8529(05)70268-1
60. Pratley, R. E., & Salsali, A. (2007). Inhibition of DPP-4: a new therapeutic approach for the treatment of type 2 diabetes. Current Medical Research and Opinion, 23(4), 919–931. https://doi.org/10.1185/030079906x162746

61. Mikhail, N. (2011). Quick-Release Bromocriptine for Treatment of Type 2 Diabetes. Current Drug Delivery, 8(5), 511–516. https://doi.org/10.2174/156720111796642255
62. American Diabetes Association. (2024). Precision Medicine in Diabetes Initiative | American Diabetes Association. Professional.diabetes.org. https://professional.diabetes.org/clinical-support/precision-medicine-diabetes-initiative-0
63. Stumvoll, M., Goldstein, B. J., & van Haeften, T. W. (2005). Type 2 diabetes: principles of pathogenesis and therapy. The Lancet, 365(9467), 1333–1346. https://doi.org/10.1016/s0140-6736(05)61032-x
64. Li, S., Couet, J., & Lisanti, M. P. (1996). Src Tyrosine Kinases, GαSubunits, and H-Ras Share a Common Membrane-anchored Scaffolding Protein, Caveolin. Journal of Biological Chemistry, 271(46), 29182–29190. https://doi.org/10.1074/jbc.271.46.29182
65. Fujimoto, T., Hiroshi Kogo, Nomura, R., & Une, T. (2000). Isoforms of caveolin-1 and caveolar structure. Journal of Cell Science, 113(19), 3509–3517. https://doi.org/10.1242/jcs.113.19.3509
66. Rothberg, K. G., Heuser, J. E., Donzell, W. C., Ying, Y., Glenney, J. R., & Anderson, R. A. (1992). Caveolin, a protein component of caveolae membrane coats. Cell, 68(4), 673–682. https://doi.org/10.1016/0092-8674(92)90143-z
67. Haddad, D., Al Madhoun, A., Nizam, R., & Al-Mulla, F. (2020). Role of Caveolin-1 in Diabetes and Its Complications. Oxidative Medicine and Cellular Longevity, 2020, 9761539. https://doi.org/10.1155/2020/9761539
68. Liu, P., Rudick, M., & Anderson, R. G. W. (2002). Multiple Functions of Caveolin-1. Journal of Biological Chemistry, 277(44), 41295–41298. https://doi.org/10.1074/jbc.R200020200
69. Bastiani, M., Liu, L., Hill, M. M., Jedrychowski, M. P., Nixon, S. J., Lo, H. P., Abankwa, D., Luetterforst, R., Fernandez-Rojo, M., Breen, M. R., Gygi, S. P., Vinten, J., Walser, P. J., North, K. N., Hancock, J. F., Pilch, P. F., & Parton, R. G. (2009). MURC/Cavin-4 and cavin family members form tissue-specific caveolar complexes. Journal of Cell Biology, 185(7), 1259–1273. https://doi.org/10.1083/jcb.200903053
70. Song, K. S., Tang, Z., Li, S., & Lisanti, M. P. (1997). Mutational Analysis of the Properties of Caveolin-1. Journal of Biological Chemistry, 272(7), 4398–4403. https://doi.org/10.1074/jbc.272.7.4398
71. Kim, C., Delepine, M., Boutet, E., Haquima El Mourabit, Soazig Le Lay, Meier, M., Nemani, M., E. Bridel, Claudia, Débora Romeo Bertola, Semple, R. K., O'Rahilly, S., Dugail, I., Capeau, J., Lathrop, M., & Jocelyne Magré. (2008). Association of a Homozygous Nonsense Caveolin-1 Mutation with Berardinelli-Seip Congenital Lipodystrophy. The Journal of Clinical Endocrinology and Metabolism, 93(4), 1129–1134. https://doi.org/10.1210/jc.2007-1328
72. Karhan, A. N., Zammouri, J., Auclair, M., Capel, E., Apaydin, F. D., Ates, F., Verpont, M.-C., Magré, J., Fève, B., Lascols, O., Usta, Y., Jéru, I., & Vigouroux, C. (2021). Biallelic *CAV1* null variants induce congenital generalized lipodystrophy with achalasia. European Journal of Endocrinology, 185(6), 841–854. https://doi.org/10.1530/EJE-21-0915
73. Razani, B., Combs, T. P., Wang, X. B., Frank, P. G., Park, D. S., Russell, R. G., Li, M., Tang, B., Jelicks, L. A., Scherer, P. E., & Lisanti, M. P. (2002). Caveolin-1-deficient Mice Are Lean, Resistant to Diet-induced Obesity, and Show Hypertriglyceridemia with Adipocyte Abnormalities. Journal of Biological Chemistry, 277(10), 8635–8647. https://doi.org/10.1074/jbc.m110970200

