

**A STUDY ON INSULIN-RELATED GENES AND THEIR VARIANTS: IN-SILICO
INSIGHTS INTO THEIR SYSTEMIC IMPACTS BEYOND GLUCOSE
METABOLISM**

**Zainab Khan¹, Ayesha Humera², Qasim Bee Ayesha³,
Khadija Faiqa Ahmed⁴**

¹Student, Microbiology dept, Shadan Degree College for Women, India

^{2,4}Student, Genetics dept, Shadan Degree College for Women, India

³Student, Chemistry dept, Shadan Degree College for Women, India

TABLE OF CONTENT

INTRODUCTION	10
GENETICS.....	10
BIOINFORMATICS	12
INSILICO ANALYSIS	12
INSULIN	15
INS.....	33
INSR	35
PPARG	37
AKT2	38
FOXP3	40
METHODOLOGY	42
INS.....	42
VARIANT.....	42
VARIANT c.143T>C (p. Phe48Ser)	43
VARIANT c.266G>A (p. Arg89His).....	44
VARIANT c.274G>T (p. Val92Leu).....	45
VARIANT c.266G>T (p. Arg89Leu)	46
INSR	47
VARIANT c.3104G>T (p. Gly1035Val)	47
VARIANT c.1459A>G (p. Lys487Glu).....	48
VARIANT c.1177G>A (p. Gly393Arg).....	49
PPARG	50
VARIANT c.362A>G (p. Tyr121Cys)	50
VARIANT c.1184G>A (p. Arg395His).....	51
AKT2	52
VARIANT c.821G>A (p. Arg274His).....	52
VARIANT c.49G>A (p. Glu17Lys)	53
VARIANT c.1112G>A (p. Arg371His).....	54
VARIANT c.904A>G (p. Ser302Gly).....	55

VARIANT c.212A>G (p. Asn71Ser).....	56
FOXP3	57
VARIANT c.224C>T (p. Pro75Leu).....	57
VARIANT c.1189C>T (p. Arg397Trp).....	58
VARIANT c.1112T>G (p. Phe371Cys).....	59
VARIANT c.1150G>A (p. Ala384Thr).....	60
RESULTS.....	61
DISCUSSION.....	99
CONCLUSION	102

LIST OF FIGURES

Figure 1. The central dogma of molecular biology. Segments of DNA, called genes, are transcribed into mRNA copies. mRNA is then “read” in three-nucleotide codons to specify the order of amino acids in a protein. (MATTAINI).....	11
Figure 2 STRUCTURE OF INSULIN.....	15
Figure 3. Schematic of the stages of insulin biosynthesis.	17
Figure 4. Schematic representation of insulin release from the pancreatic beta-cell	18
Figure 5 . Schematic of wild-type insulin interacting with the insulin receptor (IR). Displayed domains: insulin B-chain (gold), insulin A-chain (light green), IR L1 (blue), IR α CT' (pink). Insulin disulfide bonds are seen in yellow.	18
Figure 6 Schematic summary of insulin actions on the reproductive axis.	20
Figure 7 effects of insulin resistance on the nervous system.....	22
Figure 8 Diagramatic representation of insulin resistance in cells.....	25
Figure 9 . Inflammatory consequences and possible markers related to polycystic ovary syndrome (PCOS)	30
Figure 10 Protein Expression. INS chain protein expression in normal tissues and cell lines.	34
Figure 11 Protein Expression. INSR chain protein expression in normal tissues and cell lines.	36
Figure 12 Protein Expression. AKT2 chain protein expression in normal tissues and cell lines.	39

ABSTRACT

Insulin, a vital hormone produced by the pancreas, regulates glucose metabolism and energy balance in the human body. This study investigates the effects of insulin on various systems, including the immune, reproductive, and nervous systems, in addition to its traditional role in metabolism. Insulin's primary function is to promote glucose uptake into cells, but its impact extends to several physiological processes, influencing a range of systemic functions.

In the immune system, insulin regulates inflammatory responses, but insulin resistance, a hallmark of Type 2 diabetes, leads to chronic inflammation, impairing immune function and increasing infection susceptibility. The study also explores the interaction between insulin resistance and autoimmune diseases, shedding light on the complex relationship between insulin signalling and immune cell activation. In the reproductive system, insulin resistance is associated with conditions like polycystic ovary syndrome (PCOS), which causes hormonal imbalances and fertility issues. Insulin's impact on ovarian function and regulation of sex hormone levels is examined, showing how metabolic disturbances disrupt reproductive health.

The nervous system is profoundly affected by insulin, with insulin resistance linked to cognitive decline, neurodegenerative diseases such as Alzheimer's, and diabetic neuropathy. The role of insulin in brain function, synaptic plasticity, and neuroprotection is discussed, emphasizing how metabolic dysfunction influences neurological health.

We utilized various computational tools, including SIFT, PolyPhen-2, and MutationTaster, to predict the functional impact of selected variants. The study focused on key insulin-associated genes such as INS, INSR, IGF1, IGF2, and PIK3CA, which are known to influence insulin signalling and regulation across different biological systems. By applying bioinformatics techniques, we assessed the potential deleterious effects of these variants.

OBJECTIVE

- To conduct a comprehensive study on insulin-related genes and their role in various biological systems, including neurological, immune, and reproductive functions.
- To analyse the genetic variants in insulin-related genes and their potential impact on different biological systems.
- To assess the pathogenicity of these variants using in-silico tools

INTRODUCTION

GENETICS

Genetics explores the function of genes; the unique instruction sets within our bodies that define individual characteristics. Thousands of genes, each carrying a specific instruction, contribute to our individuality. Alterations in these genes can lead to genetic conditions or diseases, arising either through inheritance or spontaneous mutations.

Inheritance patterns of gene alterations vary, including autosomal dominant, autosomal recessive, X-linked, mitochondrial, and multifactorial inheritance. These patterns dictate how gene alterations are passed through families.

Genes are organized into chromosomes, and changes in chromosome number or structure can also cause developmental differences, exemplified by Down syndrome. (**MEDICINE**)

Heredity, the transmission of characteristics from parents to offspring, is governed by genes, the functional units of heritable material. While species maintain constancy through a shared gene set, individual variation arises from differences in gene forms. An organism's genotype, the inherited gene combination, contrasts with its phenotype, the observable traits resulting from gene-environment interactions. Though the genotype remains fixed, the phenotype evolves with changing environments.

Genetics, the study of genes, reveals their fundamental role in all aspects of an organism. The universality of genetic systems across diverse species underscores evolutionary relatedness, highlighting the shared genetic heritage of life. Genetic advancements, from selective breeding to recombinant DNA technology, have profoundly impacted medicine, agriculture, and industry. Understanding gene function is increasingly crucial in disease research.

Inheritance patterns, explained by Mendelian genetics and chromosome organization, underlie hereditary transmission. Gene function at the molecular level, including DNA transcription and RNA translation into proteins, elucidates the mechanisms of heredity. Finally, heredity's role in evolution underscores its central importance in biology.

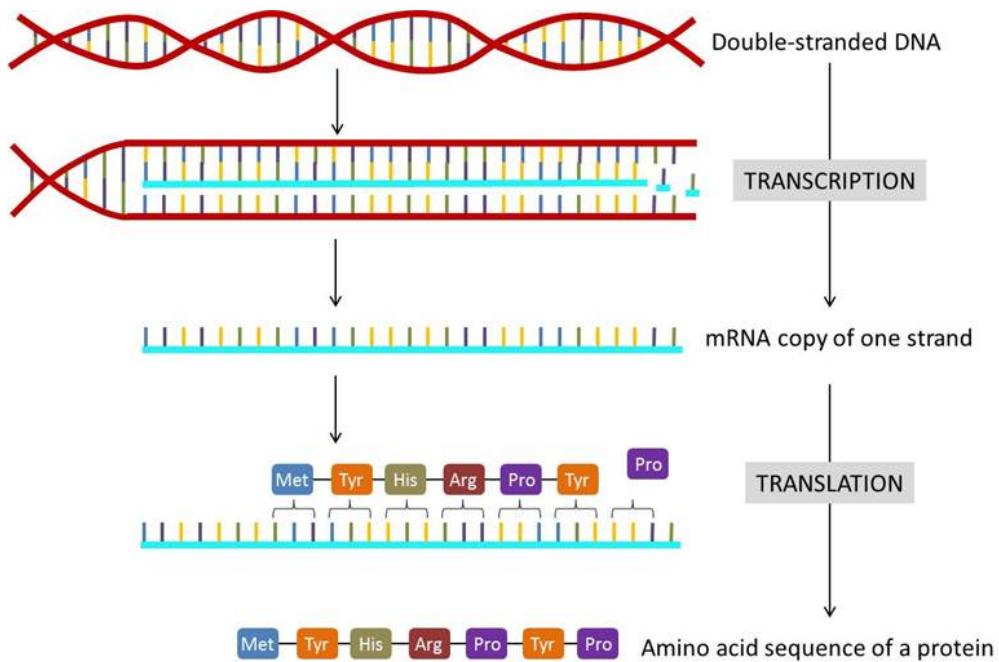


Figure 1. The central dogma of molecular biology. Segments of DNA, called genes, are transcribed into mRNA copies. mRNA is then “read” in three-nucleotide codons to specify the order of amino acids in a protein. (MATTAINI)

DNA encodes genetic information through nucleotide sequences, which are translated into proteins, the building blocks of organisms. Proteins, composed of amino acid chains, determine an organism's characteristics. The DNA nucleotide sequence is first transcribed into RNA, a single-stranded molecule, at the gene's chromosomal location. In eukaryotes, this RNA, after processing (capping, tailing, and splicing), becomes mature messenger RNA (mRNA) and moves from the nucleus to the cytoplasm.

Translation, the process of converting mRNA's nucleotide sequence into an amino acid sequence, occurs on ribosomes. mRNA codons (three-nucleotide sequences) specify amino acids. Transfer RNA (tRNA) molecules, each carrying a specific amino acid and an anticodon complementary to an mRNA codon, deliver amino acids to the ribosome. The ribosome assembles the amino acid chain, which folds into a functional protein.

The genetic code, the correspondence between codons and amino acids, is nearly universal. Ribosomes, composed of proteins and ribosomal RNA (rRNA), facilitate translation. In eukaryotes, proteins destined for secretion are synthesized on ribosomes attached to the endoplasmic reticulum (ER) and transported to the Golgi apparatus for further processing and secretion.

tRNA molecules, with anticodons matching mRNA codons, ensure the correct amino acid sequence. The ribosome moves along the mRNA, adding amino acids to the growing chain until a stop codon is reached. The resulting polypeptide chain folds into a three-dimensional protein structure, determined by amino acid properties and interactions. (**HEREDITY - GENES, DNA, INHERITANCE | BRITANNICA**)

BIOINFORMATICS

Bioinformatics integrates biology and computer science to manage and analyse vast biological datasets, primarily DNA and protein sequences. It's essential for understanding complex biological systems, identifying genes and proteins, determining their functions, and exploring evolutionary relationships. Utilizing computational techniques like pattern recognition and machine learning, bioinformatics reveals patterns within massive datasets that would be impossible to discern manually. This field underpins modern biotechnology, facilitating drug development, genetic therapies, and advancements in various biological research areas.

Historically, bioinformatics emerged in the 1960s with the use of computers to analyse protein sequences, revolutionizing the understanding of protein structure and function. The Human Genome Project, a landmark achievement, demonstrated the indispensable role of bioinformatics in large-scale biological projects, enabling the sequencing and mapping of the entire human genome. (**BIOINFORMATICS**)

INSILICO ANALYSIS

The continuous advancement in proteome research and the rapid pace of genome sequencing have resulted in an overwhelming influx of protein sequences, both experimentally derived and predicted. This surge in data presents a significant challenge for bioinformatics, necessitating the development and maintenance of comprehensive, non-redundant protein sequence databases. Furthermore, effective *in silico* proteome analysis tools are crucial for bridging the gap between sequence information and functional characterisation, particularly for human and model organisms, to gain a deeper understanding of health and disease.

Resources

Sequence Databases

Sequence databases are fundamental tools for genome and proteome analysis, serving as repositories for the vast amounts of biological data generated. These databases are essential resources for biological and medical research, especially when combined with powerful search and computational analysis tools.

Nucleotide Sequence Databases

Nucleotide sequence databases store nucleic acid sequence data from genome sequencing projects and other sequencing efforts. The International Nucleotide Sequence Database Collaboration, involving EMBL-EBI, DDBJ, and GenBank, is responsible for collecting, organizing, and distributing the majority of nucleotide sequence data. These databases strive for completeness, aiming to record and make publicly available every known nucleic acid sequence, with automatic updates occurring between the databases every 24 hours. The exponential growth of these databases, driven by technological advancements, results in a doubling of their size approximately every year.

protein Sequence Databases

Protein sequence databases store information on proteins, distinguishing between universal databases covering all species and specialised databases focused on specific protein families or organisms. Universal protein sequence databases can be further categorised into simple archives of sequence data and annotated databases, with the latter being particularly valuable for proteome analysis. PIR was the first protein sequence database, succeeded by PIR-International, while SWISS-PROT is an annotated database known for its high level of annotation and minimal redundancy. TrEMBL serves as a computer-annotated supplement to SWISS-PROT, aiming to make new sequences available quickly. SP_TR_NRDB (SWALL) combines SWISS-PROT and TrEMBL to provide a comprehensive collection of protein sequence entries. The CluSTr database offers automatic classification of SWISS-PROT and TrEMBL proteins into related groups.

Protein Tertiary Structure Databases

Protein tertiary structure databases, such as PDB, store information on known protein structures, with derived databases like DSSP, HSSP, FSSP, SCOP, and CATH enabling comparative studies of 3D structures. These databases facilitate the analysis of relationships between sequence, secondary structure elements, and 3D structure.

Proteome Analysis Databases and Tools:

Proteome Analysis Databases

Proteome analysis databases are designed to provide comprehensive statistical and comparative analyses of predicted proteomes from fully sequenced organisms. These databases, such as the proteome analysis database at EBI, integrate information from various sources to facilitate the classification of proteins in complete proteome sets. The International Protein Index (IPI) serves as a guide to the main databases describing the human and mouse proteome.

Proteome Analysis Tools

Proteome analysis tools, like those available on the ExPASy server, provide a variety of functionalities for analysing proteins, including sequence analysis, structure prediction, and post-translational modification prediction. The proteome analysis database itself offers capabilities for whole proteome comparisons, analysis of structural features, and Fasta similarity searches. These tools and databases are essential for discovering protein function, characterising proteins, and conducting comparative proteomic studies. ([PRUESS AND APWEILER](#))

INSULIN

The discovery of the insulin hormone over 100 years ago, and its subsequent therapeutic application, marked a key landmark in the history of medicine and medical research. The many roles insulin plays in cell metabolism and growth have been revealed by extensive investigations into the structure and function of insulin, the insulin tyrosine kinase receptor (IR), as well as the signalling cascades, which occur upon insulin binding to the IR. It has become clear that insulin signalling is not only fundamental in tissues previously identified as “insulin sensitive” (such as adipose, liver, and muscle tissue), but in almost all tissues of the body. (**ATAIE-ASHTIANI AND FORBES**)

Insulin is a peptide hormone secreted by the β cells of the pancreatic islets of Langerhans and maintains normal blood glucose levels by facilitating cellular glucose uptake, regulating carbohydrate, lipid and protein metabolism and promoting cell division and growth through its mitogenic effects. Insulin regulates cellular energy and macronutrient balance, driving postprandial anabolism. It facilitates glucose uptake in insulin-dependent tissues, inhibiting lipolysis and promoting lipogenesis in adipocytes, and glycogenesis and carbohydrate oxidation in myocytes. Insulin suppresses gluconeogenesis and lipolysis from muscle amino acids, while promoting protein synthesis when amino acids are available. (**WILCOX**)

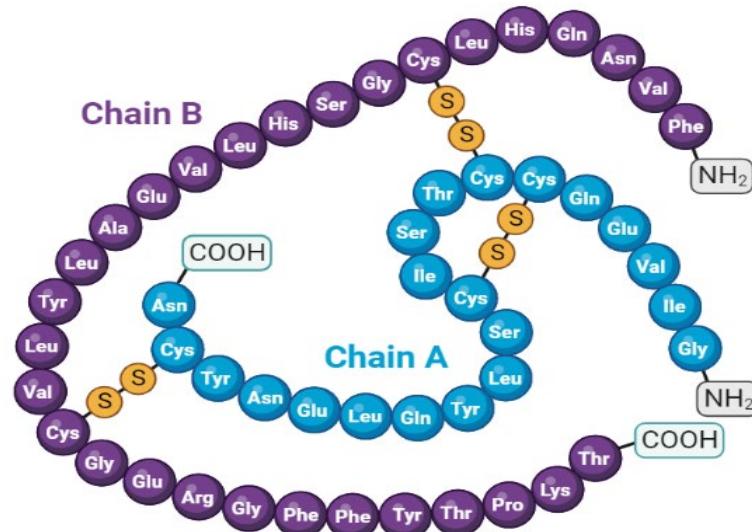


Figure 2 STRUCTURE OF INSULIN

Diabetes mellitus, a group of metabolic disorders marked by persistent high blood sugar, affects a significant portion of the global population. While most diabetes is attributed to reduced insulin production or increased insulin resistance, a smaller portion, 1-5%, is caused by single gene mutations. These monogenic forms of diabetes arise from defects in genes responsible for beta-cell function, including potassium channels, glucose sensors, transcription factors, and the insulin gene itself. Insulin gene mutations can lead to unique diabetes subtypes by disrupting insulin folding, causing endoplasmic reticulum stress and beta-cell death, or impairing insulin's interaction with its receptor. These mutations can result in neonatal-onset diabetes mellitus (NDM), Mutant Insulin-gene Induced Diabetes of Youth (MIDY), and Maturity Onset Diabetes of the Young (MODY). NDM can be transient or permanent, and both MIDY and MODY are often caused by heterozygous mutations with dominant negative effects. Monogenic diabetes is frequently misdiagnosed as type 1 or type 2 diabetes. Genetic testing and a thorough understanding of the underlying mutations are crucial for accurate diagnosis and tailored treatment. Research into the structural changes in insulin caused by these mutations is also vital for developing improved insulin analogues.

INSULIN BIOSYNTHESIS

Insulin, a crucial hormone for glucose regulation, is synthesized within the beta-cells of the pancreatic islets of Langerhans. The process begins with the transcription and translation of the insulin gene, located on chromosome 11. This gene is composed of three exons and two introns; importantly, exons 2 and 3 encode the mature insulin protein, while exon 1 plays a regulatory role. The initial product of translation is preproinsulin, a single-chain polypeptide comprising four domains: a signal peptide, the B-chain, the C-peptide, and the A-chain. Through a series of post-translational modifications, the signal peptide is cleaved, and disulfide bonds are formed. These disulfide bonds, specifically two inter-chain bonds between the A and B chains and one intra-chain bond within the A-chain, are critical for stabilizing the mature insulin's three-dimensional structure. The C-peptide is also removed, yielding the mature insulin molecule, which consists of the A-chain (21 amino acid residues) and the B-chain (30 amino acid residues). Mutations affecting the insulin gene, particularly those that disrupt the

formation of these disulfide bonds, can lead to misfolding of the protein and subsequent beta-cell dysfunction, ultimately resulting in various forms of diabetes.

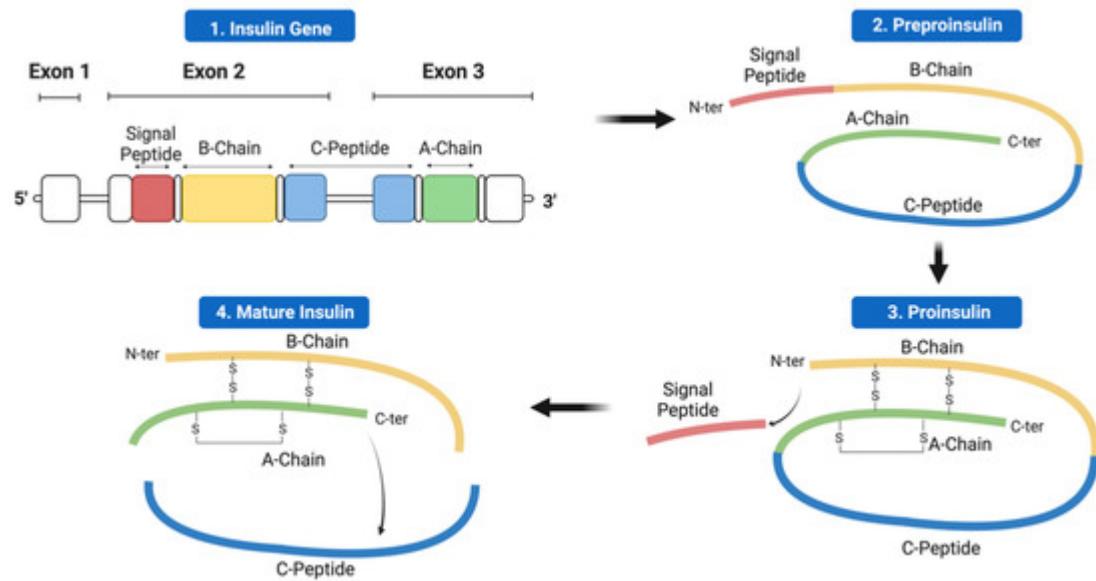


Figure 3. Schematic of the stages of insulin biosynthesis.

Insulin biosynthesis begins with gene transcription to mRNA, followed by preproinsulin translation at the endoplasmic reticulum (ER). Within the ER, the signal peptide is removed, forming proinsulin, which then folds, establishing crucial disulfide bonds. Misfolding here can trigger ER stress and beta-cell dysfunction. Properly folded proinsulin moves to the Golgi, where it's packaged into secretory vesicles. Enzymes (PC1/3, PC2, CPE) cleave the C-peptide, yielding mature insulin, stored as zinc-stabilized hexamers. Upon stimulation, insulin is released into the bloodstream. Mutations can disrupt this process, causing continuous, unregulated insulin secretion. Once released, insulin binds to its receptor (IR), initiating intracellular signalling cascades.

MECHANISM OF ACTION OF INSULIN SECRETAGOGUES

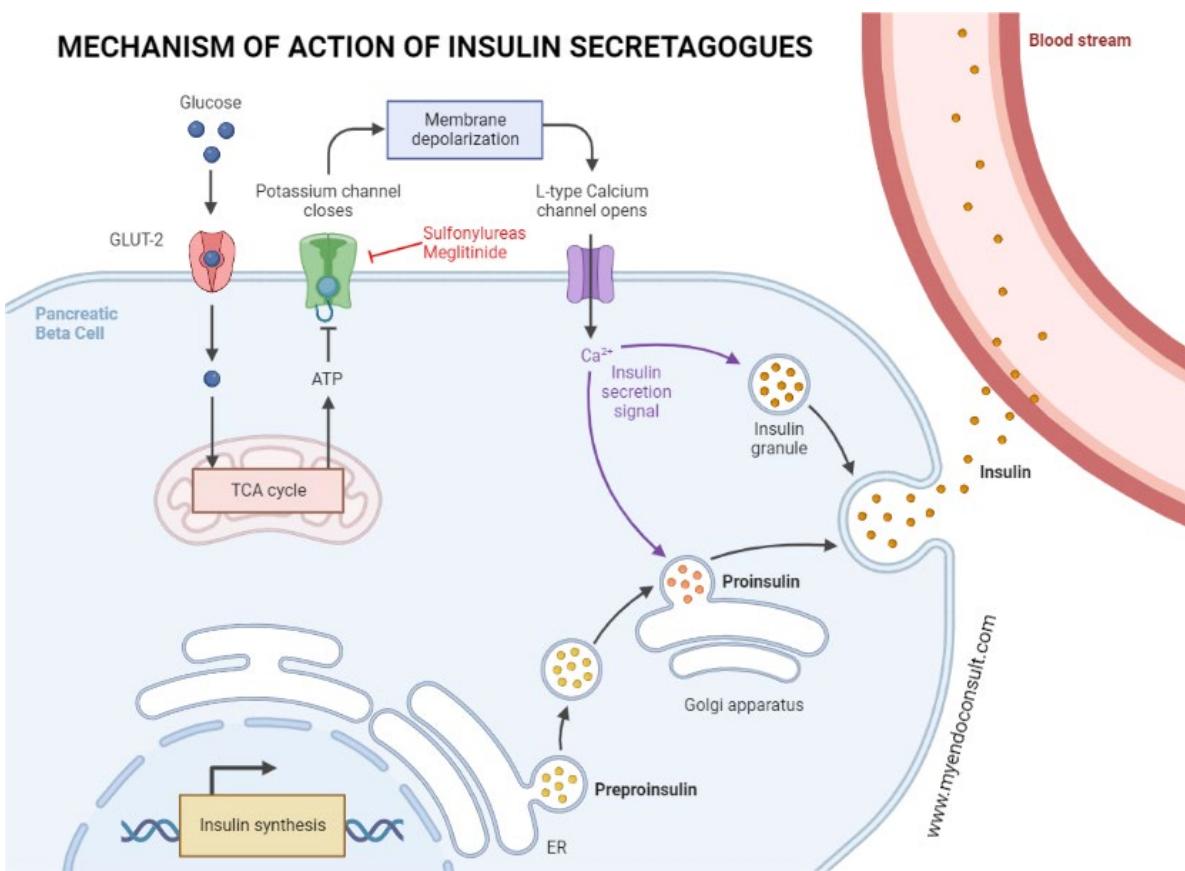


Figure 4. Schematic representation of insulin release from the pancreatic beta-cell

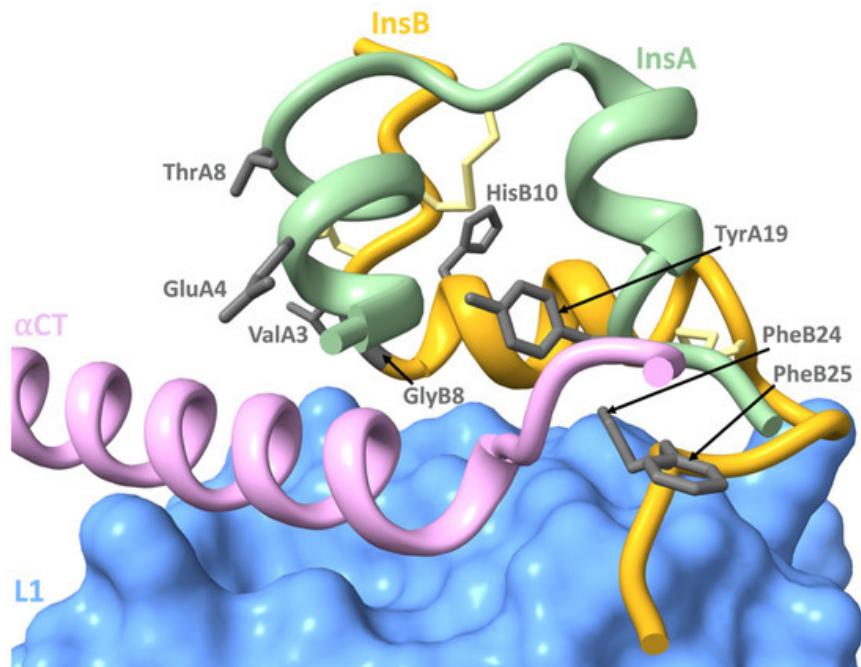


Figure 5 . Schematic of wild-type insulin interacting with the insulin receptor (IR).
Displayed domains: insulin B-chain (gold), insulin A-chain (light green), IR L1 (blue), IR αCT' (pink). Insulin disulfide bonds are seen in yellow.