74. Cao, H., Alston, L., Ruschman, J., & Hegele, R. A. (2008). Heterozygous *CAV1* frameshift mutations (MIM 601047) in patients with atypical partial lipodystrophy and hypertriglyceridemia. Lipids in Health and Disease, 7(1). https://doi.org/10.1186/1476-511x-7-3
75. Garg, A., Kircher, M., Miguel del Campo, R. Stephen Amato, & Agarwal, A. K. (2015). Whole exome sequencing identifies de novo heterozygous *CAV1* mutations associated with a novel neonatal onset lipodystrophy syndrome. American Journal of Medical Genetics - Part A, 167(8), 1796–1806. https://doi.org/10.1002/ajmg.a.37115
76. Richards, C. S., Bale, S., Bellissimo, D. B., Das, S., Grody, W. W. W., Hegde, M. R., Lyon, E., & Ward, B. E. (2008). ACMG recommendations for standards for interpretation and reporting
77. of sequence variations: Revisions 2007. Genetics in Medicine, 10(4), 294–300. https://doi.org/10.1097/GIM.0b013e31816b5cae
78. Machado, R. D., Southgate, L., Eichstaedt, C. A., Aldred, M. A., Austin, E. D., Best, D. H., Chung, W. K., Benjamin, N., Elliott, C. G., Eyries, M., Fischer, C., Gräf, S., Hinderhofer, K., Humbert, M., Keiles, S. B., Loyd, J. E., Morrell, N. W., Newman, J. H., Soubrier, F., & Trembath, R. C. (2015). Pulmonary Arterial Hypertension: A Current Perspective on Established and Emerging Molecular Genetic Defects. Human Mutation, 36(12), 1113–1127. https://doi.org/10.1002/humu.22904
79. Austin, E. D., Ma, L., LeDuc, C., Berman Rosenzweig, E., Borczuk, A., Phillips, J. A., Palomero, T., Sumazin, P., Kim, H. R., Talati, M. H., West, J., Loyd, J. E., & Chung, W. K. (2012). Whole Exome Sequencing to Identify a Novel Gene (Caveolin-1) Associated With Human Pulmonary Arterial Hypertension. Circulation: Cardiovascular Genetics, 5(3), 336–343. https://doi.org/10.1161/circgenetics.111.961888
80. Farwell, K. D., Shahmirzadi, L., El-Khechen, D., Powis, Z., Chao, E. C., Tippin Davis, B., Baxter, R. M., Zeng, W., Mroske, C., Parra, M. C., Gandomi, S. K., Lu, I., Li, X., Lu, H., Lu, H.-M., Salvador, D., Ruble, D., Lao, M., Fischbach, S., & Wen, J. (2014). Enhanced utility of family-centered diagnostic exome sequencing with inheritance model–based analysis: results from 500 unselected families with undiagnosed genetic conditions. Genetics in Medicine, 17(7), 578–586. https://doi.org/10.1038/gim.2014.154
81. Cohen, A. W., Razani, B., Wang, X. B., Combs, T. P., Williams, T. M., Scherer, P. E., & Lisanti, M. P. (2003). Caveolin-1-deficient mice show insulin resistance and defective insulin receptor protein expression in adipose tissue. American Journal of Physiology-Cell Physiology, 285(1), C222–C235. https://doi.org/10.1152/ajpcell.00006.2003
82. Otsu, K., Toya, Y., Oshikawa, J., Kurotani, R., Yazawa, T., Sato, M., Yokoyama, U., Umemura, S., Minamisawa, S., Okumura, S., & Ishikawa, Y. (2010). Caveolin gene transfer improves glucose metabolism in diabetic mice. American Journal of Physiology. Cell Physiology, 298(3), C450-456. https://doi.org/10.1152/ajpcell.00077.2009
83. Cohen, A. W., Combs, T. P., Scherer, P. E., & Lisanti, M. P. (2003). Role of caveolin and caveolae in insulin signaling and diabetes. American Journal of Physiology-Endocrinology and Metabolism, 285(6), E1151–E1160. https://doi.org/10.1152/ajpendo.00324.2003
84. Peng, H., Mu, P., Li, H., Lin, S., Lin, C., Lin, K., Liu, K., Zeng, W., & Zeng, L. (2021). Caveolin-1 Is Essential for the Improvement of Insulin Sensitivity through AKT Activation during Glargine Treatment on Diabetic Mice. Journal of Diabetes Research. https://doi.org/10.1155/2021/9943344