The insulin receptor (IR) is a disulfide-linked heterodimer, composed of two $\alpha\beta$ monomers, with the α -subunits linked by intermonomer disulfide bonds. This complex features an ectodomain with distinct subdomains (L1, CR, L2, FnIII, α CT) and an intracellular region. The α -subunits and β -subunit N-terminus contain ligand-binding sites, while the β -subunit's transmembrane helix anchors it to the membrane. Cytoplasmic domains (JM, TKD, C-tail) are phosphorylated upon insulin binding, initiating signaling. Insulin-IR interaction activates two primary pathways: the PI3K/AKT pathway, crucial for metabolism and glucose uptake, and the MAPK pathway, which drives mitogenic activity and growth. These pathways underscore insulin's critical role in regulating metabolism, growth, and development.

(ATAIE-ASHTIANI AND FORBES)

INSULIN'S MULTIFACETED INFLUENCE: BEYOND GLUCOSE REGULATION

1. REPRODUCTION

Insulin plays a crucial role in both energy homeostasis and reproductive function, with its influence on the latter being a complex and still not fully understood area. While insulin's role in glucose regulation and energy balance is well-established, its involvement in reproduction is multifaceted. Studies in animal models demonstrate insulin's impact on GnRH/LH secretion, essential for reproductive processes. Furthermore, insulin appears to be involved in the prenatal programming of adult reproductive capabilities, though the precise mechanisms are unclear. Researchers are investigating specific hypothalamic neurons as potential targets for insulin's reproductive effects. Clinically, insulin dysregulation is linked to human fertility issues and reproductive disorders. Despite evidence supporting insulin's role in reproductive neuroendocrine function, many questions remain about the specific mechanisms, target neurons, and the interplay between insulin and other metabolic/hormonal signals in both development and adulthood. More research is needed to fully understand insulin's influence on reproduction and translate this knowledge into clinical applications.

(SLIWOWSKA ET AL.)

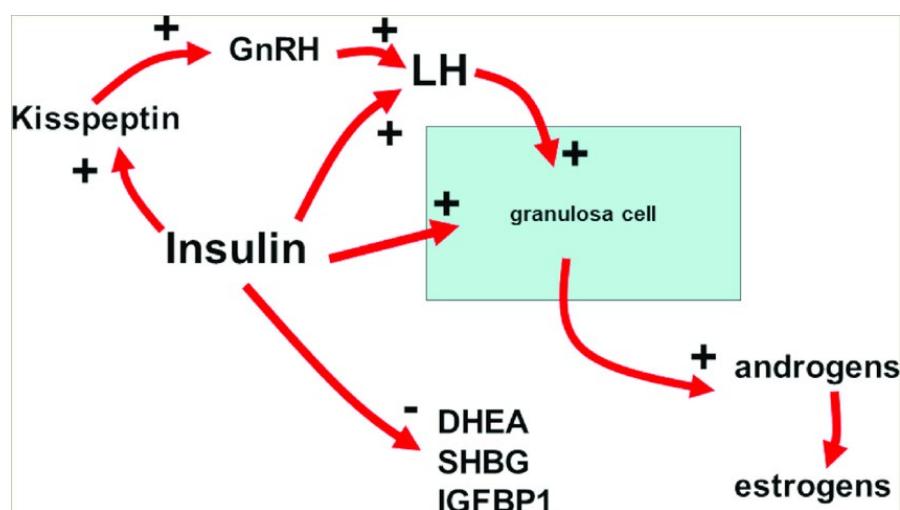


Figure 6 Schematic summary of insulin actions on the reproductive axis.

Impaired insulin signalling, independent of obesity and hyperglycemia, can significantly disrupt female reproductive function. A study using lean, normoglycemic mouse models with

varying degrees of insulin resistance demonstrated that altered insulin signalling directly affects several reproductive parameters. The insulin-resistant mice exhibited irregular estrous cycles and displayed abnormal distribution and morphology of ovarian follicles. These ovarian disruptions were not related to elevated androgen levels. Furthermore, the study revealed that compromised insulin signalling negatively impacts pregnancy, leading to decreased success in early gestation and, in successful pregnancies, reduced embryo weights and increased placental calcification. These findings highlight the crucial role of proper insulin signalling in female reproductive health, showing that disruptions in this pathway alone, without the confounding effects of excess adipose tissue or high glucose levels, can negatively impact multiple stages of reproduction, from ovarian function to successful pregnancy outcomes. (**NANDI ET AL.**)

2.NERVOUS SYSTEM

The brain, once viewed primarily as a glucose-dependent organ, is now recognized as a critical insulin-sensitive tissue with profound influence over systemic metabolic homeostasis, cognitive function, and emotional well-being. Disruptions in brain insulin signaling, frequently observed in conjunction with peripheral insulin resistance in prevalent conditions like obesity, type 2 diabetes, and increasingly recognized in neurodegenerative diseases like Alzheimer's, contribute to a cascade of debilitating health consequences. These include impaired glucose regulation, leading to hyperglycemia and increased risk of diabetes complications; dysregulated energy balance, impacting appetite, food intake, and energy expenditure, contributing to weight gain and metabolic syndrome; cognitive decline, affecting memory, learning, and executive functions, increasing the risk of dementia and Alzheimer's; and mood disorders, such as depression and anxiety, significantly impacting quality of life. Insulin's access to the central nervous system is tightly regulated by a saturable transport mechanism across the blood-brain barrier, emphasizing the importance of precisely controlled insulin delivery to the brain. Intranasal insulin administration offers a non-invasive and promising method for delivering insulin directly to the brain, bypassing peripheral metabolic effects and enabling researchers to specifically investigate the central actions of insulin. Studies employing this technique suggest that intranasal insulin holds therapeutic potential for improving glucose control, modulating energy homeostasis by influencing appetite, food intake, and energy expenditure, enhancing memory and cognitive performance, and positively impacting mood,

potentially through modulation of the hypothalamic-pituitary-adrenal (HPA) axis and its role in stress response. While further rigorous research, including large-scale clinical trials, is crucial to fully elucidate the complex underlying mechanisms, optimize dosage and treatment protocols, and establish long-term efficacy and safety, these compelling findings underscore the profound significance of brain insulin signaling in overall health and well-being, opening potential new avenues for the treatment of a range of metabolic, cognitive, and affective disorders, including diabetes, obesity, Alzheimer's disease, and depression. (**LEE ET AL.**)

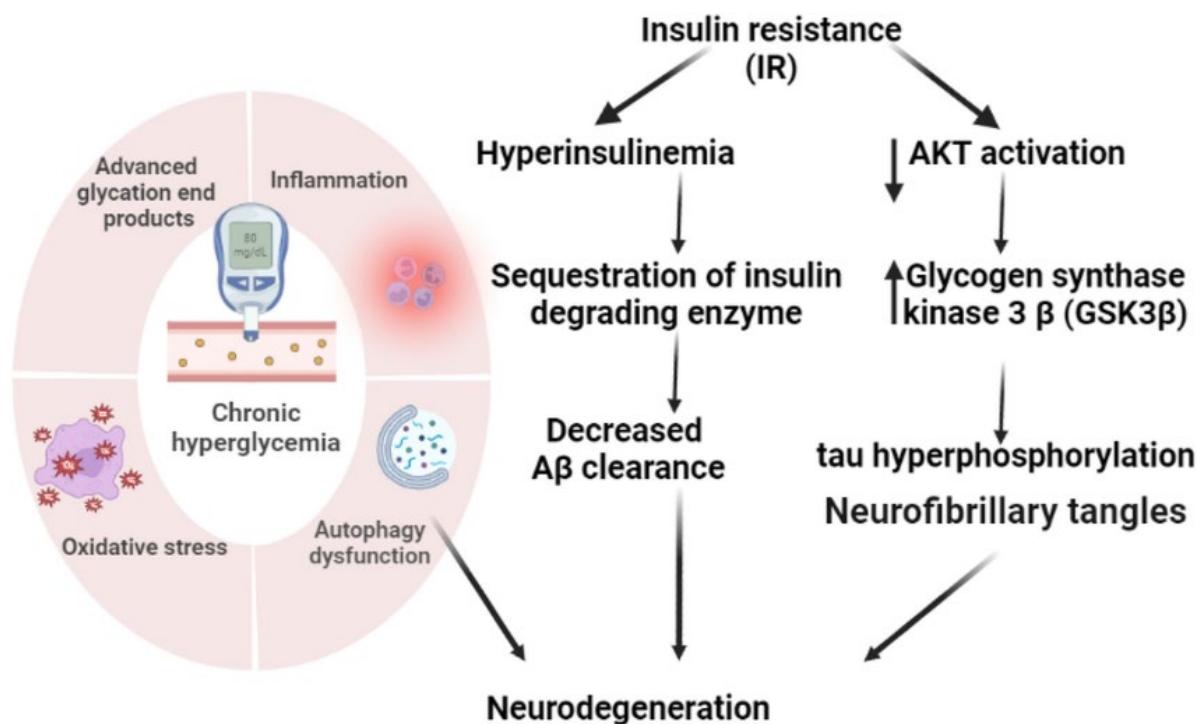


Figure 7 effects of insulin resistance on the nervous system.

3. IMMUNE SYSTEM

Insulin's influence extends far beyond its traditional role in glucose metabolism, profoundly impacting the intricate workings of the immune system. This influence encompasses a complex interplay that governs the development, differentiation, metabolic programming, and functional specialisation of diverse immune cell populations, spanning both innate and adaptive immunity. Through its receptor (InsR) and key downstream signalling cascades like PI3K/Akt/mTORC1 and MAPK/ERK, insulin exerts significant control over immune responses, dictating their intensity, direction, and duration.

Insulin's immunomodulatory effects are highly context dependent. Within innate immunity, this is clearly illustrated by its ability to promote pro-inflammatory M1 macrophage

polarisation for effective pathogen clearance while simultaneously contributing to immune tolerance under specific conditions. This dual nature is mirrored in neutrophils, where insulin enhances crucial functions like ROS production and degranulation alongside its influence on apoptosis and NET formation. Other innate immune cells, including eosinophils, dendritic cells, natural killer cells, and innate lymphoid cells, also demonstrate sensitivity to insulin signalling, although the precise mechanisms and functional consequences of this interaction require further investigation.

In the adaptive immune system, insulin's crucial role is evident in T cell activation, proliferation, and effector differentiation, particularly its support of the glycolytic metabolism essential for robust immune responses. However, the delicate balance between promoting effector T cell functions and potentially dampening regulatory T cell activity raises important questions regarding insulin's contribution to immune homeostasis and the prevention of autoimmunity. In B cells, while the intricacies of InsR signalling remain to be fully elucidated, existing evidence suggests its involvement in various stages of B cell development, differentiation, and antibody production, with potential implications for humoral immunity and class switch recombination.

The implications of dysregulated insulin signalling in various disease contexts are profound. Insulin resistance, a hallmark of obesity and type 2 diabetes, is not limited to metabolic tissues; immune cells themselves can become insulin resistant, leading to a complex interplay of impaired immune function and chronic inflammation. This "immune system IR" is a key driver of the intricate relationship between metabolic dysfunction and chronic inflammation, contributing significantly to the pathogenesis of obesity-related complications, including increased susceptibility to infections, impaired tumor surveillance, and exaggerated inflammatory responses. Furthermore, the role of insulin signalling extends to other disease states, such as cancer, pre-eclampsia, and viral infections like COVID-19, demonstrating how disruptions in this signalling pathway can profoundly influence disease progression and outcomes.

Beyond specific diseases, the influence of insulin signalling on trained immunity and aging warrants attention. Insulin's impact on metabolic reprogramming and cytokine production in innate immune cells may contribute to the long-term functional changes observed in trained immunity, with potential implications for both protective and detrimental immune responses.

In the context of aging, the chronic low-grade inflammation characteristic of immunosenescence may be linked to altered insulin signalling, highlighting the importance of understanding how insulin's role in the immune system changes over time.

The acute effects of insulin and IGF1R signaling on immune cells by cell type.

Immune cell	Adjuvant effects of acute InsR/IGF1R signalling	References
Macrophages	Insulin drives increased production of inflammatory cytokines (IL-6, TNF α) in the presence of LPS/TLR activation InsR required for M1-like polarization	(TESSARO ET AL.) ; (RATTER ET AL.) (MAUER ET AL.) ; (KNUEVER ET AL.) ; (BAUMGARTL ET AL.),
Dendritic cells	Insulin drives increased scavenger receptor expression via ERK signalling with and without TLR activation	(LU ET AL.)
Neutrophils	Insulin drives increased ROS formation via potentiation of the PI3K signal upon priming with N-formyl oligopeptide IGF1 delays apoptosis via PI3K signalling	(SAFRONOVA ET AL., 2001) (HIMPE ET AL.)
Eosinophils	Insulin drives increased peripheral eosinophil levels and mucus production in the lungs of healthy and diabetic mice upon ovalbumin allergen challenge.	(FERRIERA ET AL., 2017)
ILCs	IGF1R in ILC3s supports differentiation and function against respiratory pathogens in neonatal lungs	(OHERLE ET AL.)
T cells	Increased IL-2 responsiveness and chemotaxis, improved glycolytic and mitochondrial metabolism, increased IFNy production is mediated by InsR Th17 polarization of CD4 T cells via IGF1R Insulin drives reduction of IL-10 production by regulatory T regs	(DEBENEDETTE AND SNOW); (BERMAN AND CENTER); (FISCHER ET AL.); (TSAI ET AL., 2018) (DiTORRO ET AL., 2020) (HAN ET AL.)
B cells	Undetermined	

In summary, insulin signalling plays a fundamental role in orchestrating immune function. The intricate connections between insulin signalling and immune cell metabolism, differentiation, and effector activities underscore the therapeutic potential of targeting this pathway. A deeper understanding of the molecular mechanisms governing these interactions is essential for developing innovative strategies to combat immune dysfunction in a wide spectrum of diseases, from metabolic disorders and infections to cancer and age-related decline. Continued research focused on dissecting the complex interplay between insulin signalling and other immune regulatory networks holds tremendous promise for advancing human health. (**MAKHIJANI ET AL.**)

INSULIN RESISTANCE

Insulin resistance occurs when the body's cells don't respond properly to insulin, hindering glucose uptake and utilization. This reduced sensitivity to insulin stems from disruptions in the complex signalling pathways initiated when insulin binds to its receptors. These pathways branch into two main functions: regulating metabolic processes (like glucose control) and controlling cell growth. Notably, in type 2 diabetes, the metabolic pathway appears to be particularly impaired, while the growth-related pathway remains relatively unaffected.

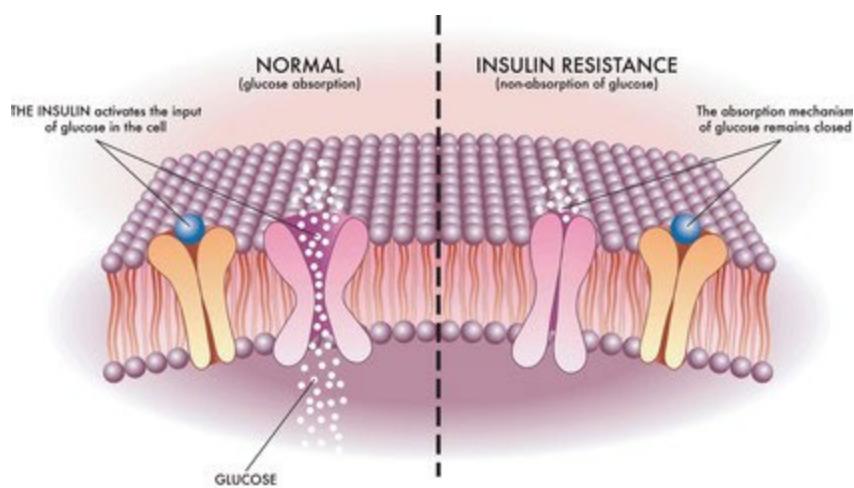


Figure 8 Diagrammatic representation of insulin resistance in cells.

Several factors contribute to insulin resistance, including:

- Genetic defects: Variations in genes encoding proteins involved in insulin signalling.
- Foetal malnutrition: Nutritional deficiencies during foetal development.

- Increased visceral fat: Excess fat accumulation around abdominal organs.

Insulin resistance is a core component of the metabolic syndrome, a cluster of risk factors that significantly increase the likelihood of developing type 2 diabetes, heart disease, high blood pressure, and polycystic ovary syndrome. The specific outcome depends on an individual's genetic predisposition. (**LEBOVITZ**).

- Research shows that reduced insulin sensitivity in muscle tissue (insulin resistance) and elevated insulin levels (hyperinsulinism) is frequently observed in conditions like high blood pressure, abdominal obesity, type 2 diabetes, unhealthy cholesterol levels (high triglycerides, low HDL), and atherosclerosis. Even in healthy middle-aged men, a significant portion (25%) exhibits insulin resistance.
- **The Genetic Insulin Resistance Hypothesis:**
 - The core idea is that a substantial portion of the population carries genetic variations affecting how insulin works. These variations lead to "primary genetic insulin resistance."
 - This primary resistance can be worsened by factors like aging, a high-fat diet, inactivity, hormonal imbalances, and certain medications, leading to "secondary insulin resistance."
- **Hyperinsulinism as a Compensatory Mechanism:**
 - When muscle tissue becomes less responsive to insulin, the pancreas produces more insulin to try and maintain normal blood sugar levels. This results in hyperinsulinism.
- **Phenotypic Expression:**
 - The specific health problems that develop due to insulin resistance and hyperinsulinism depend on an individual's genetic vulnerabilities.
 - For example:
 - If the pancreas can't produce enough insulin to compensate for resistance, blood sugar levels rise, leading to glucose intolerance and potentially type 2 diabetes.
 - In individuals genetically predisposed to high blood pressure, hyperinsulinism may exacerbate existing defects in blood pressure regulation, contributing to the development of hypertension.

In essence, genetic insulin resistance is a fundamental underlying factor that, when combined with lifestyle and environmental influences, can manifest as a range of cardiovascular and metabolic disorders. (**PEDERSEN**)

Associated Diseases (Conditions Resulting from Insulin Resistance):

Insulin resistance is strongly associated with a range of serious health conditions, including non-alcoholic fatty liver disease (NAFLD), metabolic syndrome, prediabetes or type 2 diabetes, polycystic ovarian syndrome (PCOS), and obesity. It also contributes to both microvascular complications, such as retinopathy, neuropathy, and nephropathy, which damage small blood vessels, and macrovascular complications, including stroke, peripheral artery disease (PAD), and coronary artery disease (CAD), which affect larger blood vessels.

Associated Symptoms (Signs and Clinical Manifestations):

Insulin resistance manifests through a variety of clinical and laboratory findings, including hypertension, abnormal lipid profiles (hyperlipidaemia), elevated blood glucose (hyperglycaemia), and increased uric acid levels (hyperuricemia). Systemic inflammation, endothelial dysfunction, and a prothrombotic state are also common. Furthermore, individuals may exhibit increased waist circumference, specific to gender and ethnicity, and women may display signs of Polycystic Ovary Syndrome (PCOS), such as menstrual irregularities, hirsutism, acne, and alopecia. Acanthosis nigricans, along with stigmata of rare genetic insulin resistance syndromes like Type A or Type B, can also be observed. (**FREEMAN ET AL.**)

POLYCYSTIC OVARIAN SYNDROME [PCOS]

Polycystic Ovary Syndrome (PCOS) is a common hormonal disorder affecting women of reproductive age, characterized by an imbalance in reproductive hormones that leads to a cluster of symptoms. Polycystic ovarian syndrome (PCOS) is one of the readily recognised endocrine gland illnesses in women, with an incidence range from 2.2% to 26% in India. As a leading cause of anovulatory infertility, PCOS is frequently diagnosed when women have trouble conceiving. Long-term, low-grade inflammation has emerged as a crucial factor leading to PCOS. Historically, Polycystic Ovary Syndrome (PCOS) was thought to originate from functional ovarian hyperandrogenism (FOH), resulting from an overproduction of androgens by the ovaries. Affecting a notable proportion of females, approximately one in five to six experience menstrual irregularities and fertility challenges. Stress, obesity, and hormonal imbalances are considered major contributing factors worldwide. The physiological basis of PCOS is understood to involve four key elements: disruptions in gonadotropin hormone production, the development of insulin resistance, and the amplifying effect of increased body fat. These females have erratic gonadotropin absorption and androgen biosynthesis from the adrenal and ovaries, spurred by an elevated level of insulin. According to specialists, it is regarded that the fundamental feature of PCOS is hyperandrogenism. PCOS typically begins during adolescence and is marked by significant fluctuations in various hormone levels. Disruptions in the secretion of gonadotropin-releasing hormone (GnRH), which leads to increased luteinizing hormone (LH) release, and alterations in insulin signalling, resulting in insulin resistance, are also considered key factors in the development of PCOS. While insulin resistance is a prevalent metabolic feature in many women with PCOS, often linked to genetic predisposition or obesity, it's important to note that a subset of women with classic PCOS criteria do not exhibit this characteristic. This suggests that while insulin resistance plays a significant role in many cases, it is not a universal requirement for the diagnosis or manifestation of the syndrome." PCOS significantly increases the risk of complications such as diabetes mellitus in women. (**PURWAR AND NAGPURE**)

DIAGNOSIS

PCOS diagnosis hinges on identifying at least two of three key features: ovulatory dysfunction, hyperandrogenism, and polycystic ovaries. Ovulatory dysfunction, marked by infrequent (oligomenorrhea) or absent (amenorrhea) periods, stems from hormonal imbalances disrupting egg release. Hyperandrogenism, an excess of male hormones, manifests clinically as hirsutism,

acne, or alopecia, and can be confirmed biochemically through elevated androgen levels in blood tests. Polycystic ovaries, visualized via ultrasound, reveal numerous small follicles, though their presence isn't universal in PCOS, nor does it solely indicate the condition. Recent genetic research also points towards the possibility of PCOS consisting of multiple subtypes.

Impact and Associated Symptoms:

- PCOS is a leading cause of infertility due to anovulation.
- Many women are diagnosed when they experience difficulty conceiving.
- Beyond reproductive issues, PCOS can manifest in various ways, including:
 - Increased hair growth (hirsutism)
 - Darkened skin patches (acanthosis nigricans)
 - Acne
 - Insulin resistance
 - Irregular menstrual bleeding
 - Metabolic dysfunction
 - Cardiovascular problems
 - Inflammatory responses.

In essence, PCOS is a complex hormonal disorder with diverse symptoms, impacting both reproductive and overall health. ([**\(ABOUT POLYCYSTIC OVARY SYNDROME \(PCOS\) | NICHD - EUNICE KENNEDY SHRIVER NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT\)**](#)

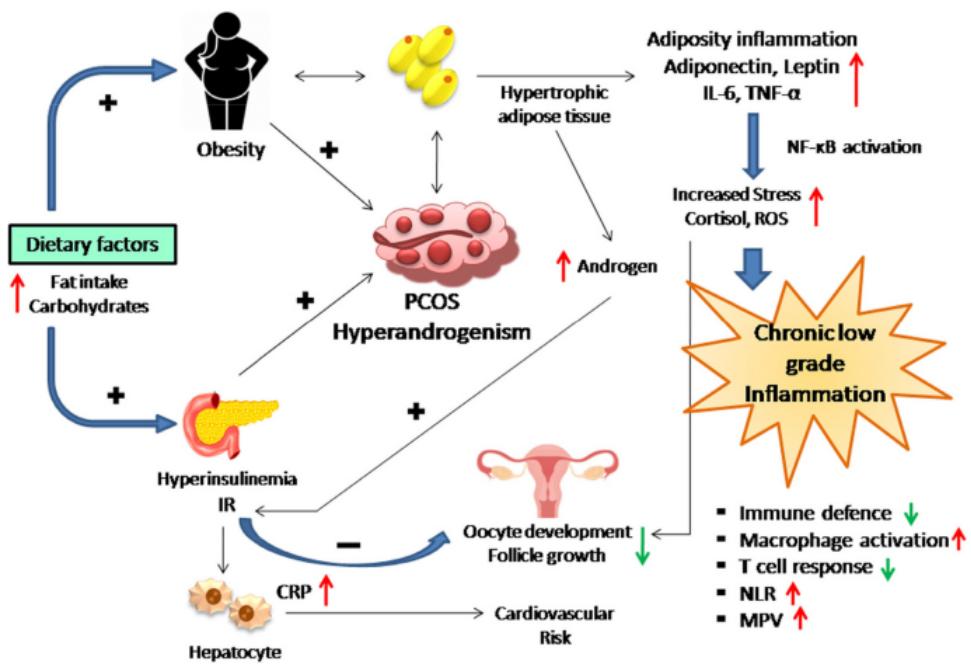


Figure 9 . Inflammatory consequences and possible markers related to polycystic ovary syndrome (PCOS)

TREATMENT

PCOS treatment is tailored to address individual concerns, focusing on managing specific symptoms such as infertility, hirsutism, acne, or obesity. Lifestyle modifications, including weight loss through a combination of a low-calorie diet and moderate exercise, are often recommended as a first-line approach. Even a modest 5% reduction in body weight can significantly improve PCOS symptoms and enhance the effectiveness of medications, while also aiding in fertility. Healthcare providers may also prescribe medications to regulate menstrual cycles, induce ovulation, or manage androgen-related symptoms. Combination birth control pills or progestin therapy can regulate periods and protect against endometrial cancer, while clomiphene, letrozole, metformin, or gonadotropins can stimulate ovulation for those seeking pregnancy. For excessive hair growth and acne, options include birth control pills, spironolactone, eflornithine cream, and hair removal procedures such as electrolysis or laser therapy. Acne treatments, including topical and oral medications, are also available. In cases of infertility, procedures like in vitro fertilization may be considered. Ultimately, a personalized treatment plan, developed in consultation with a healthcare provider, is crucial for effectively managing the diverse manifestations of PCOS. **(ABOUT POLYCYSTIC OVARY SYNDROME**

(PCOS) | NICHD - EUNICE KENNEDY SHRIVER NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT)

LINK BETWEEN PCOS AND INSULIN RESISTANCE

The interplay between Polycystic Ovary Syndrome (PCOS) and insulin resistance extends beyond a simple cause-and-effect relationship, encompassing a complex web of hormonal, metabolic, and adipokine-mediated interactions.

- Pathophysiology of Insulin Resistance in PCOS:**

- While obesity undeniably amplifies insulin resistance, PCOS exhibits an intrinsic form characterized by post-receptor defects in insulin signalling. This involves impaired phosphorylation of key downstream signalling molecules within pathways critical for glucose uptake in skeletal muscle and adipose tissue.
- This inherent resistance is not merely a quantitative reduction in insulin sensitivity but also a qualitative alteration in insulin's metabolic actions.
- Furthermore, there's evidence suggesting that specific isoforms of the insulin receptor may be differentially regulated in PCOS, contributing to tissue-specific variations in insulin responsiveness.

- Hormonal Modulation of Insulin Sensitivity:**

- Hyperandrogenism, a hallmark of PCOS, directly contributes to insulin resistance. Androgens can impair insulin signalling in skeletal muscle and adipose tissue, further exacerbating metabolic dysfunction.
- The inverse relationship between insulin and sex hormone-binding globulin (SHBG) is pivotal. Hyperinsulinemia suppresses hepatic SHBG synthesis, leading to elevated free androgen levels. This creates a vicious cycle, as increased androgens further contribute to insulin resistance.
- The increased LH from the pituitary gland, stimulated by hyperinsulinemia, further drives ovarian androgen production.

- Adipokine Signalling and Metabolic Dysregulation:**

- Adipose tissue in PCOS is not merely a passive storage depot but an active endocrine organ secreting a plethora of adipokines, many of which are dysregulated.

- The balance between insulin-sensitizing adipokines (e.g., adiponectin) and insulin-resistant adipokines (e.g., resistin, TNF- α) is shifted towards a pro-inflammatory and insulin-resistant state in PCOS.
 - Adiponectin, despite its insulin-sensitizing effects, exhibits complex interactions with PCOS. Its multimeric forms may have differential effects, and its levels are influenced by both BMI and hyperandrogenism.
 - Visfatin's actions are still being studied, and its exact role in PCOS is still being determined.
- **Dietary and Environmental Influences:**
 - The observed predilection for high-sugar diets in PCOS may reflect underlying neuroendocrine dysregulation or learned behaviours.
 - These dietary patterns contribute to hyperinsulinemia and further exacerbate metabolic disturbances.
 - Environmental endocrine disruptors may also play a role in the development and progression of PCOS, influencing both insulin sensitivity and ovarian function.
 - **Clinical Implications:**
 - The heterogeneity of PCOS underscores the importance of individualized management strategies.
 - While insulin-sensitizing agents like metformin are beneficial for many women with PCOS, they may not be universally effective.
 - Lifestyle modifications, including dietary changes and exercise, remain foundational for improving insulin sensitivity and mitigating metabolic risk.
 - The understanding of the various Adipokines, and how they interact with insulin resistance, could lead to new pharmaceutical treatments. (**PURWAR AND NAGPURE**)

INS

The human INS gene, essential for insulin production in pancreatic β cells, resides within a distinctive genomic environment that influences its precise regulation. Unlike typical genes where active histone modifications cluster near the transcription start site, the INS gene displays these modifications throughout its coding region. Notably, the INS gene is part of an approximately 80 kb region, including the TH, IGF2AS, and IGF2 genes, characterized by unusually high levels of histone acetylation and H3K4 dimethylation. This islet-specific open chromatin domain implies a coordinated regulatory mechanism for these clustered genes. Expression levels of INS, TH, and IGF2 are correlated in islets, indicating shared regulatory elements. While INS exhibits the highest expression, intergenic transcripts are detectable across this region, potentially contributing to the maintenance of the open chromatin structure. This noncoding transcriptional activity, coupled with the elevated levels of active histone modifications, may cooperatively create a permissive environment for gene activation within the cluster.

The high activity of the INS promoter in islets, along with potential long-range interactions with other regulatory elements, may establish a "bystander" activation pathway, where the open chromatin domain expands outward in both directions. This suggests that the INS gene and its neighbors may function as a single entity, regulated at least partially as a group, within an open chromatin domain that both results from and contributes to the regulatory mechanism. Understanding this complex regulation is vital for deciphering normal and abnormal pancreatic β cell function and for developing targeted therapies for diabetes and related disorders.

(MUTSKOV AND FELSENFELD)

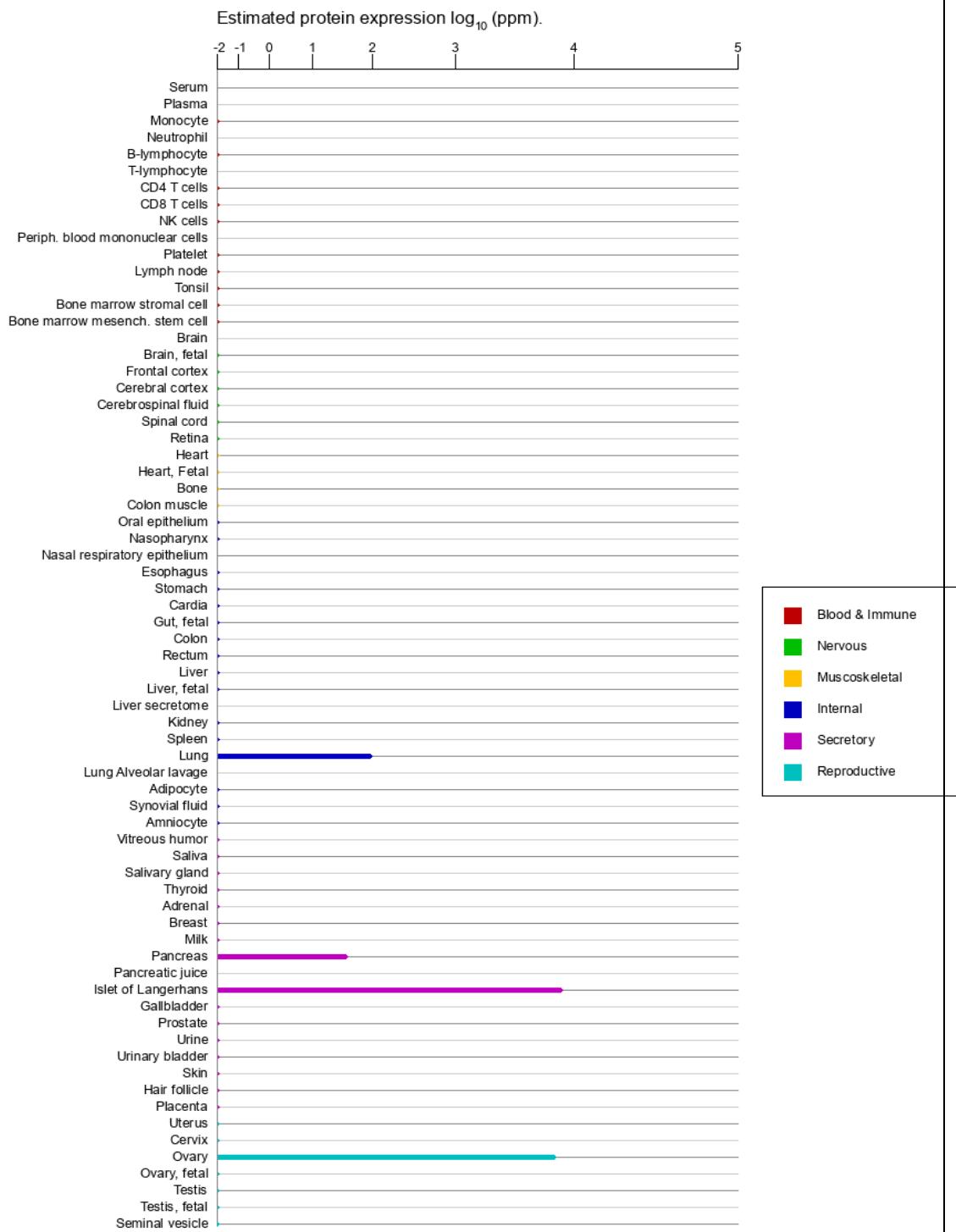


Figure 10 Protein Expression. INS chain protein expression in normal tissues and cell lines.

INSR

The insulin receptor (INSR) is a receptor tyrosine kinase that mediates the diverse actions of insulin. Upon insulin binding, INSR phosphorylates intracellular substrates like IRS1-4, SHC, and GAB1, which then recruit other signaling proteins. This activation leads to two main pathways:

- * PI3K-AKT/PKB pathway: Primarily responsible for insulin's metabolic effects, including the translocation of the GLUT4 glucose transporter to the cell membrane for glucose uptake. It also has anti-apoptotic effects and regulates gluconeogenic and lipogenic enzymes. This pathway also activates mTORC1, which regulates cell growth and metabolism.

- * Ras-MAPK pathway: Mainly involved in mediating cell growth, survival, and differentiation. INSR can also bind insulin-like growth factors (IGF1 and IGF2). Hybrid receptors containing INSR and IGF1R can bind IGF1, IGF2, and insulin with varying affinities depending on the INSR isoform. In adipocytes, INSR inhibits lipolysis.

INSR activity is regulated as follows:

- * Activation: Activated by insulin binding and subsequent autophosphorylation.

- * Inhibition:

- * Dephosphorylated by phosphatases like PTPN1, PTPRE, and PTPRF

- * Inhibited by ENPP1.

- * GRB10 and GRB14 block substrate access.

- * SOCS1 and SOCS3 bind to INSR and interfere with substrate phosphorylation.

(UNIPROT)

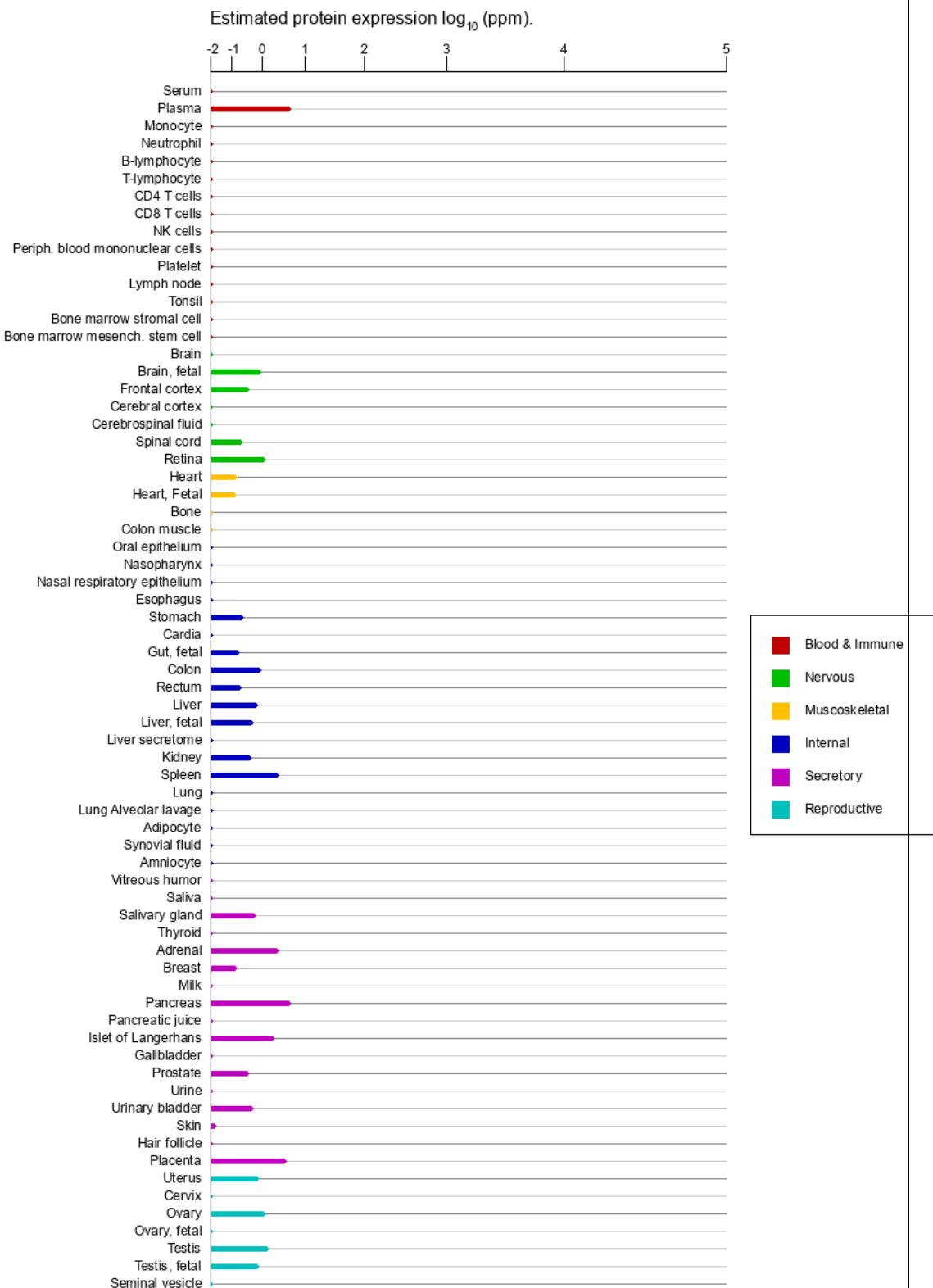


Figure 11 Protein Expression. INSR chain protein expression in normal tissues and cell lines.

PPARG

The PPARG gene, encoding the PPAR γ transcription factor, has been extensively studied for its nucleotide variations and their potential phenotypic consequences. PPAR γ regulates numerous target genes across various tissues, so genetic variants are expected to alter target gene expression. However, the exact impact of common single nucleotide polymorphisms (SNPs) on target gene expression is not fully understood.

While many studies focus on the association between PPARG variants and clinical phenotypes or biochemical markers, fewer studies directly examine the effect of these variants on PPAR γ activity itself. These studies explore alterations in PPAR γ binding affinity to PPRE, promoter efficiency, and other factors affecting transactivation.

The Pro12Ala SNP in the PPAR γ 2 isoform is the most widely studied. Functional studies suggest that Pro12Ala reduces PPAR γ 2's binding affinity to PPRE and its transactivation ability. However, in vivo studies, such as those in human adipose tissue, have shown limited differences in basal expression levels of PPAR γ target genes between Pro12Ala and Pro12Pro carriers, with some exceptions like a reduction in p85aPI3K gene expression in omental fat.

Explanations for discrepancies between in vitro and in vivo results include:

- * Differences in gene expression activation between Ala12 homozygous and heterozygous individuals.
- * Interactions between genetic and environmental factors.
- * Variations in PPAR γ 2 expression across different adipose tissue depots.

Recent research emphasizes the importance of metabolic context in modulating Pro12Ala effects. For example, gene expression changes in white adipose tissue (WAT) and muscle tissue of Ala/Ala mice differ significantly. Additionally, Pro12Ala may influence G protein function, adiponectin signaling, and cofactor recruitment.

(COSTA ET AL.)

AKT2

AKT2 is a serine/threonine kinase closely related to AKT1 and AKT3, collectively known as AKT kinase. It plays a pivotal role in regulating a multitude of cellular processes, including metabolism, proliferation, cell survival, growth, and angiogenesis. This regulation is achieved through the phosphorylation of a diverse array of downstream substrates. Notably, AKT2 is crucial for insulin-mediated glucose uptake, facilitating the translocation of the GLUT4 glucose transporter to the cell surface. It also governs glycogen storage by inhibiting GSK3A and GSK3B, promotes cell survival by phosphorylating MAP3K5, and mediates insulin-stimulated protein synthesis by activating mTORC1 signaling. Furthermore, AKT2 is involved in the phosphorylation of FOXO factors, influencing their cytoplasmic localization, and positively regulates CREB1 activity, thereby impacting the transcription of pro-survival genes. Additional functions include the potential regulation of fatty acid synthesis through ACLY phosphorylation, the inhibition of lipolysis via PDE3B activation, and the modulation of PI3P-5 activity by phosphorylating PIKFYVE. AKT2 also influences cell proliferation and growth through DLC1 phosphorylation, plays a role in adult neurogenesis, mediates the effects of various growth factors, and contributes to the antiapoptotic effects of IGF1.

Beyond these general functions, AKT2 exhibits specific roles. It is essential for SPATA13-mediated regulation of cell migration and adhesion, may be involved in placental development, and inhibits the ciliogenesis cascade in response to lysophosphatidic acid stimulation. AKT2 phosphorylates PKP1, affecting keratinocyte intercellular adhesion and PKP1 protein stability. Further, AKT2-specific substrates, such as ANKRD2, C2CD5, CLK2, and PITX2, have been identified. It may play a role in myoblast differentiation through PITX2 phosphorylation and is involved in the negative regulation of myogenesis in response to stress. AKT2 regulates insulin-stimulated glucose transport by phosphorylating C2CD5/CDP138 and participates in insulin-regulated suppression of hepatic gluconeogenesis through CLK2 phosphorylation. The activity of AKT2 is tightly regulated. Phosphorylation at Thr-309 and Ser-474 is essential for its full activation. Insulin induces this activation in adipocytes and hepatocytes. Conversely, aminofurazans, such as compound 32, act as potent inhibitors of AKT2. Insulin also induces AKT2 phosphorylation of PKP1. ([UNIPROT](#))

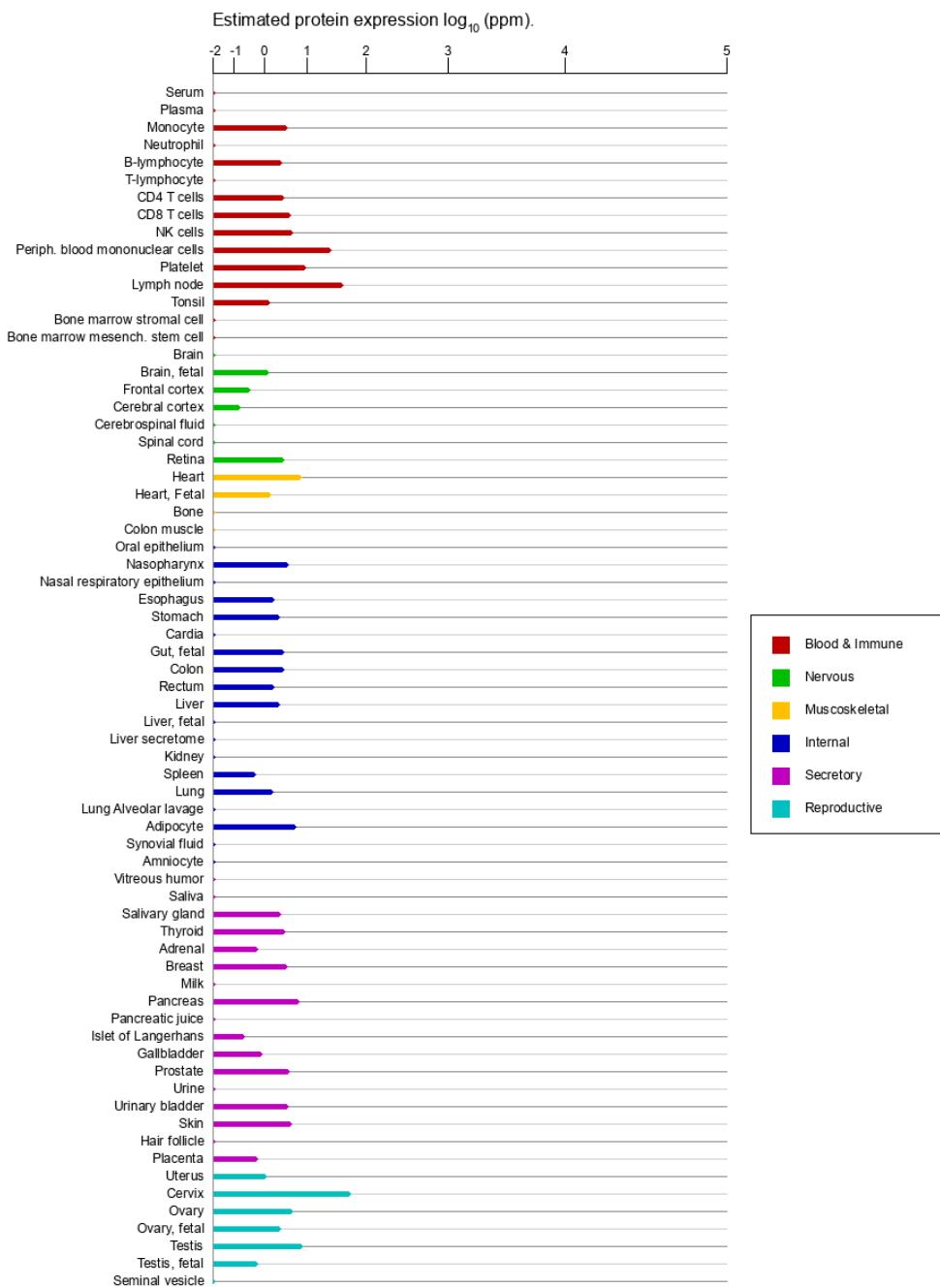


Figure 12 Protein Expression. AKT2 chain protein expression in normal tissues and cell lines.

FOXP3

FOXP3 is a master transcriptional regulator essential for the development, function, and stability of regulatory T cells (Tregs), playing a critical role in Type 1 diabetes (T1D) and celiac disease (CD). Both diseases are immune-mediated diseases characterized by complex genetic architectures. Several susceptibility loci, including *HLA*, *INS*, *CTLA4*, and *PTPN22*, have been definitively linked to T1D, while *HLA* plays a prominent role in CD susceptibility. The observation that these diseases share susceptibility genes, coupled with the increased risk of developing one disease in individuals already diagnosed with the other, strongly suggests the existence of common genetic risk factors that predispose individuals to both T1D and CD.