85. Catalán, V., Gómez-Ambrosi, J., Rodríguez, A., Silva, C., Rotellar, F., Gil, M. J., Cienfuegos, J. A., Salvador, J., & Frühbeck, G. (2007). Expression of caveolin-1 in human adipose tissue is upregulated in obesity and obesity-associated type 2 diabetes mellitus and related to inflammation. Clinical Endocrinology, 0(0). https://doi.org/10.1111/j.1365-2265.2007.03021.x
86. Frühbeck, G., López, M., & Diéguez, C. (2007). Role of caveolins in body weight and insulin resistance regulation. Trends in Endocrinology & Metabolism, 18(5), 177–182. https://doi.org/10.1016/j.tem.2007.04.001
87. Wang, H., Wang, A. X., & Barrett, E. J. (2011). Caveolin-1 is required for vascular endothelial insulin uptake. American Journal of Physiology-Endocrinology and Metabolism, 300(1), E134–E144. https://doi.org/10.1152/ajpendo.00498.2010
88. de Souza, G. M., de Albuquerque Borborema, M. E., de Lucena, T. M. C., da Silva Santos, A. F., de Lima, B. R., de Oliveira, D. C., & de Azevêdo Silva, J. (2020a). Caveolin-1 (CAV-1) up regulation in metabolic syndrome: all roads leading to the same end. Molecular Biology Reports, 47(11), 9245–9250. https://doi.org/10.1007/s11033-020-05945-y
89. Mora-García, G., Gómez-Camargo, D., Alario, Á., & Gómez-Alegría, C. (2018b). A Common Variation in the Caveolin 1 Gene Is Associated with High Serum Triglycerides and Metabolic Syndrome in an Admixed Latin American Population. Metabolic Syndrome and Related Disorders, 16(9), 453–463. https://doi.org/10.1089/met.2018.0004
90. Pojoga, L. H., Underwood, P. C., Goodarzi, M. O., Williams, J. S., Adler, G. K., Jeunemaitre, X., Hopkins, P. N., Raby, B. A., Lasky-Su, J., Sun, B., Cui, J., Guo, X., Taylor, K. D., Chen, Y.-D. I., Xiang, A., Raffel, L. J., Buchanan, T. A., Rotter, J. I., & Williams, G. H. (2011a). Variants of the Caveolin-1 Gene: A Translational Investigation Linking Insulin Resistance and Hypertension. The Journal of Clinical Endocrinology & Metabolism, 96(8), E1288–E1292. https://doi.org/10.1210/jc.2010-2738
91. Baudrand, R., Goodarzi, M. O., Vaidya, A., Underwood, P. C., Williams, J. S., Jeunemaitre, X., Hopkins, P. N., Brown, N., Raby, B. A., Lasky-Su, J., Adler, G. K., Cui, J., Guo, X., Taylor, K. D., Chen, Y.-D. I., Xiang, A., Raffel, L. J., Buchanan, T. A., Rotter, J. I., & Williams, G. H. (2015b). A prevalent caveolin-1 gene variant is associated with the metabolic syndrome in Caucasians and Hispanics. Metabolism, 64(12), 1674–1681. https://doi.org/10.1016/j.metabol.2015.09.005
92. Faezeh Abaj, Ghafour, A., & Khadijeh Mirzaei. (2021b). Mediation role of body fat distribution (FD) on the relationship between *CAV1* rs3807992 polymorphism and metabolic syndrome in overweight and obese women. BMC Medical Genomics, 14(1). https://doi.org/10.1186/s12920-021-01050-6
93. Al Madhoun, A., Haddad, D., Nizam, R., Miranda, L., Kochumon, S., Thomas, R., Thanaraj, T. A., Ahmad, R., Bitar, M. S., & Al-Mulla, F. (2022a). Caveolin-1 rs1997623 Single Nucleotide Polymorphism Creates a New Binding Site for the Early B-Cell Factor 1 That Instigates Adipose Tissue *CAV1* Protein Overexpression. Cells, 11(23), 3937. https://doi.org/10.3390/cells11233937
94. Ma, C.-X., Ma, X.-N., Guan, C.-H., Li, Y.-D., Mauricio, D., & Fu, S.-B. (2022a). Cardiovascular disease in type 2 diabetes mellitus: progress toward personalized management. Cardiovascular Diabetology, 21(1). https://doi.org/10.1186/s12933-022-01516-6
95. İlikay, S., Coşkunpınar, E., Kurnaz-gömleksiz, Ö., Buğra, Z., Eronat, A. P., Öztürk, O., & Yılmaz-aydoğan, H. (2019a). Effects of common variations of NOS3 and *CAV1* genes on
96. hypercholesterolemic profile in coronary heart disease. İstanbul Journal of Pharmacy, 49(2), 53–60. https://dergipark.org.tr/en/pub/iujp/issue/48543/616297