A critical aspect of immune regulation involves the activity of T regulatory cells (Tregs). These CD4+ T cell subsets function as negative regulators of immune responses and are essential for maintaining immune homeostasis. The *FOXP3* gene encodes FOXP3, a transcription factor that is absolutely required for the development and proper functioning of Tregs. Compelling evidence for FOXP3's importance comes from the observation that mutations in the *FOXP3* gene are responsible for IPEX syndrome (immunodysregulation, polyendocrinopathy, enteropathy, X-linked syndrome), a rare but severe recessive disorder characterized by multi-system autoimmunity, frequently including the development of T1D in infancy or early childhood. This underscores the critical role of FOXP3 in preventing autoimmunity through its influence on Treg function.

Given the suggestive linkage of both T1D and CD to the X chromosome region where the *FOXP3* gene resides, and considering the observation that T1D linkage signals frequently overlap with the *HLA*-DR3 haplotype (a major susceptibility haplotype observed in both T1D and CD), *FOXP3* emerges as a compelling candidate gene for contributing to susceptibility to both T1D and CD.

FOXP3 plays a crucial role in maintaining immune homeostasis. It functions as both a transcriptional repressor and activator, modulating gene expression through interactions with other transcription factors, histone acetylases, and deacetylases. FOXP3 orchestrates the Treg suppressive program by activating the expression of genes associated with Treg function, such as *CTLA4* and *TNFRSF18*, while simultaneously repressing the expression of genes encoding pro-

inflammatory cytokines, including IL2 and IFNG. This repression contributes to the suppression of effector T-cell function. FOXP3 also plays a role in inhibiting the differentiation of IL17-producing Th17 cells by antagonizing the function of RORC, which leads to the downregulation of IL17 expression and promotes Treg development. Furthermore, FOXP3 interacts with specific transcription factors like RUNX1 and IKZF4 to fine-tune the expression of its target genes, including IL2, IFNG, TNFRSF18, IL2RA, and CTLA4. ([UNIPROT](#)).

Type 1 diabetes (T1D) and celiac disease (CD) are both immune-mediated diseases characterized by complex genetic architectures. Several susceptibility loci, including *HLA*, *INS*, *CTLA4*, and *PTPN22*, have been definitively linked to T1D, while *HLA* plays a prominent role in CD susceptibility. The observation that these diseases share susceptibility genes, coupled with the increased risk of developing one disease in individuals already diagnosed with the other, strongly suggests the existence of common genetic risk factors that predispose individuals to both T1D and CD.

A critical aspect of immune regulation involves the activity of T regulatory cells (Tregs). These CD4+ T cell subsets function as negative regulators of immune responses and are essential for maintaining immune homeostasis. The *FOXP3* gene encodes FOXP3, a transcription factor that is absolutely required for the development and proper functioning of Tregs. Compelling evidence for FOXP3's importance comes from the observation that mutations in the *FOXP3* gene are responsible for IPEX syndrome (immunodysregulation, polyendocrinopathy, enteropathy, X-linked syndrome), a rare but severe recessive disorder characterized by multi-system autoimmunity, frequently including the development of T1D in infancy or early childhood. This underscores the critical role of FOXP3 in preventing autoimmunity through its influence on Treg function.

Given the suggestive linkage of both T1D and CD to the X chromosome region where the *FOXP3* gene resides, and considering the observation that T1D linkage signals frequently overlap with the *HLA*-DR3 haplotype (a major susceptibility haplotype observed in both T1D and CD), *FOXP3* emerges as a compelling candidate gene for contributing to susceptibility to both T1D and CD. ([BJØRNVOLD ET AL.](#))

METHODOLOGY

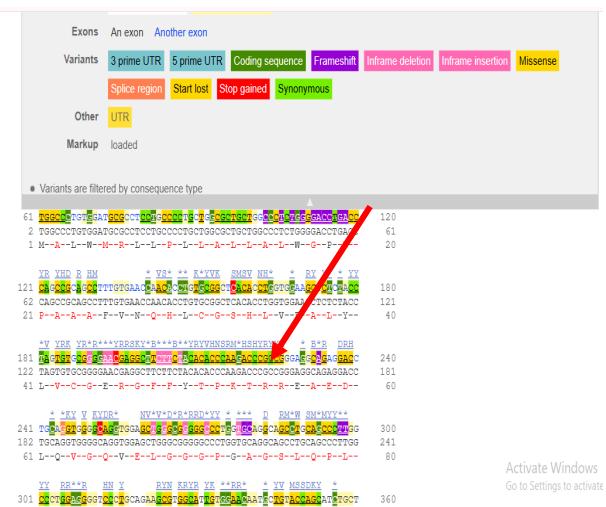
INS

VARIANT - c.147C>G (p. Phe49Leu)

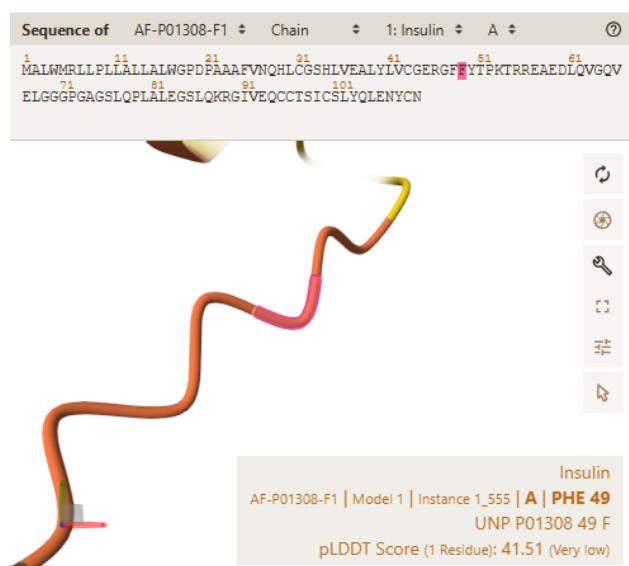
PROTEIN SEQUENCE

1 MALWMRLPLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFYPKTRREAED 60
61 LQVGVQVELGGGPAGSLQPLALEGSIQKRGIVEQCCTSICSLYQLENYCN 110

TRANSCRIPT SEQUENCE



PROTEIN STRUCTURE



VARIANT c.143T>C (p. Phe48Ser)

PROTEIN SEQUENCE

1 MALWMRLPLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERF FYTPKTRREAED 60
61 LQVQVELGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 110

TRANSCRIPT SEQUENCE

PROTEIN STRUCTURE



VARIANT c.266G>A (p. Arg89His)

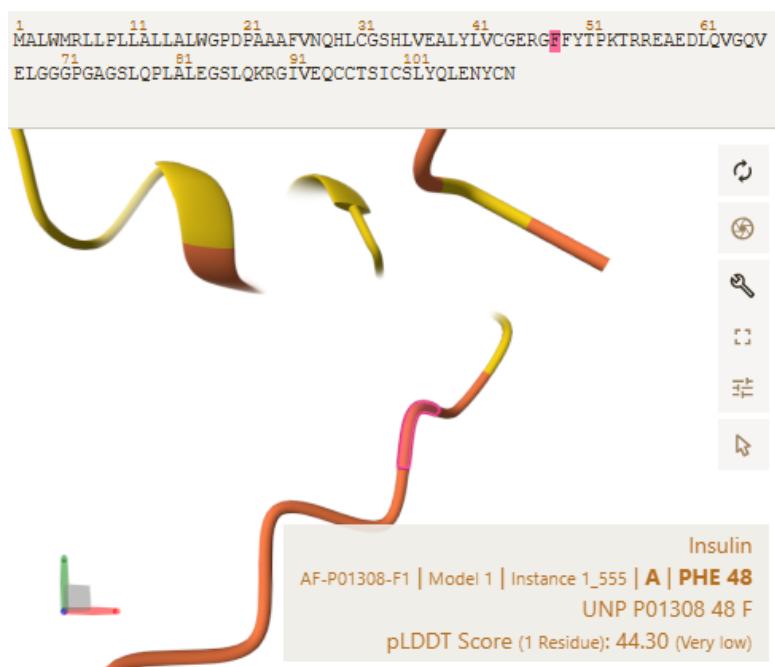
PROTEIN SEQUENCE

1 MALWMRLPLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFYTPKTRREAED 60
61 LQVGVVELGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 110

TRANSCRIPT SEQUENCE



PROTEIN STRUCTURE



VARIANT c.274G>T (p. Val92Leu)

PROTEIN SEQUENCE

1 MALWMRLPLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFYTPKTRREAED 60
61 LQVGVQVELGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 110

TRANSCRIPT SEQUENCE



PROTEIN STRUCTURE



VARIANT c.266G>T (p. Arg89Leu)

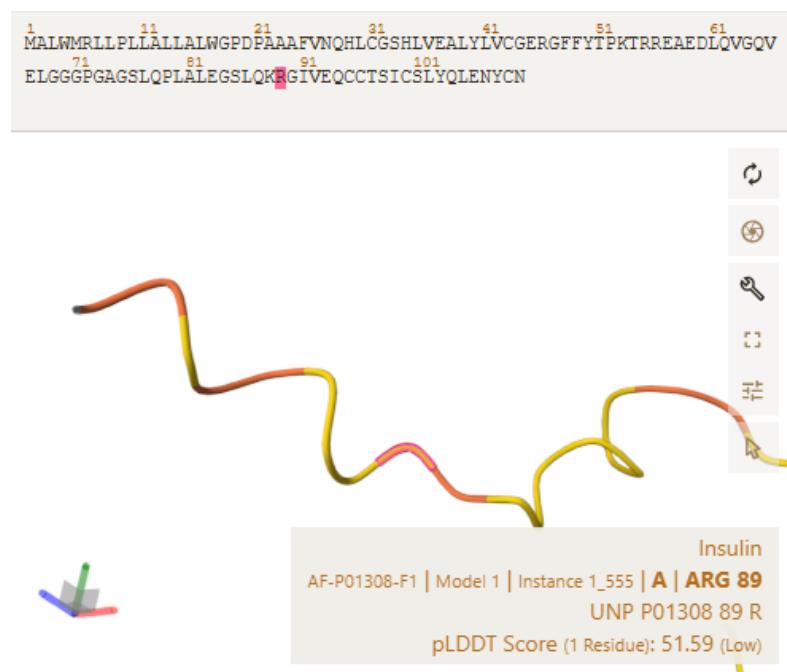
PROTEIN SEQUENCE

1 MALWMRLPLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAED 60
 61 LQV**G**VELGGPGAGSLQPLALEGS**LQK**R**GIVEQCCTSICSLYQLENYCN** 110

TRANSCRIPT SEQUENCE



PROTEIN STRUCTURE



INSR

VARIANT c.3104G>T (p. Gly1035Val)

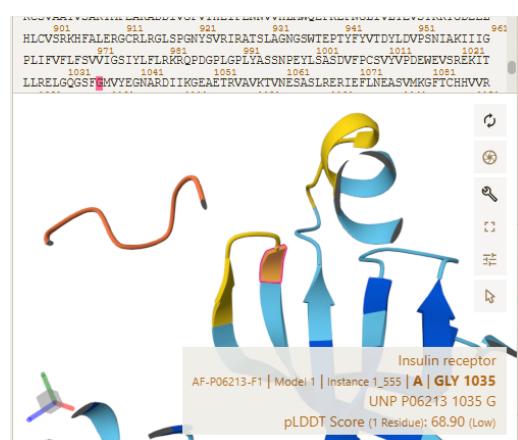
PROTEIN SEQUENCE

1 ATMTGRRRAAAPPPLVVAANLALLAGHLY PFGCWPMDNINBLTTLRHEENCSVIEHL 60
61 QILMLKQHTRKFEDPLFLSFLHMLI DDLVYRLRVEGLADLFLNLTVLNRSLRFLTNYL 120
122 VIFEMHLLGHLNLYMHLRVEKHNLCYLTADLWRSLSDEVVNLYVHNDNE 240
181 ECGDCRCATAKGKNCATPAVINGPVERCHTHSHSQCWVPCITPKHSCGACGLLCHSCS 240
241 GNC SQQPDPTTCAVTCRUFYRGLVCECT PPTFYPHQWRCWVSLQD LHHCKNSQRQ 360
301 CQHVYNNHNCICPESPGYTMNLLC TECGPVPHVCLLEGEHTJ DSATVSAQELRG 360
361 TEVNGSLIINRGNDNLAAELEALCQHNSVYIILTRASLQHPSFLKRALREI 420
421 EIGNSYFNLNDNOLRWHNSKHNTI TGQH FHHYFLCLLSEHMEVPS GTKRGKE 480
481 RND ALALKTGNCVYDLSYFISYRIS FDKLNRLLPEWYFDPDRLGMFLYR PAQY 540
541 NFEVGDGAGCSNSWTVDPLFLSDKSQHNBGWLKGKLPWJYIAVYKIVLFLYF 600
601 DERTTYKAGLNDIYVQDADTNSPVELD USVNSSSNLLWNPXKWDGPNINITYH F600
661 RQADESLFELVYCLKLKLSRSLWPSFEDS QKMRDSEASG DCCSKTSQDZL 720
721 KELEERSFSTPFETDYLNNVVFEP RTGSSTGADER PFSRRLRSGVNNTVTAEEVAF 780
781 FTTSNSTVPSFEEPERHNSVLSI SGLHRTGYTQE LQACNQHCFTEVVA 840
841 SARTHEMPLGGDAPVGTI HEIENPAGLWQHME PKENGLVLT VVSYRVEGGEIHL 900
901 SRKHFALERGRCLLGSNPNVSRVTR SLAGSHTYE FTFVYI DVLWPIAKIIG 960
961 PLI FFLVFSVLSYFSLRQRPDGELGFLASNSPEYLSVAS CFVSVPEDEWV03 1020
1020 EKTIHLRELQGSGS YMNEYARNDI IGEKARTA VAKVTKNSAESLRLERI FLENAWSVNG 1080
1081 FCTHVRVBLWPSVQSKQFLVMLMHNADLSKYLSPVWQHDFRPPFTLQEMIQA 1140
1140 AEADOGMAYNRAKWHVBDLAAARNMHDFTVWIKGD FGMTRDIVEY TRYNGKKGLD 1200
1200 RWHAFES DLEGFTVTTSSPFSGVWLNEIQLQGPGLSNQEWFLVMDGQYDQ 1200
1261 CPER DTLMWCNSQCPHNPBMRFLWELNLDQDLPFS FVPS FFSNEAEPESELE 1320
1320 FEIDMENYFLDRSNSHCQEAGRGDSGSLGFKRSYEE HIIYTHMMGKGHRNLLPRLR 1380
1380 PS 1382

TRANSCRIPT SEQUENCE



PROTEIN STRUCTURE



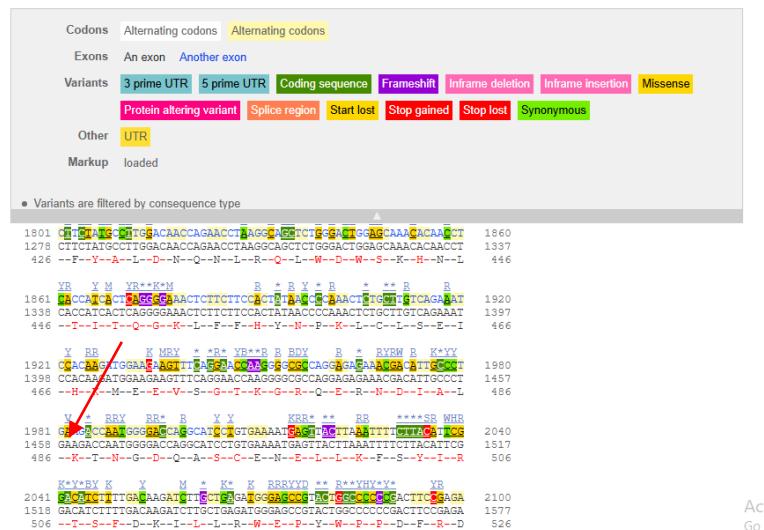
VARIANT c.1459A>G (p. Lys487Glu)

PROTEIN SEQUENCE

```

1 MATGGRRGAAAPLLVAVAALLLGAAGHLYPGEVC POMDIRNNNTRLHELENCVIEGHL 60
61 QILLMFKTRPEFDRLSFFKLIMITYDYLIFRVYGLESLKDFPLNLTVRSRFLFVNAYAL 120
121 VIFEMVHLREIGLYNLMITRGSVRIEKNELCYLATIWSRILDSVEVDNTIVLNMKDN 180
181 ECGDICPGTAGKTCNPATVINGQFVERCWTSHCQVCFTICKSHGCTAEGLCCHSECL 240
241 DAKLWVQKQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQ 300
301 CHQYVIIHNNKICPECPSCYTMSSNLLCTPCLGPCHVCLLSEKITLDSTISACELRCG 360
361 TIVNGSLIINIRGGRNLLAELEANGLLIEEISGYLKRARSVALVLSLSFRKLRLRSEI 420
421 EIGHYSYFALDQNQLRQLWWSKHKHLLTTOQKLFLHHVFLKLLSEIHOMEVSGTKRQE 480
481 RNDLWVQKQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQ 540
541 NVTEDFDQDAGGSNSWIVDIDPPLSNPDPSNHNPGLMRLKPKTQYALFYVKTLVTF 600
601 DERRTYGAKSDII XVQTDATAN SVFLDOPISVNSQSSLILWKGKPSDPNGHIIHVVLVFM 660
661 RQAEDSELFELDYCLAGKLPSRTWSPPFSEEDSQKHQSEYEYDSAGECCSCPFTDSQIL 720
721 KEELESSFRKTFRDYLHNVFVTRKTSSTGAEDPRSRKRSLSLDGVGVAVATEVRAAF 780
781 PNTTQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQVYVQ 840
841 SARTTHTPGRDIDVQEPPTVIEIENNNVHISQEGPQVQVYVQVYVQVYVQVYVQVYVQ 900
901 KKHVFALECRCLRLGSLPCHNSVRIDASLGNQSHTEPFTTYVTDYLDVPSIHKLIIIG 960
961 PL1FVFLFSVVI6SIVYLFLRQPDGLPGLPYASSNFEYLSSASWVFCVSVVDEMEVSR 1020
1021 EKTTLRLERLQQSGFMVYEGNARDIKEGEETRVAVAKTYVNEASASLRERIEFLNEASVMHG 1080
1081 FCTCHVVRLLGVVSKGGPFTLVVMEIAGDLKSYLSLRLPEAANNPGRRPPFTLCEMIQMA 1140
1141 AI IADGMAYLNAGKGVVRDLAARNVVAHDFTVKGIGFGMTDIYETDVRGGKQULLPV 1200
1201 RWHAPEFSLKDGVFTTSS INMSFGVVWLWEISLAEQPYQGLSNEXQVQLKFMDDGVLQPN 1260
1261 RVTDLMRWWQFNPNMFTFLEVNLKLDLHPSFPEVWVFFHSEENGAPSEELME 1320
1321 PEDMENVPLDRSSHCQREAAAGRDDGS SLGFKRSYEEHPIYTHMMNGKRGNGRILTPRSN 1380
1381 PS 1382
|
```

TRANSCRIPT SEQUENCE



PROTEIN STRUCTURE



VARIANT c.1177G>A (p. Gly393Arg)

PROTEIN SEQUENCE

```

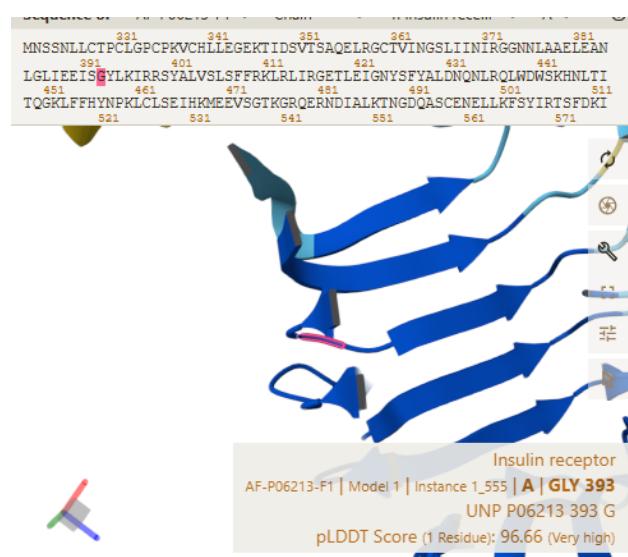
1 MATGGRRGAAAPLLVAVAALLGAAGHLYPGEVCPGMDIRNNLTRLHELENCSVIEGLH 60
61 QILLMFKTRPDEFLSFLKLMITDYLRLRIVYGELESLKULFNLTVIIGSRLFNHYAL 120
121 VIFEMWHLREGLINLMNITINGSVRIERKNELCYLATTIDWSRILDSVEVDNYIVLNHDDNE 180
181 ECGDCIPGTAKGKTCATVINGQFVERCWHSHQKVCPCTICKSHGCTAEGLCCSECL 240
241 GNC SQQDPPTKCVACRGRYLIGRCVETCPPYYHDMRCVNFSFCQDLHHRKNSRRQQ 300
304 CHQYVTHNNCICPECPGTTINS SNLLCTPCGSPCRVCHLLEGEK109VTSQAELRGCG 360
361 TVNGSLIINRGGNNLAELAEANLGLIEEIS YLIRIRSTAYLVLSSLFFRKRLIRGETL 420
421 EIGAFSTYALDQNQQLWQWWDQWQWQWQWQWQWQWQWQWQWQWQWQWQWQWQWQWQ 480
482 ANDTAAAGTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATG 540
541 NEDFDODACCSHSVTVDDPLPSLSDNQVQNHFGHIMQHGWVQHWTQVADLWV 600
601 DERTYVGKNDI11VQTDATPSVLDLFSVSNSSQJLILMWPFPSSDWGNNTVHIVWNE 660
661 RQADESELFLDVCCLGKLERTHTSPFTESDQGHMNOSEYEDEACECCSCBPKIQSIL 720
721 KELAESSEFRGTFEVLYHNVWNEKTS5GCTGEDRBSRBRZLGVNGHTVAVTPVDAF 780
781 PNTSITSPVSEEEHRRPPEKVNKESLV1SGLRRTGYRIELACQNDTPEERCSVAAYV 840
841 SARMPTEKAADDIVGPVTHEIENNNVHLWQPKPNWGLIVLVEVSYRVYDQEELHLBCV 900
901 SRKHFALERCGRCLGLSPGNVSVIRATISLAGNSSTEPTFYVVDLIDVENSIAIIIG 960
961 FLLIVFLFWVIGSITYLFLRKRPGFLGLFLVASNPVYLGSADWFFCSVVVFWDEWEVR 1020
1021 EKITLRELGGSGFGMVYEGNARDIIGEAETAVVKVTWESASLRLERIEFLNEASVMG 1080
1081 FTCHDRVLLVVSRSQFPTLVVMEIMAHGDLKSYLRSRFEANNPGRPPPTLQEMOMQ 1140
1141 AEIAOGMAYNAKMKPHRDLAARUCHMVAHDFTVKLGDFGHTRDIVETDYYRKGGHGLPV 1200
1201 RHMAMESLKDGWTTTSSSSSFGVWLEITSLAEQPYQQLSNEVWKPMWMSGYLQDPD 1260
1261 CPERVTDLMRMWCQFNPKMPTFLEVNLLKDLRPSFPEVFFFHSEENKAPESELEME 1320
1320 FEDMENVFLDRSSHQRREEAGGRDGGSLSGFKRSYEEHIPYTHMNNGKKRNGRILTLPRSN 1380
1381 PS

```

TRANSCRIPT SEQUENCE



PROTEIN STRUCTURE



PPARG

VARIANT c.362A>G (p. Tyr121Cys)

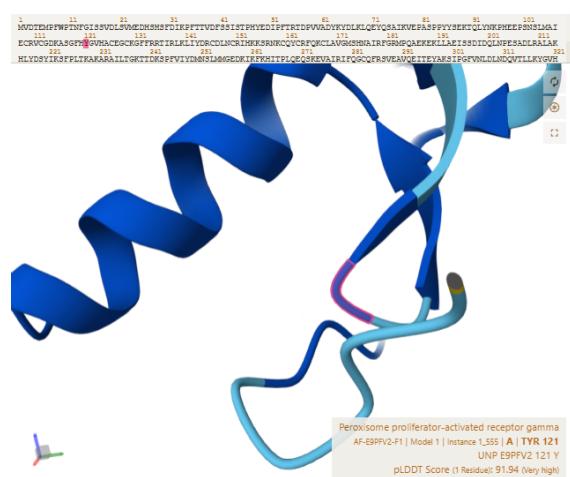
PROTEIN SEQUENCE

1 MVDTEMPFWPTNFGISSVDSLVMEDHSHSFDIKPFTTVDSSISTPHYEDIPFTRTDPPV 60
61 ADYKYDLKLQEYQSAIKVEPASPPYYSEKTQLYNKPHEEPSNSLMAIECRVCVGDKASGFH 120
121 CGHVACEGCKGFRRRTIRLKLHYDRCDLNCRIHKKSRNKQCYCRFKQKCLAVGMSHNAIRF 180
181 GRMPOAEKEKLLEISSLIDQLNPESADLRLAKHLYDSYIKSFPLTKAKARAILTGKTT 240
241 DKSFPVFIYDMNSLMMGEDKIKFKKHITPLQEQSKEVAIRIFQGCQFRSVEAVQEITEYAKS 300
301 IPGFVNLDLNDQVTLLKGVHEIIYTMLASLMNKGDVLISEGQGFMTRFELKSLRKPFGD 360
361 FMEPKFEFAVKFNAALELDDSDLAIFIAVIILSGD~~RPG~~LLNVKPIEDIQDNLLQALELQLK 420
421 LNHPESSQLFAKLQLQKMTDLRQIVTEHVQLLOVKKTEDMSLHPLLQEYIKDLY 475

TRANSCRIPT SEQUENCE

Codons	Alternating codons	Alternating codons	11				
Exons	An exon	Another exon					
Variants	3 prime UTR	5 prime UTR	Coding sequence	Frameshift	Inframe deletion	Inframe insertion	Missense
			Protein altering variant	Splice region	Start lost	Stop gained	Synonymous
Other	UTR						
Markup	loaded						
• Variants are filtered by consequence type							
	B R M R M D B K P * Y V B H R Y D K K H Y S R B						
361	T C C G A T G E C G A T C T A G C G G A C C T G C T C G C T A T A T G T G A G A G G C T C G						420
214	T A C C A A R G T G C A A T C A A A G T G G A G C C T G C A T C T C C A C C I T A T T A T T C T G A G A A G C T C A G						273
72	-T -Q -S -A -I -K -V -E -P -A -S -P -P -Y -Y -S -E -K -T -Q -						91
	Y S B R R M S Y Y * V Y R V Y H B S S R H R * * Y Y R R *						
421	C T G C A T A G C T A A G A G G C T T C C A C G C T C G C T A T S C C A T T C H A T G T G S H C						480
274	C T C T A C A T T A A G C C T C A T G A G A G C C T T C C A C T C C T C A T G G C A A T T G A A T G T C G T G I C						333
92	-L -Y -N -K -P -H -E -E -P -S -N -S -L -M -A -I -E -C -R -V -						111
	* B R K Y * R * R X K F N W * K Y Y R V * * K Y						
481	I T G C A T A A G A G C T C T C A T T C A C T G C A T G A G A G C T C A T G A G G C A T C A G G G						540
334	T T C G G A G A T A A G C T T C T G G A C T A C T A G A T G G A G T C A T G C T T G T G A A G G A T G C A A G G G I						393
112	-C -G --D -K -A -S -G -F -H -Y -G --V -H -A -C -E -G -C -K -G -						131
	K * * Y Y R K S R R S * D K R M S R K S R R Y R * Y M S R						
541	H C D C G A G A C T A T A G T G A C T A T G A C A T G A T G C A T T A A C T G T C G						600
394	T T C T C C G G A G A C A T C A G A T T G R A G C T A T T C A T G A C A G A T G T G A T C T T A A C T G T C G G						453
132	-P -F -F -R -R -T -I -R -L -K -L -I -Y -D -R -C -D -D -L -N -C -R -						151
	Y M R * * * * Y R B D M * * * Y R D * * * Y R R B						
601	A T C C G A A A A A A G A S A A A T A G T G C A G T A C T C C G G T T T G A G A A T G C T I G C G T G						660
454	A T C C A C A A A A A A A G T G A A A T A A T G C T A G T G C T G G T T T C A G A A A T G C C T T C A G T G						513
152	-I -H -K -K -S -R -N -K -C -Q -Y -C -R -F -Q -K -C -L -A -V -						171

PROTEIN STRUCTURE

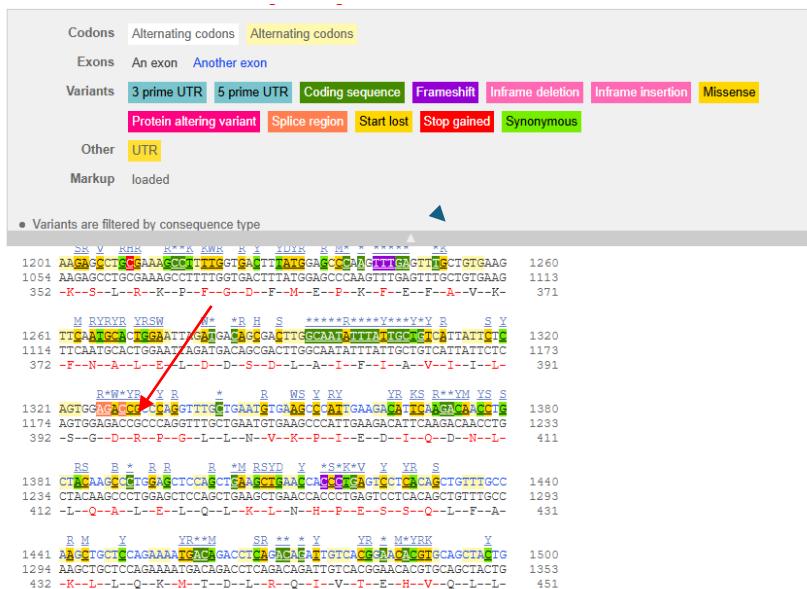


VARIANT c.1184G>A (p. Arg395His)

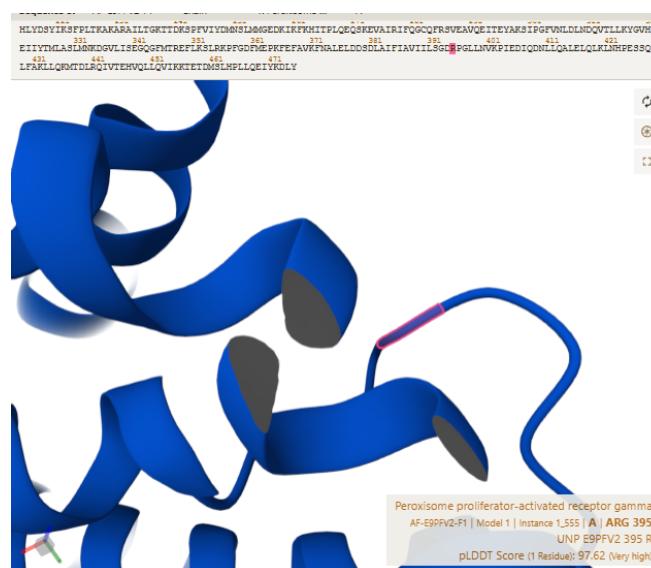
PROTEIN SEQUENCE

```
1 MVDTEMPFWPTNFGISSVDSLVMEDHSHSF DIKPFITVDFSSISTPHYEDIPFTRTDPPV 60
61 ADYKYDLKLQEYQS A I K V E P A S P P Y Y S E K T Q L Y N K P H E E P S N S L M A I E C R V C G D K A S G F H 120
121 YGVHACEGCKGFFRRTIRLKLHYDRCDLNCRHKKSRNKCQYCRFQKCLAVGM SHNA I R F 180
181 GRMPQA E K E K L L A E I S S D I D Q L N P E S A D I R A L A K H L Y D S Y I K S F P L T K A K A R A I L T G K T T 240
241 D K S P F V I Y D M N S L M M G E D K I K F K H I T P L Q E Q S K E V A I R I F Q G C Q F R S V E A V Q E I T E Y A K S 300
301 I P G F V N L D L N D Q V T L L K Y G V H E I I Y T M L A S L M N K D G V L I S E G Q G F M T R E F L K S L R K P F G D 360
361 F M E P K F E F A V K F N A L E L D D S L A I F I A V I I L S G D H P G L L N V K P I E D I Q D N I L Q A L E L Q L K 420
421 L N H P E S S Q L F A K L L Q K M T D L R Q I V T E H V Q L L Q V I K K T E D M S L H P L L Q E I Y K D L Y 475
```

TRANSCRIPT SEQUENCE



PROTEIN STRUCTURE



AKT2

VARIANT c.821G>A (p. Arg274His)

PROTEIN SEQUENCE

```

1 MNEVSVIKEGWLHKRGEYIKTWRPRYFLLKSDGSFIGYKERPEAPDQTLPLNNFSVAEC 60
61 QLMKTERPRPNTIVRCLQQTIVIERIFVFDVSDFDERKEENWRAIQMVANSLQRAFGEDPM 120
121 DYCNGSPSDSTTEEMEVAVSKARAKVIMMDFDYLKLLKGTFGKVILVREKATGRYYAM 180
181 KILRKVEVIIAKDEVAHVTIESRVLQNTRHPLTALKYAFQTHDRLCFVMEYANGELFFH 240
241 LSERERVTEERARFYGAEVISALEYLHSRDVVYRDIKLENIMLDKDHIKITDFGLCKEG 300
301 ISDGATMKTFCGTPEYLAPEVLEDNDYGRAVDWGGLGVVIMEMCGLLPFVQHQHERLFE 360
361 LIIMEEIRFFRTLSPEAKSLLAGLLKDPKQRLGGPSDAKEVMERFLSINQDVVK 420
421 KLLPPFKPQVTSVDFTRYDEFIAQSITITPPDRYDLSGLLELDQRTFPQFSYASIR 480
481 E

```

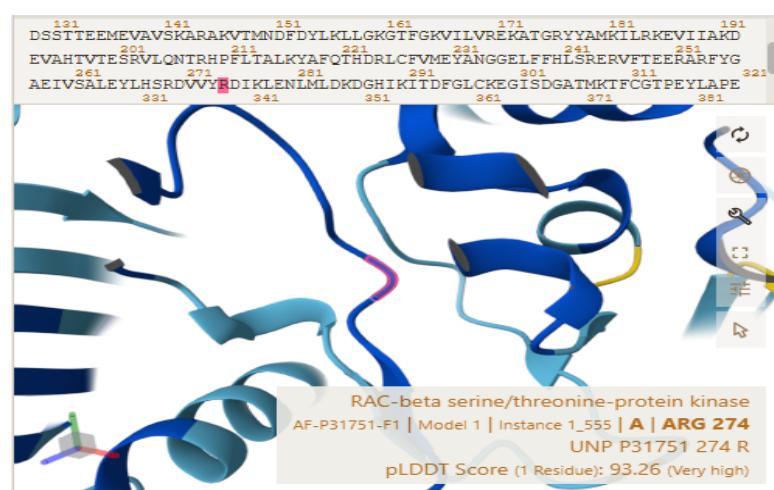
TRANSCRIPT SEQUENCE

Codons Alternating codons Alternating codons
 Exons An exon Another exon
 Variants 3 prime UTR 5 prime UTR Coding sequence Frameshift Inframe deletion Missense Splice region
 Start lost Stop gained Synonymous
 Other UTR
 Markup loaded

• Variants are filtered by consequence type

*****Y*SMITYSWYRHHDYD	RY	**	RY	Y	*RVBRRK	K	*					
901	GCGCCTG AACCAAGACCCG TGTCG TTTGTTG ATGGG GT TGCGA CGG GCG GA GACTG	960										
652	GCCTTCCAGACCCAGACCGCCTGCTGCTTGTGATGGAGATATCCAAACGGGGTGACCTG	711										
218	-A-F-Q-T-H-D-R-L-C-F-V-M-E-Y-A-N-G-G-E-L-	257										
	Y	Y	R	YRK	DYRM*	*						
961	TGCTTCCACG TGCCG GACCGTC CTTG ATCGA GA ATCGG CCCG TTTA GG CGA	1020										
712	TCTTCCACCTGTCGGAGCGCTGCTTACACAGGAGCGGCCGTTATGGTGC	771										
238	-F-F-H-L-S-R-E-R-V-F-T-E-E-R-A-R-F-Y-G-A-	257										
	R	*	YRB	Y	RW	YY	Y	HIDNR	*HR*	R****YRBD	*	
1021	GAGA T TCTG GGG CTCTG CT GTG AT GTG CA TGCG TA CGC TA CGT AT CGA	1080										
772	GAGATTGTCGCGCTCTGAGTACTGCACTCCCGGAGCTGTTACCGGAGCATCGAAG	831										
258	-E-I-V-S-A-L-E-Y-L-H-S-R-D-V-V-Y-R-D-I-K-	277										
	K	SS	Y	RR	W	KMY	*	RRR	*	Y	S	R
1081	CTGGAAN CG CTGCT GT CAAGAT GG CGACAT GT CGT ACT TTGG T TC	1140										
832	CTGGAACACTCATGCTGGACAAGATGGCCACATCAAGATCACTGACTTTGGCTCTGC	891										
278	-L-E-N-L-M-L-D-K-D-G-H-I-K-I-T-D-F-G-L-C-	297										
	K	RVR	YR	BDR	VYY	SYRY	Y	YY	*B	BMSVYSR		
1141	AAAG GG CGCA AT GGGACCG GG CCG CD AAACAC T GTG GG CGCG GT TACCTG	1200										
892	AAAGGGGACATCAGTGACCGGGCACCAAGAACCTCTGTTGGACCCGGGTACCTG	951										
298	-K-E-G-I-S-D-G-A-T-M-K-T-F-G-P-E-Y-L-	317										

PROTEIN STRUCTURE



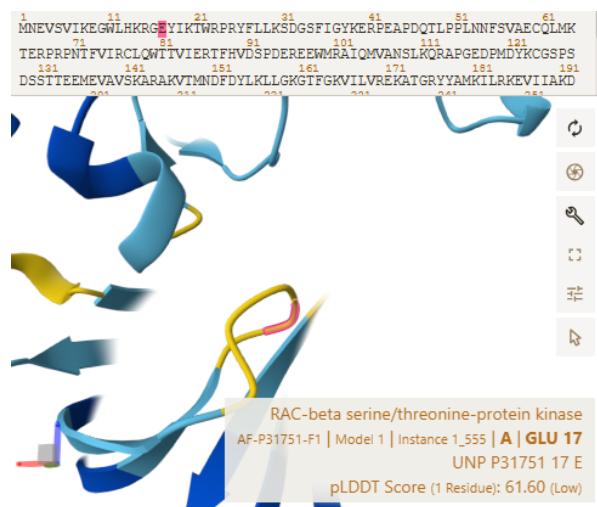
VARIANT c.49G>A (p. Glu17Lys)

PROTEIN SEQUENCE

1 MNEVSVIKEGWLHKRG**E**YIKTWRPRFYFLKSDGSFIGYKERPEAPDQTLPPNNFSVAEC 60
61 QLMKTERPRPNTFVIRCLQWTTVIERTFHVDSPDR**E**EEWMRAIQMVANSLKQRAPGEDPM 120
121 DYKGCGPSDSSTTEMEEVAVSKARAKVTMNDFDYLNLGLKGTFGKVILVREKATGRYYAM 180
181 KILRKKEVIAKIADVEAHTVTEESRVLQNTRHFLTAALKYAFQTQHDLRCFVMEYANGELFFFH 240
241 LSRRVFTEERARFYGAEIVSALEYLHSRDVDVYRDIKLENLMLDKDGHKITDFGLCKEG 300
301 ISDGATMKFCGTPEYLAPEVLEDNDYGRAWDWWGLGVMMYEMMCGRLPFYQNQDHRLF 360
361 LILMEEIRFPRTLSPEAKSLLAGLKKDPQKR~~L~~GGGPDAAKEVMEMHRFFLISINWQDVVK 420
421 KLLPPFKPKQVTSEVDTRYFDDEFTAQSITITPPDRYDSLGLLELDQRTHFPQFSYSASIR 480
481 E 481

TRANSCRIPT SEQUENCE

PROTEIN STRUCTURE



VARIANT c.1112G>A (p. Arg371His)

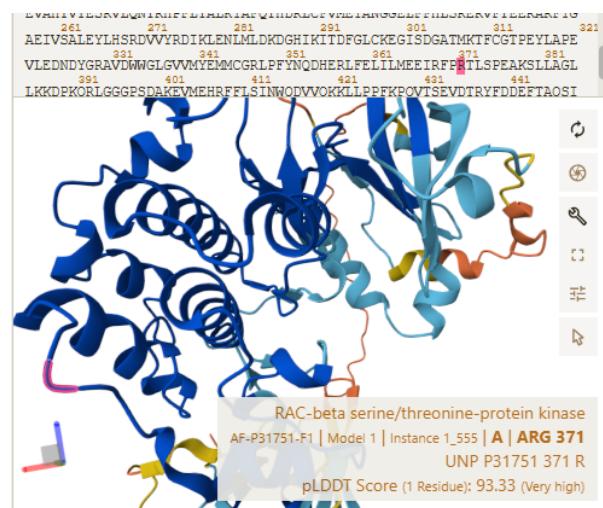
PROTEIN SEQUENCE

1 MNEVSVIKEGWLHKRGEYIKTWRPRYFLLKSDFSGSFIGYKERPEAPDQTLPPLNNFSVAEC 60
61 QLMKTERPRPNTFVICRLQWTTIVERTFHFSVPDEREEWMRAIQMVANSLSQRKAPGEDPM 120
121 DYKCGPSDSSTTEEMEVAVSKARAKTMDDFDYLKLGGKTFGVKVILVREKATGTRYVAM 180
181 KILRKEVIAKDEAVHTTESRVLNTRHFFLTALKYAFQTHDRLCFVMYEANGELFFF 240
241 LSRRERVTEERARFYGAETIVSALEYLHSRDVVYRDIKLEMMLDKGDHGKIKITDFGLCKEG 300
301 ISDGATMKTFCGFTPEYLAPEVLEDNDYGRADVNDWGLGVVMMYEMCMGRLFVNQNDHERLF 360
361 LILMEEIRFPFLISPEAKSLLAGLKKDPKQRLLGGGPSDAKEVMEHRRFLSINWQDVVKQ 420
421 KLLPPFKPKQVTSVEVDTRYFDDDETAQSITTPPDYRDSLGLLELDQRTHFQFSYSASIR 480
481 E 481

TRANSCRIPT SEQUENCE



PROTEIN STRUCTURE



VARIANT c.904A>G (p. Ser302Gly)

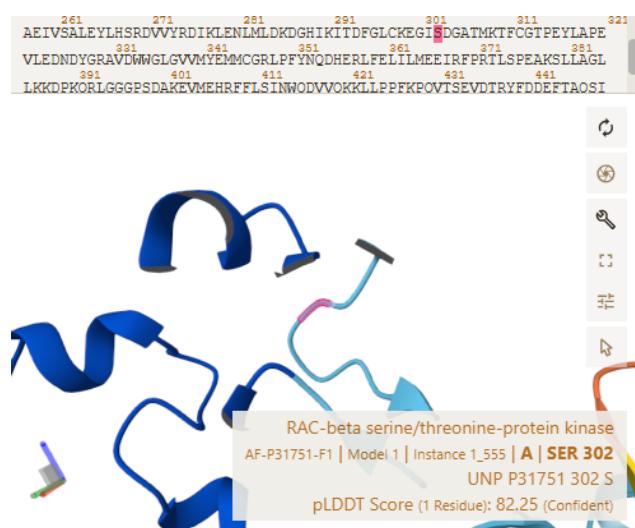
PROTEIN SEQUENCE

1 MNEVSVIKEGWLHKRGEYIKTWRPRYFLLKSDGSFIGYKERPEAPDQTLPPLNNFSVAEC 60
61 QLMKTERRPNTFVICLQWTIVIERTFHVDSPDEREEWMARIAQMVANSLSQRAGPEDPM 120
121 DYKCGPSDSSTTEEMEVAVSKARAKTVMKLLKGKTFGKVILVREKATGTRYVAM 180
181 KILRKVEIIAKDEAVHTTESRVLNTRHFFLTALKYAFQTHDRLCFVMYEANGELFFF 240
241 LSRRERVTEERARFYGAETVSAYEYLHSLRDVVYRDIKLENLMLDKGDGHKITDFGLCKEG 300
301 ISDGATMTKFCGTPEYLAPEVLEDNDYGRAVDWWNLGLVVMYEMMCGRLPFYQNQDHRLFE 360
361 LIILMEEIRFPRTLSPEAKSLLAGLKKDPQRLLGGGPSDAKEVMEHRFFLSINWQDVVKQ 420
421 KLLPPFKPKQVISEVDTRYFDDEFTAQSISITPPDRYDSLGLLELDQRTHPQFYSASIR 480
481 E 481

TRANSCRIPT SEQUENCE



PROTEIN STRUCTURE



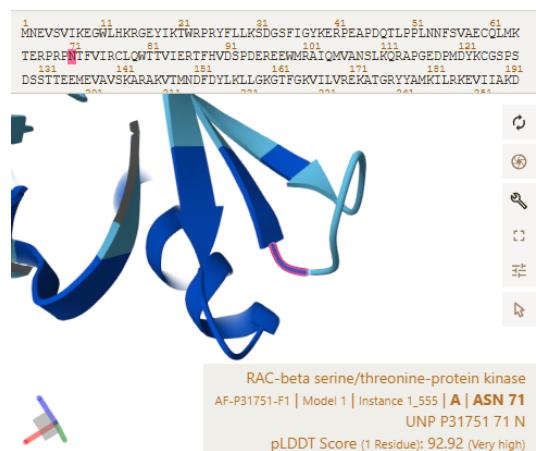
VARIANT c.212A>G (p. Asn71Ser)

PROTEIN SEQUENCE

1 MNEVSVIKEGWLHKRGEYIKTWRPRYFLLKSDGSFIGYKERPEAPDQTLLPPLNNFSVAEC 60
61 QLMKTERPRPNTFVIRCLQTTWVIERTFHWDSPDREREWMRAIQMVANSLKQRAGPEDPM 120
121 DYKCGPSDSSTDTEEMEVAVSKVAKTMDNFDYLKLLKGKTFGKVILVREKATGRYYAM 180
181 KILRKEVIAKDEAVHTTESRVLQNTRHFFLTALKAYQFTHDRLCFMEYANGELFFFH 240
241 LSRRERVTEERARFYGAIEIVSALEYLHSRDVVYRDIKLENLMLDKDGHKITDFGLCKEG 300
301 ISDGATMKTFCGTFEYLAPEVLEDNDYGRADVWGLGVVMYEMCMGRLPFYQNDQHFERLFE 360
361 LILMEEIRFPRTLSPEAKSLLAGLKKDPKQRLLGGPSDAKEVMEMHRFFLSINWQDVVK 420
421 KILPPFKPKQVISEVDTRYFDDETAQSISITTPDRYDSDLGLLELDQRTFHQPFSYASIR 480
481 E 481

TRANSCRIPT SEQUENCE

PROTEIN STRUCTURE



FOXP3

VARIANT c.224C>T (p. Pro75Leu)

PROTEIN SEQUENCE

```
1 MPNRPKGPSAFLALGPSPGASPWSRAAPKASDLLGARGPGGTQQRDLRGGAHASSSS 60
61 LNFMPPSQLQLPTLVMVAPSGARLGPLPHQLQALLQDRPHFMHQLSTDAHARTPVLVQV 120
121 HPLESPAMISLPTPTATGVFSLKARPGLPEGIVASLEWVSREPALLCTFPNPSAPRKD 180
181 STLSAVPQSSYVPLANGCKWPCEVKFEEPEDFLKHCQADHLLDEKGRAQCLLQREMVQ 240
241 SLEQLVLKEKEKLSMASQAHLAGMALTKASSVASSDKGSCCIVAAGSQGVVPAWSGPR 300
301 AFDSLFAVRHLGSHGNSTFPEFLHNMDYFKFHNMRPFTYATLIRWAILEAPEKQR 360
361 NEIYHWFTRMFAFFRNHPATWKNAIRHNLSLHKCFVRESEKGAVWIVDELFRKRSQR 420
421 PSRCSNPTPGP 431
```

TRANSCRIPT SEQUENCE

Codons Alternating codons Alternating codons

Exons An exon Another exon

Variants 3 prime UTR 5 prime UTR Coding sequence Frameshift Inframe deletion Missense Splice region