97. Chen, S., Wang, X., Wang, J., Zhao, Y., Wang, D., Tan, C., Fa, J., Zhang, R., Wang, F., Xu, C., Huang, Y., Li, S., Yin, D., Xiong, X., Li, X., Chen, Q., Tu, X., Yang, Y., Xia, Y., & Xu, C. (2016b). Genomic variant in *CAV1* increases susceptibility to coronary artery disease and myocardial infarction. Atherosclerosis, 246, 148–156. https://doi.org/10.1016/j.atherosclerosis.2016.01.008
98. Ketteler, J., & Klein, D. (2018b). Caveolin-1, cancer and therapy resistance. International Journal of Cancer, 143(9), 2092–2104. https://doi.org/10.1002/ijc.31369
99. Chang, W.-S., Lin, S.-S., Li, F.-J., Tsai, C.-W., Li, L.-Y., Lien, C.-S., Liao, W.-L., Wu, H.-C., Tsai, C.-H., Shih, T.-C., & Bau, D.-T. (2013a). Significant Association of Caveolin-1 (*CAV1*) Genotypes with Upper Urothelial Tract Cancer. Anticancer Research, 33(11), 4907–4912. https://ar.iiarjournals.org/content/33/11/4907.short
100. Khan, N., Paterson, A. D., Roshandel, D., Maqbool, S., Fazal, N., Ali, L., Khurram, R., Ijaz Maqsood, S., Hafiza Benish Ali, S., Khan, H., Tariq, K., Bano, S., Azam, A., Muslim, I., Maria, M., Agha, Z., Babar, M., Khalida Waheed, N., Azam, M., & Qamar, R. (2023a). Role of 19 SNPs in 10 genes with type 2 diabetes in the Pakistani population. Gene, 848, 146899. https://doi.org/10.1016/j.gene.2022.146899
101. Yan, C., Sun, C., Ding, X., Rizeq, F. K., Ren, M., Yang, F., Chen, Y., & Wang, B. (2019a). Association of *CAV1* polymorphisms with the risks of breast cancer: A systematic review and meta-analysis. Pathology - Research and Practice, 215(9), 152518. https://doi.org/10.1016/j.prp.2019.152518
102. Zhang, J., Xue, F., Chen, S., Zhang, D., Lu, C., & Tang, G. (2017b). The influence of caveolin-1 gene polymorphisms on hepatitis B virus-related hepatocellular carcinoma susceptibility in Chinese Han population. Medicine, 96(42), e7359. https://doi.org/10.1097/md.0000000000007359
103. Zhao, X., Pan, G., Yuan, Q., Mu, D., Zhang, J., Cui, T., Zhang, J., & Zhang, L. (2014a). Genetic variations of *CAV1* gene contribute to HCC risk: a case–control study. Tumor Biology, 35(11), 11289–11293. https://doi.org/10.1007/s13277-014-2428-5
104. Fan, S., Meng, J., Zhang, L., Zhang, X., & Liang, C. (2019a). *CAV1* polymorphisms rs1049334, rs1049337, rs7804372 might be the potential risk in tumorigenicity of urinary cancer: A systematic review and meta-analysis. Pathology - Research and Practice, 215(1), 151–158. https://doi.org/10.1016/j.prp.2018.11.009
105. Wang, M., Tian, T., Ma, X., Zhu, W., Guo, Y., Duan, Z., Fan, J., Lin, S., Liu, K., Zheng, Y., Sheng, Q., Dai, Z.-J., & Peng, H. (2017a). Genetic polymorphisms in caveolin-1 associate with breast cancer risk in Chinese Han population. Oncotarget, 8(53), 91654–91661. https://doi.org/10.18632/oncotarget.21560
106. Garcia, G. J., Ruiz, M. S., Gomez, D. E., & Gomez, C. J. (2017a). Frequency of common polymorphisms in Caveolin 1 (*CAV1*) gene in adults with high serum triglycerides from
107. Colombian Caribbean Coast. Colombia Médica, 48(4), 167–173. https://doi.org/10.25100/cm.v48i4.2625
108. Chen, P., Zhang, Y.-L., Xue, B., & Wang, J.-R. (2021). *CAV1* rs7804372 (T29107A) polymorphism might be a potential risk for digestive cancers: A protocol for systematic review and meta analysis. Medicine, 100(24), e26186. https://doi.org/10.1097/MD.0000000000026186
109. Li, Q., Bagrodia, A., Cha, E. K., & Coleman, J. A. (2016b). Prognostic Genetic Signatures in Upper Tract Urothelial Carcinoma. Current Urology Reports, 17(2). https://doi.org/10.1007/s11934-015-0566-y