Start lost Stop gained Stop lost Stop retained Synonymous

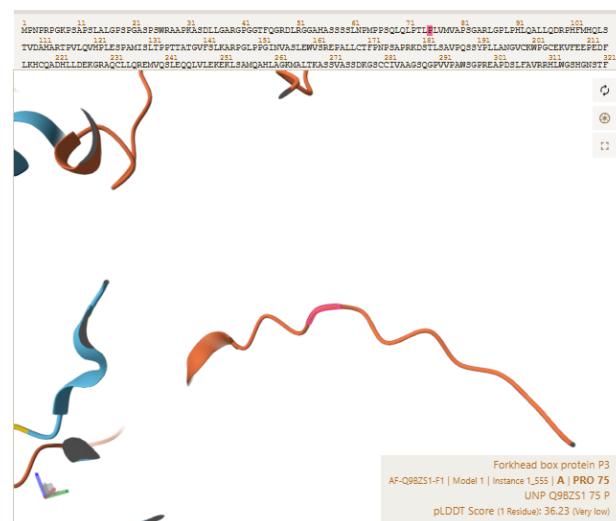
Other UTR

Markup loaded

• Variants are filtered by consequence type

```
121 DN CCCAGCTCCAGGTCGCAGCTGGAGGCTCACCCAAAGCCTCAGTTCTTG 180
49 GGCCCATCCCCAGGGCCTCGCCCAGCTGGAGGGCTGACCCBAACCCTAGACCCTG 108
17 -G->-S->-F->-A->-S->-P->-S->-W->-R->-A->-A->-P->-K->-A->-S->-D->-L->-L-
      DRN VV R * *R*B**RRS K K YR K *R K R*Y * Y ** 37
      DRN VV R * *R*B**RRS K K YR K *R K R*Y * Y ** 180
181 GGGGCCCGGGGGGTGGCAGCTGGAGGCTCACCCAAAGCCTCAGTTCTTG 240
109 GGGGCCCGGGGGCCCAGGGGGACCTTCAGGCCCAGGGCCGGGAGATCTTCGAGGCCGGGCCATGCC 168
37 -G->-A->-R->-G->-P->-G->-T->-P->-Q->-G->-R->-D->-L->-R->-G->-G->-A->-H->-A-
      *VVVY VV K Y ***S P YD S * SY S Y D VVS 56
      *VVVY VV K Y ***S P YD S * SY S Y D VVS 180
241 GCCTCCAGGTCGCAGCTGGAGGCTCACCCAAAGCCTCAGTTCTTG 300
169 TCCTCTCTCTCTTGAACCCCATGCCACCATCCAGCTGCCAGCTGCCACACTGCCCTTA 228
37 -S->-S->-L->-N->-P->-M->-P->-S->-Q->-L->-Q->-T->-L->-P->-L-
      * K YD Y SHDS YD R Y S * K MM* YH * H 76
      * K YD Y SHDS YD R Y S * K MM* YH * H 180
301 TCCAGGCCCGGGTGGCAGCTGGAGGCTCACCCAAAGCCTCAGTTCTTG 360
228 GTCATGTGGCCCAGGTGGCAGCTGGAGGCTCACCCAAAGCCTCAGTTCTTG 288
77 -V->-R->-P->-S->-A->-R->-L->-Q->-P->-H->-L->-Q->-A->-L->-L-
      R S B * RY * * M YD K RMY *HD* YR 96
      R S B * RY * * M YD K RMY *HD* YR 180
361 CAGCAGCCCAGGTGGCAGCTGGAGGCTCACCCAAAGCCTCAGTTCTTG 420
289 CAGACAGCAGCACATTCATGTGACCACGCTTCACCCAAAGCCTCAGTTCTTG 348
97 -Q->-D->-P->-H->-M->-H->-Q->-L->-S->-T->-V->-D->-A->-H->-A->-R->-T->-P->-G 116
      A G
```

PROTEIN STRUCTURE



VARIANT c.1189C>T (p. Arg397Trp)

PROTEIN SEQUENCE

```

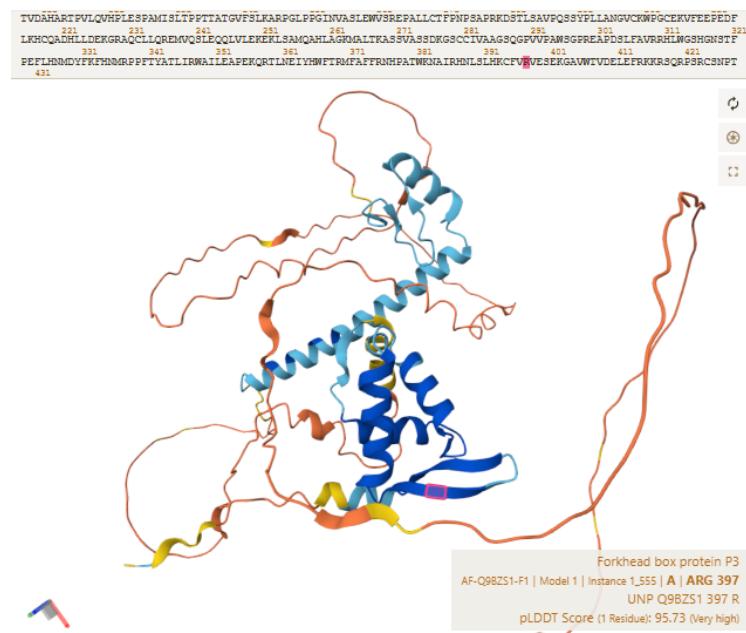
1 MPNPRPGKPSAPSIALGPSPGASPWSRAAPKASDLLGARGPGGTQGRDLRGGAHASSSS 60
61 LNPMPPSQLQLTPLPLVMVAPSGARLGPLPHIQLLQDRPHFMHQIQLSTVDAAHARTPVLOV 120
121 HPLESPAMISLTPPTATGVFSLKARPGLPPGINVASLEWVSREPALLCTFPNPSAPRKD 180
181 STLSAVQSSYPLLANGCKWPGCERVFEEDFLKHCQADHLLDEKGRACQCLLQREMVQ 240
241 SLEQQLVLEKEKSMQAHLAGMALTKASSVASSDKGSCCIVAGSQGVVPFWNSGPRE 300
301 APDSLFAVRHLWGSHGNSTFPEFLRNMDYFKFHNMRPFYTATLIRWAIALEPKQRTL 360
361 NEIYHWHTRMFAFRNHPATWKNAIRHNLSLHKCFVWESEKGAVWTDELEFRKKRSQR 420
421 PSRCSNPTGPG 431

```

TRANSCRIPT SEQUENCE



PROTEIN STRUCTURE



VARIANT c.1112T>G (p. Phe371Cys)

PROTEIN SEQUENCE

```
1 MPNPRPGKPSAPSALALGPSPGASPSWRAAPKASDLLGARGPGGFTQGRDLRGGAHASSSS 60
61 LNMPMPSQLQLPTLPLVMAPSGARGLPFLHQALLQDRPHMFNLSTVDHAHPTVPLQV 120
121 HPLEPSAMISLTPTTATGVFLSKARPGFLPPGINVASVWHSREPAALLCFTPNPSAPRKD 180
181 STLSAQVSPQSYYLLANGCWVKPGCEKVFFEEPDFLKHCOADHLDEKGRAQCQLQREMO 240
241 SLEQQQLVLKEKELSAMQAHQLAGKMLTAKSASSVASSDKGSGCCIVAAQSGQEVPAWSGPRE 300
301 APDSLAVFVRHLWGHSGNSTFPEFLHNMDFYFKFHNNMRPFVTIATLIRWAILEAPEKQRL 360
361 NEIYHWFTRMCAFFRNHPATWKNAIRHNLSLHKCFVRVESEKGAVWTDELEFRKRSQR 420
421 PSRCSNTPGP 431
```

TRANSCRIPT SEQUENCE

Codons	Alternating codons	Alternating codons					
Exons	An exon	Another exon					
Variants	3 prime UTR	5 prime UTR	Coding sequence	Frameshift	Inframe deletion	Missense	Splice region
	Start lost	Stop gained	Stop lost	Stop retained	Synonymous		
Other	UTR						
Markup	loaded						

- Variants are filtered by consequence type

M	YR	M	M	V	*	R	B	Y	Y	R	Y	R	R	W	M	W							
1021	GGA	A	CGC	GCAT	TG	CGA	TCT	C	AC	AT	AT	AT	G	AT	CCAC	AC	ATG	1080					
949	GGAA	ACAG	CACAT	CCC	GAG	T	TCT	C	CC	AC	ATG	GA	T	CT	CCAC	AC	ATG	1008					
317	-G	-N	-S	-T	-F	-P	-E	-F	-L	-H	-N	-M	-D	-Y	-F	-K	-F	-H	-N	-M	336		
	H	R	S	Y		YR	HR*	R*	BR	K*	S	*Y			*V**								
1081	CG	CCC	G	T	T	C	CC	CA	CG	T	C	G	C	T	G	G	GTC	C	AG	GA	AG	1140	
1009	CG	AA	CG	AC	AT	CC	CC	G	AG	CT	CC	G	CC	AT	CC	TG	GG	CC	AG	AG	AG	1068	
337	-R	-P	-F	-T	-Y	-A	-T	-L	-I	-R	-W	-A	-I	-L	-E	-A	-P	-E	-K	-	356		
	YB	Y	*YS	R***	R	*	*	*Y*S	YR	R	***K	*H***Y***	R										
1141	CAG	CG	GA	C	AT	CG	TG	T	TC	AC	CC	T	AT	T	CG	CT	TC	TA	AA	AC	AC	1200	
1069	CA	GC	GG	AC	CT	AA	GT	CC	AC	TC	AT	TC	CG	CA	CT	TC	TC	TC	CA	GA	AC	1128	
357	-Q	-R	-T	-L	-N	-E	-I	-Y	-H	-W	-F	-T	-R	-M	-F	-A	-F	-F	-R	-N	-	376	
	SR	RM	*	R	*	YRY	**Y*		*RY	*	R	Y											
1201	CA	T	CT	CC	CA	CT	G	AA	GA	AT	CC	CA	AT	TC	AC	CA	AG	TG	TT	GT	TG	1260	
1129	CA	T	CT	CC	GC	CA	CT	G	AA	GA	AC	CT	CC	CA	AC	TC	AC	AG	TG	TT	GT	1188	
377	-H	-P	-A	-T	-W	-K	-N	-A	-I	-R	-H	-N	-L	-S	-L	-H	-K	-C	-F	-V	-	396	
	YD	RHR	***	Y	*	*	BR	Y	****	*YRM	*	WW											
1261	CG	GT	GG	GA	GA	AA	GG	GG	G	T	G	TG	AC	CG	TG	GA	T	GAG	GT	CCCC	AC	GG	1320
1189	CG	GG	GG	GA	GA	AA	GG	GG	G	T	G	TG	AC	CG	TG	GA	T	GAG	GT	TC	CA	AG	1248
397	-R	-V	-E	-S	-E	-K	-G	-A	-V	-W	-T	-V	-G	-D	-A	-G	-T	-G	-A	-G	-A	416	

PROTEIN STRUCTURE



VARIANT c.1150G>A (p. Ala384Thr)

PROTEIN SEQUENCE

```

1 MPNPRPGKPSAPSILALGPSPGASPSSWAAPKASDLLGARGPGGTFQGRDLRGGAHASSS 60
61 LNMPMSQLPLPVLVVAFSGARLGPLPHQLALLQDRPHFMHQLSTDAHARTPVLOV 120
121 HPLESPAMISLTPPTATGVFSLKARPGLPPGINVASLEWVSREPALLCTFPNPSAPRKD 180
181 STLSAVFQSSYLLANQVKWPGECKVFEEPEDFLKHCQADHLLDEKGRAOCLLQREMVQ 240
241 SLEQOLVLEKEKLSAMQAHLAGMALTKASSVASSDKGSCCIVAAGSQGPVVPAWSGPRE 300
301 APDSLFAVRHRHLWGSHNSTFPEFLHNMDYFKFHNMRFPFTYATLIRNAILEAPEKQRIL 360
361 NEIYHWFTRMFAFFRHNHATWKNTIRHNLSLHKCFVRVESEKGAVWTVDELEFRKKRSQR 420
421 PSRCSNFTPGP 431

```

TRANSCRIPT SEQUENCE

Codons Alternating codons Alternating codons

Exons An exon Another exon

Variants 3 prime UTR 5 prime UTR Coding sequence Frameshift Inframe deletion Missense Splice region

Start lost Stop gained Stop lost Stop retained Synonymous

Other UTR

Markup loaded

- Variants are filtered by consequence type

	Sequence	Position
1081	GG CCCG A TTCACTTACGCCAGTC GG CGAT GG CTGGAGGT GG GAAG	1140
1009	CGACCCCCCTTCACTTACGCCAGTC GG CGAT GG CTGGAGGT GG GAAG	1068
337	-R--P--P--F--I--Y--A--T--L--I--R--W--A--I--L--E--A--P--E--K-	356
1141	YR Y *YS R** *R * * *YS YR R **K *H***Y*** R	1200
1069	CAGCGGAC A CTGAT GG CTTG GG ACACCC GG GTG GG TC GG AAAC	1128
357	-Q--R--T--L--I--Y--H--W--F--T--R--M--F--A--F--R--N-	376
1201	SR RM * R * VYV *Y* *RY * R Y	1260
1129	CATCCTGCCACTTGGAAAGAACCCATCCGGCCACACCTGAGTC GG ACAA GG CTT GG AAAC	1188
377	-H--P--A--I--R--N--L--S--L--H--K--C--F--V-	396
1261	YDD RHR *** Y * * BR *Y * * * YRM * W W	1320
1189	CGGTGGAGAGCAGAAGGC GG CTTG GG AC GG CTGG GG AG GG GT GG CC GG AA GG AA	1248
397	-R--V--E--S--E--K--G--A--V--W--T--V--D--E--L--E--F--R--K--K-	416
1321	YRR VS Y R *** RMSS Y Y* * * * S+R*****	1380
1249	CGGAGCCAGGGCCCCAC GG CTTG GG CC GG AC GG CT GG CA GG AA.....	1296
417	-R--S--Q--R--P--S--R--C--S--N--P--T--P--G--P--*	431

PROTEIN STRUCTURE

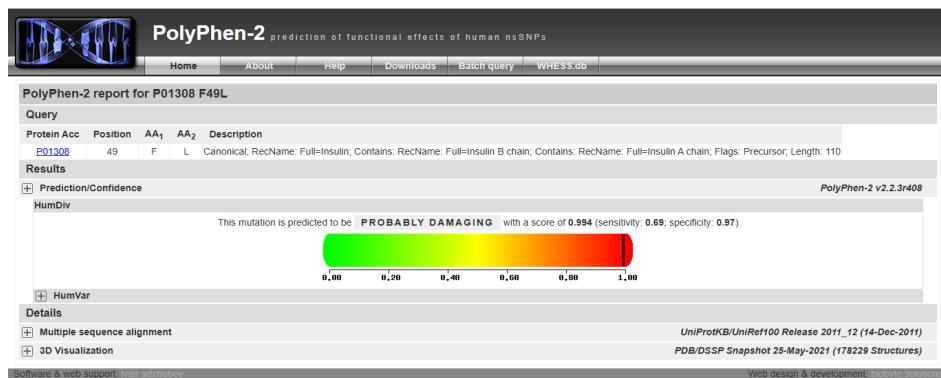


RESULTS

INS

VARIANT c.147C>G (p. Phe49Leu)

POLY PHEN 2



SIFT- Sorting Intolerant From Tolerant

Predicted the variant under tolerated

w h y f m i q r n d e l c k v t g	33S	1.00	P A S
m w i f v c l y p r q a t e k s g d	34H	1.00	N H
y w v t s r q p n m k i h g f e d c a	35L	1.00	L
y w t s r q p n m l k i h g f e d c a	36V	1.00	V
c w m f i y l v h r g t n s p a k q	37E	1.00	D E
w h y f m q r n d i c e k l p v g s	38A	1.00	T A
y w v t s r q p n m k i h g f e d c a	39L	1.00	L
w v t s r q p n m l k i h g f e d c a	40Y	1.00	Y
d h g n c e w k r p y q t a v i	41L	1.00	S F M L
w h y d g n r q f e k c m p s l a i	42V	1.00	T V
y w v t s r q p n m l k i h g f e d a	43C	1.00	C
w m i f y c h l v r t e k p d n s a	44G	1.00	Q G
c w f m i y l v h g t n s a k q	45E	1.00	R P D E
w c f m y i h v l p g a q e n	46R	1.00	D S T K R
y w v t s r q p n m l k i h f e d c a	47G	1.00	G
y w v t s r q p n m l k i h g e d c a	48E	0.97	F
y w v t s r q p n m l k i h g e d c a	49F	1.00	F
w v t s r q p n m l k i h g f e d c a	50Y	1.00	Y
w	51T	1.00	c m p g h d f v l I e k Q R A Y N S T
y w v t s r q n m l k i h g f e d c a	52P	1.00	P
c w f d y i v g p s n a t l e q	53K	0.97	H r M K
	54T	0.81	w c m F h y p I g l V n r Q D T k e S A
c w d f m y v g p s h n a l t e q	55R	0.75	I K R
y w v t s q p n m l k i h g f e d c a	56R	0.91	R
c w m f i y l v r h t p s a k n q	57E	1.00	G E D
g w h d y n r q e s k c p f t	58A	1.00	M I A L V
c w m f i y l v r h t p s n a k q	59E	1.00	G D E
w m c f i v	60D	0.97	l y r H t a k s Q P G E N D
w	61L	0.88	h y n c d Q r f e P k m t g s i V A L
w c y h f m	62Q	0.72	p d n i v e g r t k a s L Q
h w d p q c e n r k s y t	63V	0.78	f M A G L I V
w	64G	0.75	c y m F h i v P l n t R q G d S A K E
w h y d g r n	65Q	0.75	w c m h i F p y v g L n n r t D S Q A K E
w h y d g r n	66V	0.72	f q e k c m p S i T A L V

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.



mutation t@sting

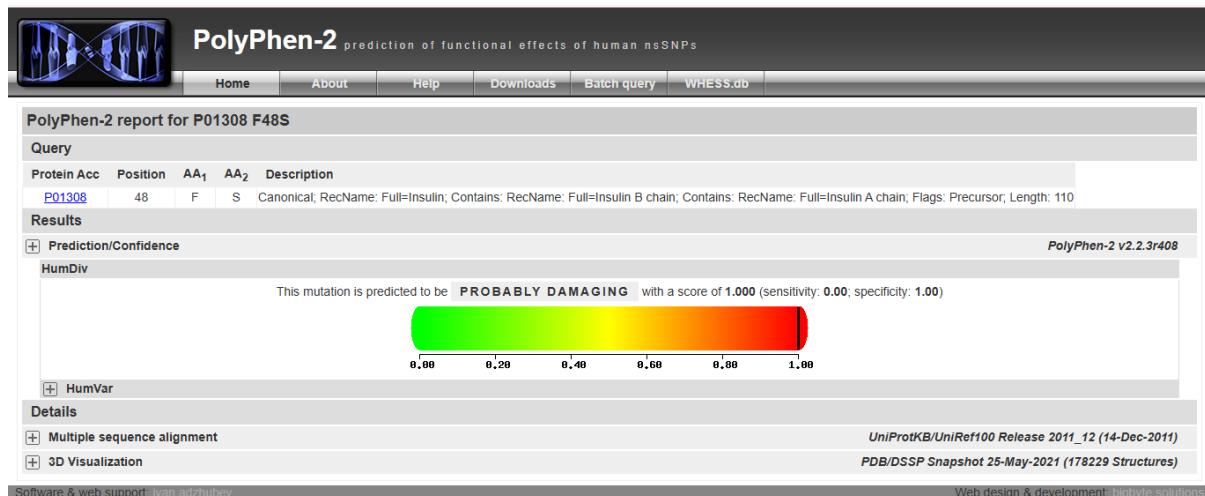
Prediction:	Deleterious	Permalink
Summary:	<ul style="list-style-type: none"> Amino acid sequence changed Known disease mutation at this position (HGMD CM810002) Known disease mutation: ClinVar ID 13377 (pathogenic) Protein features (might be) affected 	<ul style="list-style-type: none"> Model: simple_aae Tree vote: 100 0 (del benign) ? Automatic classification due to ClinVar
Analysed issue	Analysis result	
Phys. location	chr11:2182055G>C show variant in all transcripts IGV	
Gene symbol	INS	
EXAC LOF metrics	LOF: 0.73, missense: 0.57, synonymous: -0.51	
Ensembl transcript ID	ENST00000381330.4	
Genbank transcript ID	NM_00207 (exact from MANE)	
UniProt peptide	P01308	
Variant type	Single base exchange	Activate W Go to Setting
Gene region	CDS	
DNA changes	c.147C>G g.517C>G	

Protein conservation	Species	Match	Gene	AA Alignment
	Human	not conserved		49 ALYLVCGERGFYYTPKTRREADL
	mutated			49 ALYLVCGERFLYTPKTRREAD
	Prooglyotes	no homologue		
	Mmulatta	all identical	ENSMUUG00000031267	49 ALYLVCGERGFYYTPKTRREAD
	Fcatus	no homologue		
	Mmusculus	all identical	ENSMUSG00000035804	49 ALYLVCGERGFYYTPKSRREVED
	Ggalus	all identical	ENSGALG00000006552	49 ALYLVCGERGFYYSPKARRDVEQ
	Trubripes	all identical	ENSTRUG000000030504	50 ALYLVCGDGFYYNPK---RDVD
	Dreno	all identical	ENSDARG00000035350	49 ALYLVCGDGFYYNPK---RDVEP
	Dmelanogaster	no homologue		
	Celegans			
	Xtropicalis	all identical	ENSXETG00000014029	49 ALYLVCGDGFYYYPKIKRDIQ

Original gDNA sequence snippet	TGCGGGAAACGAGGCCCTTCTTACACACCCAAGACCCGCCG
Altered gDNA sequence snippet	TGCGGGAAACGAGGCCCTTCTTACACACCCAAGACCCGCCG
Original cDNA sequence snippet	TGCGGGAAACGAGGCCCTTCTTACACACCCAAGACCCGCCG
Altered cDNA sequence snippet	TGCGGGAAACGAGGCCCTTCTTACACACCCAAGACCCGCCG
Wildtype AA sequence	HALWIRLLPL LALLLWGPD PAAAFVNQHL CGSHLVEALY LVCGERGFY TPKTRREADL LQVGQVELGG GPAGSLQLP ALEGSILQRKG IVEQCTCTIC SLYQLENYN *
Mutated AA sequence	HALWIRLLPL LALLLWGPD PAAAFVNQHL CGSHLVEALY LVCGERGFY TPKTRREADL LQVGQVELGG GPAGSLQLP ALEGSILQRKG IVEQCTCTIC SLYQLENYN *

VARIANT c.143T>C (p. Phe48Ser)

POLY PHEN 2 – Predicted the variant to be a probably damaging kind with a score of 1.000.



SIFT- Sorting Intolerant From Tolerant

Predicted the variant under tolerated

Sequence alignment showing the variant at position 48 (F to S) circled in red:

Residue Position	Residue	Score	Conservation
33S	S	1.00	PAS
34H	H	1.00	NH
35L	L	1.00	L
36V	V	1.00	V
37E	E	1.00	DE
38A	A	1.00	TA
39L	L	1.00	L
40Y	Y	1.00	Y
41L	L	1.00	SFML
42V	V	1.00	TV
43C	C	1.00	C
44G	G	1.00	QG
45E	E	1.00	RPDE
46R	R	1.00	DSTKR
47G	G	1.00	G
48F	F	0.97	E
49F	F	1.00	F
50Y	Y	1.00	Y
51T	T	1.00	cmpghdfvlIEkQRAYNST
52P	P	1.00	P
53K	K	0.97	HrMK
54T	T	0.81	wcmFhypIglvnrrQDTkesA
55R	R	0.75	IKR
56R	R	0.91	R
57E	E	1.00	GED
58A	A	1.00	MIALV
59E	E	1.00	GDE
60D	D	0.97	lyRHaksQPGEND
61L	L	0.88	hyncdQrfEPkmgsival
62Q	Q	0.72	pdnivegrtkasLQ
63V	V	0.78	fMAGLIV
64G	G	0.75	cymFhivP1ntRqGdsAKE

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.

Mutation Taster

mutation t@sting

Prediction: **Deleterious** [Permalink](#)

Summary:

- Amino acid sequence changed
- Known disease mutation at this position (HGMD CM074289)
- Known disease mutation at this position (HGMD CM830002)
- Known disease mutation: ClinVar ID 13378 (pathogenic)
- Protein features (might be) affected

Analysed issue	Analysis result
Phys. location	chr11:2182059A>G show variant in all transcripts IGV
Gene symbol	INS
ExAC LOF metrics	LOF: 0.73, missense: 0.57, synonymous: -0.51
Ensembl transcript ID	ENST00000381330.4
Genbank transcript ID	NM_000207 (exact from MANE)
UniProt peptide	P01308
Variant type	Single base exchange
Gene region	CDS
DNA changes	c.143T>C g.513T>C

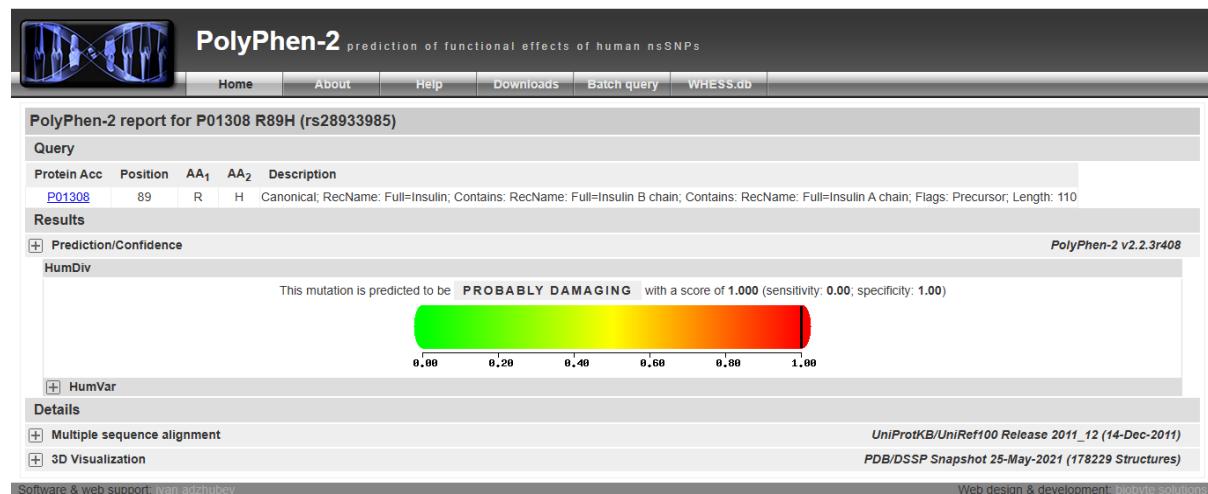
Activate W [Go to Settings](#)

Protein conservation	Species	Match	Gene	AA Alignment
Human	Human	not conserved		EALYLVCGERGFYTPKTRREAED
mutated				EALYLVCGERGSFYTPKTRREA
Ptroglydotes	Ptroglydotes	no homologue	ENSMMMUG00000031267	EALYLVCGERG[FYTPKTRREA
Mmulatta	Mmulatta	all identical		
Fcatus	Fcatus	no homologue		
Mmusculus	Mmusculus	all identical	ENSMUSG00000035804	EALYLVCGERG[FYTPKSREVE
Ggalus	Ggalus	all identical	ENSGALG00000006552	EALYLVCGERG[FYSPKARRDOVE
Trubripes	Trubripes	all identical	ENSTRUG00000030504	DALYLVCGDRG[FYNPK--RDV
Drerio	Drerio	all identical	ENSDARG00000035350	DALYLVCGPTG[FYNPK--RDVE
Dmelanogaster	Dmelanogaster	no homologue		
Celegans	Celegans	no homologue		
Xtropicalis	Xtropicalis	all identical	ENSXETG00000014029	EALYLVCGDRG[FYPPIKRDI

Original gDNA sequence snippet	AGTGTGCCGGAAACGAGGCT T CTTCTACACACCCAAGACCC
Altered gDNA sequence snippet	AGTGTGCCGGAAACGAGGCT C TTCTACACACCCAAGACCC
Original cDNA sequence snippet	AGTGTGCCGGAAACGAGGCT T CTTCTACACACCCAAGACCC
Altered cDNA sequence snippet	AGTGTGCCGGAAACGAGGCT C TTCTACACACCCAAGACCC
Wildtype AA sequence	MALWIRLLPL LALLALWQD PAAAFVNQHL CGSHLVEALY LVCGERGFY TPKTRREAED LQVGQVELGG GPGAGSLQLP ALEGSLQRK6 IVEQQCTSIC SYLQLENYN *
Mutated AA sequence	MALWIRLLPL LALLALWQD PAAAFVNQHL CGSHLVEALY LVCGERGSFY TPKTRREAED LQVGQVELGG GPGAGSLQLP ALEGSLQRK6 IVEQQCTSIC SYLQLENYN *

VARIANT c.266G>A (p. Arg89His)

POLY PHEN 2 – Predicted the variant to be a probably damaging kind with a score of 1.000.



SIFT- Sorting Intolerant From Tolerated

Predicted the variant under tolerated

wc	y	78Q	1.00	wcmiphFvGNlRDYtsQkAE
wy	fcmh	79P	1.00	mhiFvlnrGsQDkTAEP
wcfymihvlgpns	wyf	80L	0.97	cwpdMqnEKrGitsVAhFLY
wcfymihvlgpns	wyf	81A	0.97	ilVPnrgDTsQKEA
wcfymihvlgpns	wyf	82L	1.00	cwpmeknrdtiQsGvhAYFL
wcfymihvlgpns	wyf	83E	1.00	TDAKQE
wcfymihvlgpns	wyf	84G	1.00	cwpMDqENKGritsVAHlFY
wcfymihvlgpns	wyf	85S	1.00	chMI1VPgnrTDQSAKE
wcfymihvlgpns	wyf	86L	0.97	cwdPMgnisetvhfaQYLRK
wcfymihvlgpns	wyf	87Q	0.97	ihdpqlinstaMVERKQ
wcfymihvlgpns	wyf	88K	0.91	K
wcfymihvlgpns	wyf	89R	0.91	R
wyvtsrqpnmlkihgfedca	wyvtsrqpnmlkihgfedca	90G	0.97	G
wyvtsrqpnmlkihgfedca	wyvtsrqpnmlkihgfedca	91I	1.00	VI
wyvtsrqpnmlkihgfedca	wyvtsrqpnmlkihgfedca	92V	1.00	V
wyvtsrqpnmlkihgfedca	wyvtsrqpnmlkihgfedca	93E	1.00	QDE
wyvtsrqpnmlkihgfedca	wyvtsrqpnmlkihgfedca	94Q	1.00	EQ
wyvtsrqpnmlkihgfedca	wyvtsrqpnmlkihgfedca	95C	1.00	C
wyvtsrqpnmlkihgfedca	wyvtsrqpnmlkihgfedca	96C	1.00	C
wyvtsrqpnmlkihgfedca	wyvtsrqpnmlkihgfedca	97T	1.00	cfiylrvqpHEkgAdNST
wyvtsrqpnmlkihgfedca	wyvtsrqpnmlkihgfedca	98S	1.00	ivylptHaqeRdGKSN
wyvtsrqpnmlkihgfedca	wyvtsrqpnmlkihgfedca	99I	1.00	feKcmsPlTaIV
wyvtsrqpnmlkihgfedca	wyvtsrqpnmlkihgfedca	100C	1.00	C

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.

Mutation Taster

mutation t@sting

Prediction: **Deleterious** [Permalink](#)

Summary:

- Amino acid sequence changed
- Known disease mutation at this position (HGMD CM850007)
- Known disease mutation at this position (HGMD CM920374)
- Known disease mutation at this position (HGMD CM970784)
- Known disease mutation: ClinVar ID 13380 (pathogenic)
- Model: simple_aae
- Tree vote: 97|3 (del | benign)
- Automatic classification due to ClinVar

Analysed issue	Analysis result
Phys. location	chr11:2181149C>T show variant in all transcripts IGV
Gene symbol	INS
ExAC LOF metrics	LOF: 0.73, missense: 0.57, synonymous: -0.51
Ensembl transcript ID	ENST00000381330.4
Genbank transcript ID	NM_000207 (exact from MANE)
UniProt peptide	P01308
Variant type	Single base exchange
Gene region	CDS
DNA changes	c.266G>A g.1423G>A

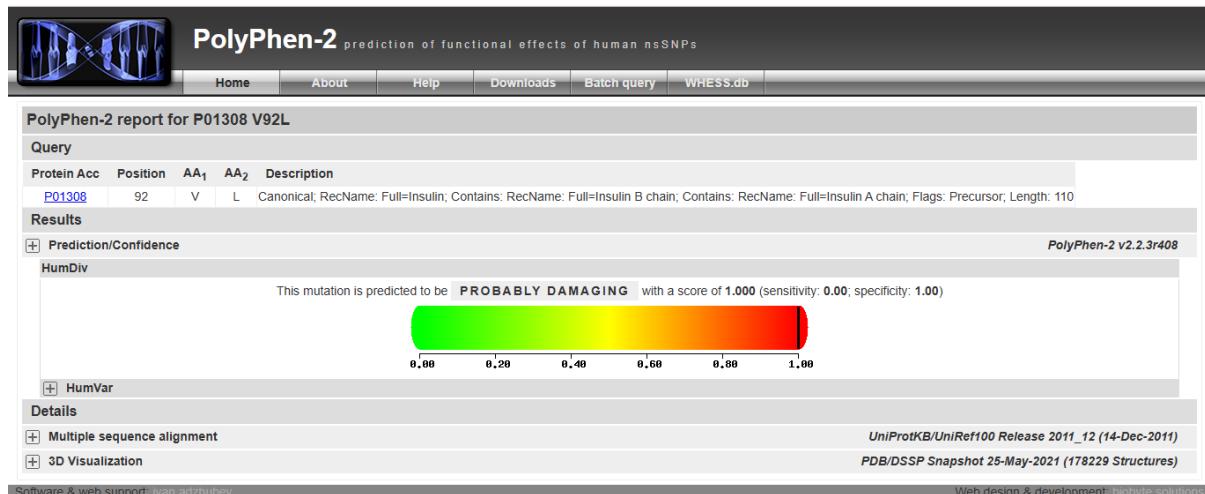
Activate V
Go to Setting

Protein conservation	Species	Match	Gene	AA Alignment
Human	not conserved			89 QPLAL EGS LQ KRGIVEQCCTSICS
mutated				89 QPLAL EGS LQ KRGIVEQCCTSIC
Ptroglobutes	no homologue			
Mmulatta	all identical	ENSMUML00000031267	89 QPLAL	EGS LQ KRGIVEQCCTSIC
Fcatus	no homologue			
Mmusculus	all identical	ENSMUSG00000035804	89 QTAL	EVA RQ KRGIVDQCCTSIC
Ggallus	all identical	ENSGAL00000006552	89 PFQEE	EYE KV KRGIVEQCCHNTC
Tribripes	all identical	ENSTRUG00000030504	87 AEYAF	KDQ MEIMVNGIVEQCCLRPC
Driero	all identical	ENSDAARG00000035350	86 ETEVA	DFAFKDHAELIR KRGIVEQCCHKPC
Dmelanogaster	no homologue			
Celegans	no homologue			
Xtropicalis	all identical	ENSXETG00000014029	88 DGIQLQPQEYQ	KH KRGIVEQCCHSTC

Original gDNA sequence snippet	GGAGGGGCCCTGCAGAACG GT TGGCATTGTGGAAACAATGCT
Altered gDNA sequence snippet	GGAGGGGCCCTGCAGAAC CAT TGGCATTGTGGAAACAATGCT
Original cDNA sequence snippet	GGAGGGGCCCTGCAGAAC GT GGCATTGTGGAAACAATGCT
Altered cDNA sequence snippet	GGAGGGGCCCTGCAGAAC AT GGCATTGTGGAAACAATGCT
Wildtype AA sequence	MALINWRLPL LALLALGPD PAAAFVNQHL CGSHLVEALY LVGERGFY TPKRREAED LQVQVQVELGG GPGAGSLQPL ALEGSLQKRG IVEQCCTSIC SLYQLENYN *
Mutated AA sequence	MALINWRLPL LALLALGPD PAAAFVNQHL CGSHLVEALY LVGERGFY TPKRREAED LQVQVQVELGG GPGAGSLQPL ALEGSLQ KHG IVEQCCTSIC SLYQLENYN *

VARIANT c.274G>T (p. Val92Leu)

POLY PHEN 2 – Predicted the variant to be a probably damaging kind with a score of 1.000.



SIFT- Sorting Intolerant From Tolerated

Predicted the variant under tolerated

wcfymihvlgrpns	83E	1.00	TDAKQE
	84G	1.00	cwpMDqENKGritsVAHlFY
wyf	85S	1.00	chMILVPgnrTDQSAKE
	86L	0.97	cwdPMgnIsetvhfaQYLRK
cwfy	87Q	0.97	ihdpglnstaMVERKQ
wfcmyihdvlnpteqsgnA	88K	0.91	K
ywvtsqpnlkihgfedca	89R	0.91	R
ywvtsrqpnmlkihgfedca	90G	0.97	G
hwqdnpercgksytfmal	91I	1.00	VI
vwtsrqpnmlkihgfedca	92V	1.00	V
cwfmiylvhrgntspak	93E	1.00	QDE
cwmfiyvhltgnsprakd	94Q	1.00	EQ
ywvtsrqpnmlkihgfedca	95C	1.00	C
whmyfqqrkednilpvgsaT	96C	1.00	C
	97T	1.00	cfiylrqvppHEkgAdNST
wmcf	98S	1.00	ivylptHaqeRdGKSN
gwhydnrq	99I	1.00	fekcmsPlTaIV
ywvtsrqpnmlkihgfedca	100C	1.00	C

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.



mutation t@sting

Prediction: Deleterious [Permalink](#)

Summary:

- Amino acid sequence changed
- Known disease mutation at this position (HGMD CM860013)
- Known disease mutation: ClinVar ID 13381 (pathogenic)
- Protein features (might be) affected
- Model: simple_aae
- Tree vote: 100|0 (del | benign) ?
- Automatic classification due to ClinVar

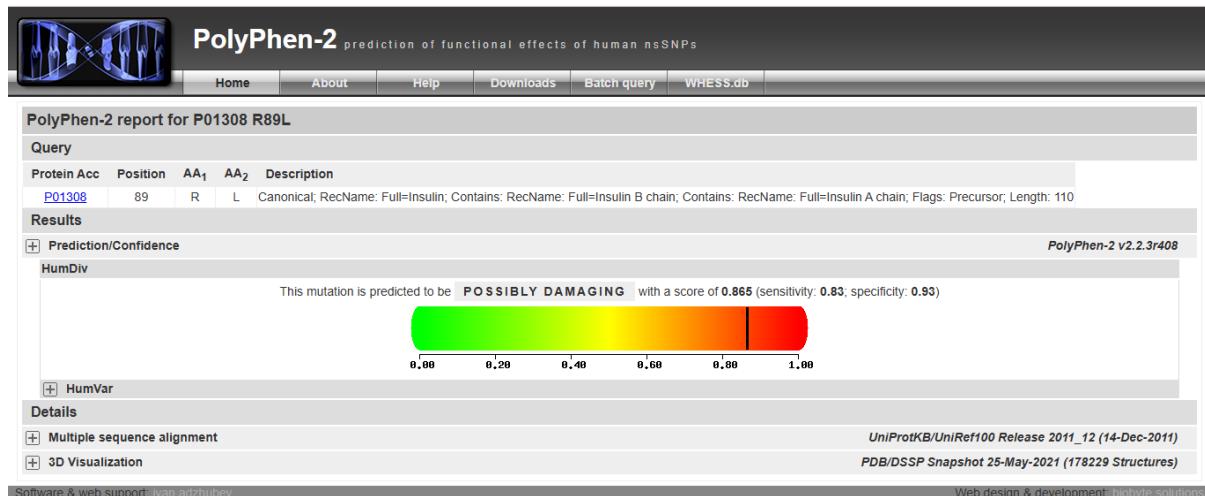
Analysed Issue	Analysis result
Phys. location	chr11:2181141C>A show variant in all transcripts IGV
Gene symbol	INS
ExAC LOF metrics	LOF: 0.73, missense: 0.57, synonymous: -0.51
Ensembl transcript ID	ENST00000381330.4
Genbank transcript ID	NM_000207 (exact from MANE)
UniProt peptide	P01308
Variant type	Single base exchange
Gene region	CDS Activate WiFi
DNA changes	c.274G>T g.1431G>T Go to Settings

Protein conservation	Species	Match	Gene	AA Alignment	LQ	KRGIVEQQCTTSICSLY
Human				92 AL EGS	LQ	KRGIVEQQCTTSICSLY
mutated	all conserved			92 AL EGS	LQ	KRGIEQQCTTSICSLY
Prologodytes	no homologue					
Mmulatta	all identical		ENSMUMUG00000031267	92 AL EGS	LQ	KRGIEQQCTTSICSLY
Fcatus						
Mmusculus	all identical		ENSMUSG00000035804	92 AL EVA	RQ	KRGIVQQCTTSICSLY
Ggalus	all identical		ENSGALG00000006552	92 QE EYE	KV	KRGIVQQCHNTCSLY
Tribripes	all identical		ENSTRUG00000030504	90 AF KDQ	HEMWVKRGIVQQCLRPNCLL	
Drerio	all identical					
Dmelanogaster	no homologue		ENSDARG00000035350	89 VA DFAFKDHAILIR	KRGIVQQCHHKPCSF	
Celegans	no homologue					
Xtrropicalis	all identical		ENSXFTG00000014029	91 QLOPOFYQ	KM	KRGIVQQCHSTCSLF

Original gDNA sequence snippet	CCCTGCAAGCGTGGCATTTGTGGAAACAATGCTGTACCAAGC
Altered gDNA sequence snippet	CCCTGCAAGCGTGGCATTTGTGGAAACAATGCTGTACCAAGC
Original cDNA sequence snippet	CCCTGCAAGCGTGGCATTTGTGGAAACAATGCTGTACCAAGC
Altered cDNA sequence snippet	CCCTGCAAGCGTGGCATTTGTGGAAACAATGCTGTACCAAGC
Wildtype AA sequence	MAWIRLLPL LALLALWGPD PAAAFVNQHL CGSHLVEALY LVCGERGFY TPKTRREAED LQVGQVELGG GPGAGSLQPL ALEGSILQKRG IVEQCTCTIC SLYQLENYCN *
Mutated AA sequence	MAWIRLLPL LALLALWGPD PAAAFVNQHL CGSHLVEALY LVCGERGFY TPKTRREAED LQVGQVELGG GPGAGSLQPL ALEGSILQKRG ILEQCTCTIC SLYQLENYCN *

VARIANT c.266G>T (p. Arg89Leu)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 0.865.



SIFT- Sorting Intolerant From Tolerated

Predicted the variant under tolerated

wc	y	78Q	1.00	wcmiPhFvGN1RDYtsQkAe
wy	f	79P	1.00	mhiFvlnnGsQDkTAEP
wy	f	80L	0.97	cwpdMqnEKrGitsVAhFLY
wy	f	81A	0.97	ilVPnrGDTsQKEA
wy	f	82L	1.00	cwpmeknRDtIQsGvhAYFL
wy	f	83E	1.00	TDAKQE
wy	f	84G	1.00	cwpMDqENKGritsVAHlFY
wy	f	85S	1.00	chMILVPgnrTDQSAKE
wy	f	86L	0.97	cwdPMgnIsetvhfaQYLRK
wy	f	87Q	0.97	ihdpqlnstamVERKQ
wf	cm	88K	0.91	K
wy	vtsqpnmlkihgfedca	89R	0.91	R
wy	vtsrqpnmlkihgfedca	90G	0.97	G
hw	qd	91I	1.00	VI
yw	tsrqpnmlkihgfedca	92V	1.00	V
cwf	miylvhrgntspak	93E	1.00	QDE
cwmfiyvhltgnsprakd	94Q	1.00	EQ	
ywv	tsrqpnmlkihgfedca	95C	1.00	C
whmy	fqrkednilpvgsaT	96C	1.00	C
wm	97T	1.00	cfiylrqvppHEkgAdNST	
wmc	f	98S	1.00	ivylptHaqeRdgKSN
gwhydnrq	99I	1.00	fekcmsPlTaIV	
ywv	tsrqpnmlkihgfedca	100C	1.00	C

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.



mutation t@sting

Prediction:
Deleterious
[Permalink](#)

Summary:

- Amino acid sequence changed
- Known disease mutation at this position (HGMD CM860007)
- Known disease mutation at this position (HGMD CM920374)
- Known disease mutation at this position (HGMD CM970784)
- Known disease mutation: ClinVar ID 13382 (pathogenic)
- Model: simple_aae
- Tree vote: 100|0 (del | benign) 
- Automatic classification due to ClinVar

Analysed issue	Analysis result
Phys. location	chr11:2181149C>A show variant in all transcripts IGV
Gene symbol	INS
ExAC LOF metrics	LOF: 0.73, missense: 0.57, synonymous: -0.51
Ensembl transcript ID	ENST00000381330.4
Genbank transcript ID	NM_000207 (exact from MANE)
UniProt peptide	P01308
Variant type	Single base exchange
Gene region	CDS
DNA changes	c.266G>T g.1423G>T

[Activate V](#) [Go to Setting](#)

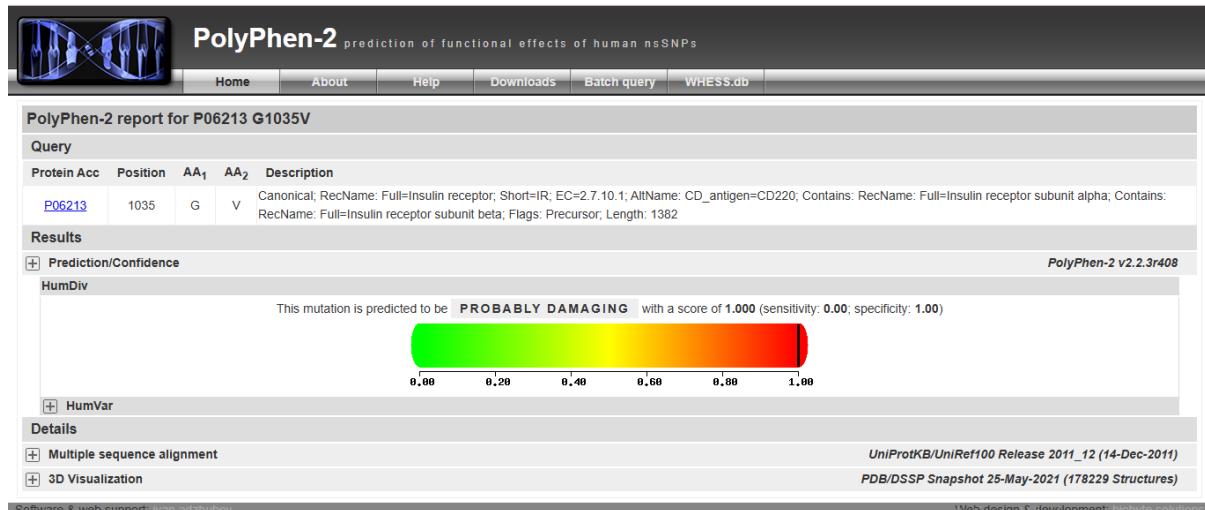
Protein conservation	Species	Match	Gene	AA Alignment	AA	Conservation
	Human	not conserved		89 QPLAL EGS	LQ	KRGIVEQCCTTSIC
	mutated	not conserved		89 QPLAL EGS	LQ	KLGIVEQCCTTSIC
	Prooglyotes	no homologue				
	Mmulatta	all identical	ENSMUJG00000031267	89 QPLAL EGS	LQ	KRGIVEQCCTTSIC
	Fcatus	no homologue				
	Mmusculus	all identical	ENSMUSG00000035804	89 QTLAL EVA	RQ	KRGIVDQCCTTSIC
	Ggalus	all identical	ENSGALG00000006552	89 PFQQE EYE	KV	KRGIVFQCCHNTC
	Trubripes	all identical	ENSTRUG00000030504	87 AEYAF KDQ	MMPVW	KRGIVFQCCLRPC
	Drerio	all identical	ENSDARG00000035350	86 ETEVA DFAFKDHAELIR	KM	KRGIVFQCCHKPC
	Dmelanogaster	no homologue				
	Celegans	no homologue				
	Xtropicalis	all identical	ENSXETG00000014029	88 DGHQLOPQEYQ	KM	KRGIVEQCCHSTC

Original gDNA sequence snippet	GGAGGGGTCCTGCAGAAC G CTGGCATTGTGGAAACATGCT
Altered gDNA sequence snippet	GGAGGGGTCCTGCAGAAC G CTTGGCATTGTGGAAACAATGCT
Original cDNA sequence snippet	GGAGGGGTCCTGCAGAAC G CTGGCATTGTGGAAACAATGCT
Altered cDNA sequence snippet	GGAGGGGTCCTGCAGAAC G CTTGGCATTGTGGAAACAATGCT
Wildtype AA sequence	MAIWRRLPL LALLLWGPD PAAAFVNQHL CGSHLVEALY LVCGERGFY TPKTRREAED LQVGQVELGG GPAGASLQPL ALEGSLQKG IVEQCCTSIC SLYQLENYN *
Mutated AA sequence	MAIWRRLPL LALLLWGPD PAAAFVNQHL CGSHLVEALY LVCGERGFY TPKTRREAED LQVGQVELGG GPAGASLQPL ALEGSLQ KL G IVEQCCTSIC SLYQLENYN *

INSR

VARIANT c.3104G>T (p. Gly1035Val)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 1.000.