110. Tsou, Y.-A., Tsai, C.-W., Tsai, M.-H., Chang, W.-S., Li, F.-J., Liu, Y.-F., Chiu, C.-F., Lin, C.-C., & Bau, D.-T. (2011). Association of caveolin-1 genotypes with nasopharyngeal carcinoma susceptibility in Taiwan. Anticancer Research, 31(10), 3629–3632. https://pubmed.ncbi.nlm.nih.gov/21965789/

111. Wu, H.-C., Chang, C.-H., Tsou, Y.-A., Tsai, C.-W., Lin, C.-C., & Bau, D.-T. (2011). Significant association of caveolin-1 (*CAV1*) genotypes with prostate cancer susceptibility in Taiwan. Anticancer Research, 31(2), 745–749. https://pubmed.ncbi.nlm.nih.gov/21378366/

112. Luo, S., Yang, M., Zhao, H., Han, Y., Jiang, N., Yang, J., Chen, W., Li, C., Liu, Y., Zhao, C., & Sun, L. (2021). Caveolin-1 Regulates Cellular Metabolism: A Potential Therapeutic Target in Kidney Disease. Frontiers in Pharmacology, 12. https://doi.org/10.3389/fphar.2021.768100

113. Sim, N.-L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., & Ng, P. C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Research, 40(W1), W452–W457. https://doi.org/10.1093/nar/gks539

114. Calabrese, R., Capriotti, E., Fariselli, P., Martelli, P. L., & Casadio, R. (2009). Functional annotations improve the predictive score of human disease-related mutations in proteins. Human Mutation, 30(8), 1237–1244. https://doi.org/10.1002/humu.21047

115. Johnson, A. D., Handsaker, R. E., Pulit, S. L., Nizzari, M. M., O'Donnell, C. J., & I.W, P. (2008). SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics, 24(24), 2938–2939. https://doi.org/10.1093/bioinformatics/btn564

116. Pejaver, V., Urresti, J., Lugo-Martinez, J., Pagel, K. A., Lin, G. N., Nam, H. J., Mort, M., Cooper, D. N., Sebat, J., Iakoucheva, L. M., Mooney, S. D., & Radivojac, P. (2020). Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. Nature communications, 11(1), 5918. https://doi.org/10.1038/s41467-020-19669-x

117. Cheng, J., Randall, A., & Baldi, P. (2005). Prediction of protein stability changes for single-site mutations using support vector machines. Proteins: Structure, Function, and Bioinformatics, 62(4), 1125–1132. https://doi.org/10.1002/prot.20810

118. Ben Chorin, A., Masrati, G., Kessel, A., Narunsky, A., Sprinzak, J., Lahav, S., Ashkenazy, H., & Ben-Tal, N. (2019). ConSurf-DB: An accessible repository for the evolutionary conservation patterns of the majority of PDB proteins. Protein Science, 29(1), 258–267. https://doi.org/10.1002/pro.3779

119. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols, 10(6), 845–858. https://doi.org/10.1038/nprot.2015.053

120. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. Nature methods, 12(1), 7–8. https://doi.org/10.1038/nmeth.3213

121. Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nature protocols, 5(4), 725–738. https://doi.org/10.1038/nprot.2010.5

122. Zhang Y. (2008). I-TASSER server for protein 3D structure prediction. BMC bioinformatics, 9, 40. https://doi.org/10.1186/1471-2105-9-40

123. Ko, J., Park, H., Heo, L., & Seok, C. (2012, May 30). GalaxyWEB server for protein structure prediction and refinement. Oxford Academic; Nucleic Acids Research. https://academic.oup.com/nar/article/40/W1/W294/1078340

124. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--A visualization system for exploratory research and analysis. Journal of Computational Chemistry, 25(13), 1605–1612. https://doi.org/10.1002/jcc.20084
125. Zhang, Y., & Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. Nucleic acids research, 33(7), 2302–2309. https://doi.org/10.1093/nar/gki524
126. Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. Protein science: a publication of the Protein Society, 2(9), 1511–1519. https://doi.org/10.1002/pro.5560020916
127. Markus Wiederstein, Manfred J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, Nucleic Acids Research, Volume 35, Issue suppl_2, 1 July 2007, Pages W407–W410, https://doi.org/10.1093/nar/gkm290
128. Sippl, M. J. (1993). Recognition of errors in three-dimensional structures of proteins. Proteins: Structure, Function, and Genetics, 17(4), 355–362. https://doi.org/10.1002/prot.340170404
129. Deng, W., Wang, Y., Ma, L., Zhang, Y., Ullah, S., & Xue, Y. (2016). Computational prediction of methylation types of covalently modified lysine and arginine residues in proteins. Briefings in Bioinformatics, bbw041. https://doi.org/10.1093/bib/bbw041
130. Blom, N., Gammeltoft, S., & Brunak, S. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. Journal of Molecular Biology, 294(5), 1351–1362. https://doi.org/10.1006/jmbi.1999.3310
131. Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T-B. G., Lavrsen, K., Dabelsteen, S., Pedersen, N. B., Marcos-Silva, L., Gupta, R., Paul Bennett, E., Mandel, U., Brunak, S., Wandall, H. H., Levery, S. B., & Clausen, H.
132. (2013). Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. The EMBO Journal, 32(10), 1478–1488. https://doi.org/10.1038/emboj.2013.79
133. Gupta, R., & Brunak, S. (2002). Prediction of glycosylation across the human proteome and the correlation to protein function. Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing, 310–322. https://pubmed.ncbi.nlm.nih.gov/11928486/
134. Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C. T., Maitland, A., Mostafavi, S., Montojo, J., Shao, Q., Wright, G., Bader, G. D., & Morris, Q. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Research, 38(suppl_2), W214–W220. https://doi.org/10.1093/nar/gkq537
135. Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & Mering, C. von. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Research, 47(Database issue), D607–D613. https://doi.org/10.1093/nar/gky1131
136. Landrum, M. J., Lee, J. M., Riley, G. R., Jang, W., Rubinstein, W. S., Church, D. M., & Maglott, D. R. (2014). ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Research, 42(Database issue), D980–D985. https://doi.org/10.1093/nar/gkt1113
137. Sabiha, B., Bhatti, A., Roomi, S., John, P., & Ali, J. (2021). In silico analysis of non-synonymous missense SNPs (nsSNPs) in CPE, GNAS genes and experimental validation in type II diabetes mellitus

through Next Generation Sequencing. Genomics, 113(4), 2426–2440. https://doi.org/10.1016/j.ygeno.2021.05.022

138. Gong, H., Zhang, H., Zhu, J., Wang, C., Sun, S., Zheng, W.-M., & Bu, D. (2017). Improving prediction of burial state of residues by exploiting correlation among residues. BMC Bioinformatics, 18(S3). https://doi.org/10.1186/s12859-017-1475-5
139. Carugo, O., & Pongor, S. (2008). A normalized root-mean-square distance for comparing protein three-dimensional structures. Protein Science, 10(7), 1470–1473. https://doi.org/10.1110/ps.690101
140. Ardito, F., Giuliani, M., Perrone, D., Troiano, G., & Muzio, L. L. (2017). The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). International Journal of Molecular Medicine, 40(2), 271–280. https://doi.org/10.3892/ijmm.2017.3036
141. de Almeida, C. J. G. (2017). Caveolin-1 and Caveolin-2 Can Be Antagonistic Partners in Inflammation and Beyond. Frontiers in Immunology, 8, 1530. https://doi.org/10.3389/fimmu.2017.01530

PLAGIARISM REPORT

Molecular Screening of Caveolin 1 _(CAV1)_ Gene Polymorphisms in Patients with Type II Diabetes Mellitus

by Shanza Ayub

Submission date: 16-May-2024 11:56AM (UTC+0300)
Submission ID: 2380834994
File name: UG_FYP_THESIS_Group_01_FINAL.docx (20.99M)
Word count: 16186
Character count: 95785

Molecular Screening of Caveolin 1 _(CAV1)_ Gene Polymorphisms in Patients with Type II Diabetes Mellitus

ORIGINALITY REPORT

13%	10%	10%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.ncbi.nlm.nih.gov Internet Source	1%
2	Submitted to Cranfield University Student Paper	1%
3	www.mdpi.com Internet Source	1%
4	www.nature.com Internet Source	<1%
5	link.springer.com Internet Source	<1%
6	Submitted to Higher Education Commission Pakistan Student Paper	<1%
7	www.frontiersin.org Internet Source	<1%
8	orca.cardiff.ac.uk Internet Source	<1%