SIFT- Sorting Intolerant From Tolerant

Predicted the variant under tolerated

wfymhcgblr	1024T	1.00	qpvdekekNIAST
dghnceanwsrkyp	1025L	1.00	afqvMIL
	1026L	1.00	wcpdhgqrenkmysfatvIL
cwfmiyvhplgstdnaq	1027R	1.00	KER
ywvtsrqpnmlkihgfdca	1028E	1.00	E
ywvtsrqpnmkihgfedca	1029L	1.00	L
ywvtsrqpnmlkihfedca	1030G	1.00	G
ywvtsrqpnmlkihfedca	1031Q	1.00	Q
ywvtsrqpnmlkihfedca	1032G	1.00	G
ywvtsrqpnmlkihfedca	1033S	1.00	S
ywvtsrqpnmlkihgfedca	1034F	1.00	F
ywvtsrqpnmlkihfedca	1035G	1.00	G
ywvtsrqpnmlkihgfedca	1036M	1.00	M
ywvtsrqpnmlkihgfedca	1037V	1.00	V
wvtsrqpnmlkihgfdca	1038Y	1.00	Y
cwfmyivhlgtnpsraq	1039E	1.00	dKE
ywvtsrqpnmlkihfedca	1040G	1.00	G
wch	1041N	1.00	pgdqrNsykmeftalIV
whyndrqcekf	1042A	1.00	mptgsivLA
cwfmdyivgphslna	1043R	1.00	eqTRK
wmifcvl	1044D	1.00	yrhptaksQenGD
hwdqnperecgskytafm	1045I	1.00	ILV

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.



mutation t@sting

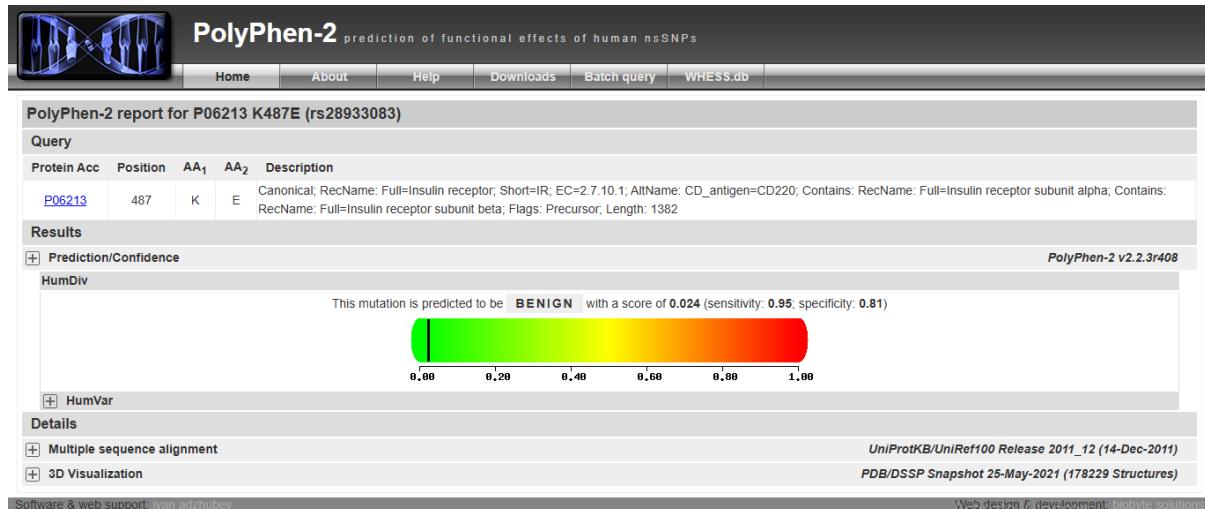
Prediction:	Deleterious	Permalink
Summary:		
	<ul style="list-style-type: none"> Amino acid sequence changed Known disease mutation at this position (HGMD CM890077) Known disease mutation: ClinVar ID 14679 (pathogenic) Protein features (might be) affected 	
Analysed issue	Analysis result	
Phys. location	chr19:7125448C>A show variant in all transcripts IGV	
Gene symbol	INSR	
ExAC LOF metrics	LOF: 0.19, missense: 5.21, synonymous: 0.66	
Ensembl transcript ID	ENST00000302850.5	
Genbank transcript ID	NM_000208 (exact from MANE)	
UniProt peptide	P06213	
Variant type	Single base exchange	
Gene region	CDS	Activate W Go to Settings
DNA changes	c.3104G>T g.168598G>T	

Protein conservation	Species	Match	Gene	AA Alignment
Human mutated	not conserved			1035 TLLRELQGGSFIVYEGIARDIIK
Ptroglobutes	all identical	ENSPTRG00000010386	1035 TLLRELQGGSFIVYEGIARDIIK	
Mmulatta	all identical	ENSMUG00000026907	697 TLLRELQGGSFIVYEGN	
Fcatus	all identical	ENSPCA00000003030	1034 TLLRELQGGSFIVYEGIARDIV	
Mmusculus	no alignment	ENSMUSG000000005534	n/a	
Ggallus	all identical	ENSALG000000040758	982 TLLRELQGGSFIVYEGIAKDIV	
Trubripes	all identical	ENSTRUG00000007600	1020 LOGQTIVYEGIAKDIV	
Driero	all identical	ENSDARG00000071524	1001 MRELQGGSFIVYEGIAKDII	
Melanogaster	no homologue			
Celegans	no homologue			
Xtropicalis	all identical	ENSXETG00000013518	1027 SLLRELQGGSFIVYEGIAKDII	

Original gDNA sequence snippet	GCTGGGGCAGGGCTCTTCGGCATGGTGTATGAGGGCAATG
Altered gDNA sequence snippet	GCTGGGGCAGGGCTCTTCGTCATGGTGTATGAGGGCAATG
Original cDNA sequence snippet	GCTGGGGCAGGGCTCTTCGGCATGGTGTATGAGGGCAATG
Altered cDNA sequence snippet	GCTGGGGCAGGGCTCTTCGTCATGGTGTATGAGGGCAATG
Wildtype AA sequence	MATGGRRGAA AAPLIVAWAA LLGGAGAHHY PGEVCPGNDI RNNLRLHIEL ENCSVIEGHL QLLMPKTRP EDIFDLSPFK LDKITDYLIL FRRVYGLESIK DLFPNLTVIR GSILFNFYAL VIFEMMHLKE LGLYNLMNLT RSQVRIEKNM ELCYLATIOM SRILSDEVN YVILNKDNE ECGDLCPITA KGGKNCIPATV INQFQVERHN THSHCQKICP TICKSHGCTA EGLCCHSEL QNCSSQDPTT KVA�CNHYL DQCVETCPY PYMMHQWRC VNFSFCQDQLH HKKNSRSHQG CHQVYHNNK CIECPGSGYT PNSNSLCLCT CLGPCKVCH LLEGEKTIDS VISAQEJNGE TVINGSLLIN IRGNNLAAKE LEANLGLIEE ISGYLK1KRS YALVSLSSFR KHLINGETL EIGWYFVAL DNQNLHQDQ WSKHMLLTQ GKLFFHYNPK LCLSEEHMKH EVSGTKHQEQE RNDIALKING DQASCENELL KFSYVLTTSF KILLRHEPMY PPOFDHLLQF MLYKEAHPYQ MVEFPGQDA CGSNSTVNAIDPPLRSNPD KSQNHPHQWL RGLKPNQYQA IPVKTLVTF5 DERRTYGAKS DIZYVQDATT MHSVPLDPI5 VSNESSQIQL KWPKPSDPMG NTHYLVWHE RQADESELFFE LDYCLKGKLK PSRWTSPPIPE SEDSQKHMQS EYEDASAECC SCPKTDSQIL KELEESFSRK TFEDYLNHVW FWRKTSQSGT GAEDHPMSRK RHLSDQYDVY TVAPTVAAF PNTSISVSPY SPEEHPRPEK VNKESLVSIS GLRHFHTGKRI ELQACNQDQTP EERCSVAAYV SARTMPEAKA DDIIVGPVTHA IFENNWHMHM WQEKEPMPEK LVWVSYRHY YQEELHLCY SIRKHFALERG CBLRGLSPN YSVRILATSL AGNSKSTEPF YFYVTDYLVQ PNSIAKLIIG PLIPVFLFSV VZGSYLYFLFR KNPQDQPLGP LYASSSMEYL SASOYPCVCS VYDNEWEVSR EKITTLLRELG QGSFWMYEG NARQKLGKA ETRVAKVTKN ESASLNERIE FNEASWSWKG F7CHWYVILL GVSKCQGPTL VNHLMAMQD LIKSYLSRSP EABMNCRP PTLQEMDQMA AEIADQHAYL NAKKFVPHD AHNINCWHD FTVKIQDHM TROIYETDY HKKGKGLLPV RWMPAESLKD GUFTTSZDM4 SFGVLMVETT SLAEQYVQGL SNEQVLKFWM DGGYDQDOPN OPERVTDLNR MCQWNPKNPK PTLFEVNLW KDLQHPSFPE VSFFFHSEEMN APSEELLEHE FEDMENVPLD RSHSCOREEA GORDDGSSLQ KRISYEEHID YTMHNGKKN GRLLTLPNSR PS*
Mutated AA sequence	MATGGRRGAA AAPLIVAWAA LLGGAGAHHY PGEVCPGNDI RNNLRLHIEL ENCSVIEGHL QLLMPKTRP EDIFDLSPFK LDKITDYLIL FRRVYGLESIK DLFPNLTVIR GSILFNFYAL VIFEMMHLKE LGLYNLMNLT RSQVRIEKNM ELCYLATIOM SRILSDEVN YVILNKDNE ECGDLCPITA KGGKNCIPATV INQFQVERHN THSHCQKICP TICKSHGCTA EGLCCHSEL QNCSSQDPTT KVA�CNHYL DQCVETCPY PYMMHQWRC VNFSFCQDQLH HKKNSRSHQG CHQVYHNNK CIECPGSGYT PNSNSLCLCT CLGPCKVCH LLEGEKTIDS VISAQEJNGE TVINGSLLIN IRGNNLAAKE LEANLGLIEE ISGYLK1KRS YALVSLSSFR KHLINGETL EIGWYFVAL DNQNLHQDQ WSKHMLLTQ GKLFFHYNPK LCLSEEHMKH EVSGTKHQEQE RNDIALKING DQASCENELL KFSYVLTTSF KILLRHEPMY PPOFDHLLQF MLYKEAHPYQ MVEFPGQDA CGSNSTVNAIDPPLRSNPD KSQNHPHQWL RGLKPNQYQA IPVKTLVTF5 DERRTYGAKS DIZYVQDATT MHSVPLDPI5 VSNESSQIQL KWPKPSDPMG NTHYLVWHE RQADESELFFE LDYCLKGKLK PSRWTSPPIPE SEDSQKHMQS EYEDASAECC SCPKTDSQIL KELEESFSRK TFEDYLNHVW FWRKTSQSGT GAEDHPMSRK RHLSDQYDVY TVAPTVAAF PNTSISVSPY SPEEHPRPEK VNKESLVSIS GLRHFHTGKRI ELQACNQDQTP EERCSVAAYV SARTMPEAKA DDIIVGPVTHA IFENNWHMHM WQEKEPMPEK LVWVSYRHY YQEELHLCY SIRKHFALERG CBLRGLSPN YSVRILATSL AGNSKSTEPF YFYVTDYLVQ PNSIAKLIIG PLIPVFLFSV VZGSYLYFLFR KNPQDQPLGP LYASSSMEYL SASOYPCVCS VYDNEWEVSR EKITTLLRELG QGSFWMYEG NARQKLGKA ETRVAKVTKN ESASLNERIE FNEASWSWKG F7CHWYVILL GVSKCQGPTL VNHLMAMQD LIKSYLSRSP EABMNCRP PTLQEMDQMA AEIADQHAYL NAKKFVPHD AHNINCWHD FTVKIQDHM TROIYETDY HKKGKGLLPV RWMPAESLKD GUFTTSZDM4 SFGVLMVETT SLAEQYVQGL SNEQVLKFWM DGGYDQDOPN OPERVTDLNR MCQWNPKNPK PTLFEVNLW KDLQHPSFPE VSFFFHSEEMN APSEELLEHE FEDMENVPLD RSHSCOREEA GORDDGSSLQ KRISYEEHID YTMHNGKKN GRLLTLPNSR PS*

VARIANT c.1459A>G (p. Lys487Glu)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 0.024.



SIFT- Sorting Intolerant From Tolerant

Predicted the variant under tolerated

w	478R	1.00	cdfymhnpgesqtIlvAkR
wcfmy	479Q	1.00	hgpnvltsraIkdeQ
wy	480E	1.00	fcmhIpvlgrnqTDASkE
wcfymhi	481R	1.00	vlpgntsqaRDEK
w	482N	1.00	cfymhivplrGqTNdsAkE
cwmfilylvhrtgspnakq	483D	1.00	ED
whcqydrenkpmfg	484I	1.00	atlvsI
w	485A	1.00	fcmiyihlvpргQtdeASN
w	486I	1.00	hycdqrgfenPmkiasvLT
wcfy	487K	1.00	mhivpgndtLqasERK
wfymhclireqgkvkpda	488T	1.00	NST
ywvtsrqpmmlkihgfedca	489N	1.00	N
ywvtsrqpnmlkihfedca	490G	1.00	G
wcmfivlyrhptakqsg	491D	1.00	NED
cwfdyivgphsn	492Q	1.00	taleMkQR
ywvtsrqpnmlkihgfedc	493A	1.00	A
wfyhmicl	494S	1.00	rvndpketgQAS
ywvtsrqpnmlkihgfed	495C	1.00	C
wcfymihvl	496E	1.00	pgrntakdQSE
	497N	1.00	wcmphIFgvrlqlyDNkeAST

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.



mutation t@sting

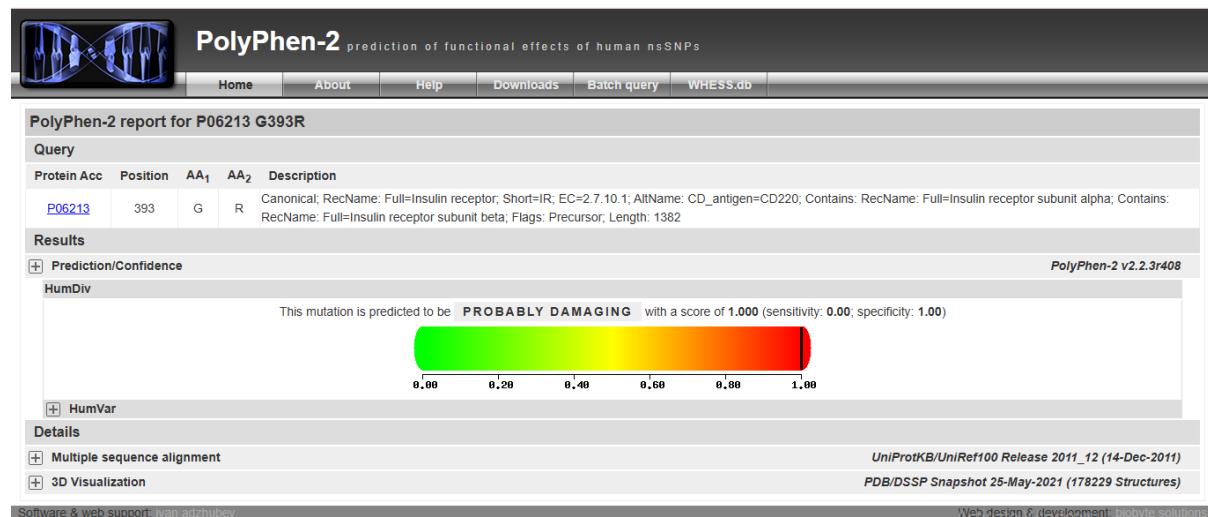
Prediction:	Deleterious	Permalink
Summary:	<ul style="list-style-type: none"> Amino acid sequence changed Known disease mutation at this position (HGMD CM880048) Known disease mutation: ClinVar ID 14680 (pathogenic) Protein features (might be) affected 	<ul style="list-style-type: none"> Model: simple_aae Tree vote: 85 15 (del benign) Automatic classification due to ClinVar
Analysed issue	Analysis result	
Phys. location	chr17:170572T>C show variant in all transcripts IGV	
Gene symbol	INSR	
ExAC LOF metrics	LOF: 0.19, missense: 5.21, synonymous: 0.66	
Ensembl transcript ID	ENST00000302850.5	
Genbank transcript ID	NM_000208 (exact from MANE)	
UniProt peptide	P06213	
Variant type	Single base exchange	
Gene region	CDS	Activate W Go to Settings
DNA changes	c.1459A>G g.123474A>G	

Protein conservation	Species	Match	Gene	AA	Alignment
Human mutated	all conserved			487	KGRQERNIDALTKINGQDQASCNEEL
Prologotes	all identical	ENSPTRG00000010386		487	RNDIALTNTNGQDQASCNE
Mmulatta	all identical	ENSMUUG00000028907		435	RNDIALTNTNGQDQASCNE
Fcatus	all identical	ENSFAGM00000003030		487	RNDIALTNTNGQDQASCNE
Mmusculus	no alignment	ENSMUSG00000005534		n/a	
Ggalus	all identical	ENSGALG00000040758		447	KGRQERNIDALTKINGQDQASCNE
Trubripes	all identical	ENSTRUG00000007600		489	KQRQVRNDIASNTNGDQA
Dreio	not conserved				ENSDARG00000071524
Dmelanogaster	no homologue			476	KNRH--NDIS--NEKDQ
Celebensis	no homologue				
Xtropicalis	all identical	ENSXETG00000013518		496	KGRQDKNDISTNTNGQDQASCEDN

Original gDNA sequence snippet	AGAGAAAAGCACATTGCCCTG A AGACCAATGGGGACCAAGCA
Altered gDNA sequence snippet	AGAGAAAAGCACATTGCCCT G AGACCAATGGGGACCAAGCA
Original cDNA sequence snippet	AGAGAAAAGCACATTGCCCTG A AGACCAATGGGGACCAAGCA
Altered cDNA sequence snippet	AGAGAAAAGCACATTGCCCT G AGACCAATGGGGACCAAGCA
Wildtype AA sequence	MATGORRGA AAPLIVAVAA LLGLGAQHLS PGEVCGNDI RNNLTLRHL ENCSVIEGH QILLMPKTRP EDFDOLSPFK LMDITDILL FRVYGELESLK DLFLNMLTVIR GSRLFFNVAL VIFEMMHILKE LGGLNLMNIT RGSVRLVKEKNE ELYCLATIOMN SHILDSVDEN YZVLNKDNE EGCQICPGTA KOKTNCPATV INQQRVERCH THSHQCKVCP TICKSHGCTA EGLGCHESEL GNC5QDQPT KCVACHNFYL DRCVETCIP PYHHFQDQWRC VNF-SFCQDQLH HKCKNSRROQ CHOVYHNMNK CIEPCPSGTY MNGSNLLCTP CLGPQPKVCH LLEGEKTIDS VTSQAELRQC TVNGSLLIN INGGONLAAE LEANLGLIEE ISGYLKIRRS YALVSLSPFR KLRLIRGETL EIGNYSFVAL DNQNLHQWD WSKHMLTITQ GKLFHYHPX LCLSLSEHME EVSGTKRQE RNDIALTKING QDASECNELL KFSYVIRTSFQ KILLINWEPWY PFDORFLLG MFYFKEAPYQ MTVEFDGQDA CGSNSMVTWD IDPPLRSNDI KSONHPGMLN RGLKPATQYA IPVKTLTVFS DERRTYGAKS DIIVYQDABT NPSVPLDPI5 VSNSSSQJII WNKPPSDNG NITHYLVFWNE RQAEDSELF EDVCLGLKL PSRTNSPFFE SEDSOXKHOOS EYESDAGECC SCPKTDQSQL KELEESFSRK TFEQYLNHVV FVPRTKTSGT GAEDPRPSRK RHSLGQDWNV TVAWPTVAAF PNTS5TSPVTP SPEEHRIPEK VVVKESLVI5 GLRHFGTGYZL ELQACQDQTP EERCSVAAYV SARTWPEAKA DDIVGIVTHE IFENNWHUH WQEPKEEPLG IVLVEVSYRR YQDEELHLCV SIRKHFALERG CHRLGLSPQN YSVRLRATSL AONGSWTEPT YFYVTDYLVW PSNLAKIIIG PLIFVFLPSV VIGSILYFLR KHOQDPGLP LYASSNPYEL SASOVPCSV YVPMEMEVSR EKITLRLRELQ QGSFGMVYVEG NARDIKIGEA ETIRAVKTVN ESASLHIEE FLINEASVMKG FTCHMMVILL GVVKSGQQTl VMELMNHGD LKSVLRSLRP EAENNPGRPP PTLOEMIQMA AEIADDMAYL NAKKFVHMDL AARNCMVWAHD FTVKLGQFON TRDIZETDYY RKKGGKLPP RMAPAESLKD GVFTTSDWM SFGVVLWEIT SLAEQPYOGI SNEQVLFKVN DGGYLDQDQN OPERVTDLWR MCWQHNPMMR PTFLIEVNL KDLHSPFPE VSFHSEENK APESEELEME FEDMENVPLD RSSHCQHEEA GGRDGSSSLG FKSYEEHL P YTHMGQKKN GRILTLPHSN PS*
Mutated AA sequence	MATGORRGA AAPLIVAVAA LLGLGAQHLS PGEVCGNDI RNNLTLRHL ENCSVIEGH QILLMPKTRP EDFDOLSPFK LMDITDILL FRVYGELESLK DLFLNMLTVIR GSRLFFNVAL VIFEMMHILKE LGGLNLMNIT RGSVRLVKEKNE ELYCLATIOMN SHILDSVDEN YZVLNKDNE EGCQICPGTA KOKTNCPATV INQQRVERCH THSHQCKVCP TICKSHGCTA EGLGCHESEL GNC5QDQPT KCVACHNFYL DRCVETCIP PYHHFQDQWRC VNF-SFCQDQLH HKCKNSRROQ CHOVYHNMNK CIEPCPSGTY MNGSNLLCTP CLGPQPKVCH LLEGEKTIDS VTSQAELRQC TVNGSLLIN INGGONLAAE LEANLGLIEE ISGYLKIRRS YALVSLSPFR KLRLIRGETL EIGNYSFVAL DNQNLHQWD WSKHMLTITQ GKLFHYHPX LCLSLSEHME EVSGTKRQE RNDIALETING QDASECNELL KFSYVIRTSFQ KILLINWEPWY PFDORFLLG MFYFKEAPYQ MTVEFDGQDA CGSNSMVTWD IDPPLRSNDI KSONHPGMLN RGLKPATQYA IPVKTLTVFS DERRTYGAKS DIIVYQDABT NPSVPLDPI5 VSNSSSQJII WNKPPSDNG NITHYLVFWNE RQAEDSELF EDVCLGLKL PSRTNSPFFE SEDSOXKHOOS EYESDAGECC SCPKTDQSQL KELEESFSRK TFEQYLNHVV FVPRTKTSGT GAEDPRPSRK RHSLGQDWNV TVAWPTVAAF PNTS5TSPVTP SPEEHRIPEK VVVKESLVI5 GLRHFGTGYZL ELQACQDQTP EERCSVAAYV SARTWPEAKA DDIVGIVTHE IFENNWHUH WQEPKEEPLG IVLVEVSYRR YQDEELHLCV SIRKHFALERG CHRLGLSPQN YSVRLRATSL AONGSWTEPT YFYVTDYLVW PSNLAKIIIG PLIFVFLPSV VIGSILYFLR KHOQDPGLP LYASSNPYEL SASOVPCSV YVPMEMEVSR EKITLRLRELQ QGSFGMVYVEG NARDIKIGEA ETIRAVKTVN ESASLHIEE FLINEASVMKG FTCHMMVILL GVVKSGQQTl VMELMNHGD LKSVLRSLRP EAENNPGRPP PTLOEMIQMA AEIADDMAYL NAKKFVHMDL AARNCMVWAHD FTVKLGQFON TRDIZETDYY RKKGGKLPP RMAPAESLKD GVFTTSDWM SFGVVLWEIT SLAEQPYOGI SNEQVLFKVN DGGYLDQDQN OPERVTDLWR MCWQHNPMMR PTFLIEVNL KDLHSPFPE VSFHSEENK APESEELEME FEDMENVPLD RSSHCQHEEA GGRDGSSSLG FKSYEEHL P YTHMGQKKN GRILTLPHSN PS*

VARIANT c.1177G>A (p. Gly393Arg)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 1.000.



SIFT- Sorting Intolerant From Tolerated

Predicted the variant under tolerated

384N	1.00	wmciivFprlqhyekatgsdN
dhgnceswrkypqtafvi	385L	1.00 ML
wmfihyclrvektpdn	386G	1.00 aSG
whcdyeyqrnkpg	387L	1.00 tfmvaisl
hwqdpnercrgksymfatl	388I	1.00 VI
wcfmyihlvrqpn	389E	1.00 kqadTE
w	390E	1.00 yhfcmgrnpqkdlsiaTEV
hwqdpnercrgksymfat	391I	1.00 lVI
wfymc	392S	1.00 hilvprqdenGaKST
wmifcvlyqrptek	393G	1.00 asdnHG
qknrhdgepctsmaiviwl	394Y	1.00 FY
hdwgnercsqykpta	395L	1.00 fmIVL
wfcmyihdlvnpt	396K	1.00 qgesrAK
hwqdpnercrgksymftal	397I	1.00 VI
cwfd	398R	1.00 myigphsnlvtaeqRK
cwfmi	399R	1.00 ydvgpslatneqHkR
ywvtrqpnmlkihgfedca	400S	1.00 S

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.



mutation t@sting

Prediction:	Deleterious	Permalink
Summary:	<ul style="list-style-type: none"> Amino acid sequence changed Known disease mutation at this position (HGMD CM920377) Known disease mutation: ClinVar ID 14681 (pathogenic) Protein features (might be) affected 	<ul style="list-style-type: none"> Model: simple_aae Tree vote: 98 2 (del benign) Automatic classification due to ClinVar
Analysed issue	Analysis result	
Phys. location	chr19:7172392C>T show variant in all transcripts IGV	
Gene symbol	INSR	
ExAC LOF metrics	LOF: 0.19, missense: 5.21, synonymous: 0.66	
Ensembl transcript ID	ENST00000302850.5	
Genbank transcript ID	NM_000208.(exact from MANE)	
UniProt peptide	P06213	
Variant type	Single base exchange	
Gene region	CDS	Activate Wi Go to Settings
DNA changes	c.1177G>A g.121654G>A	

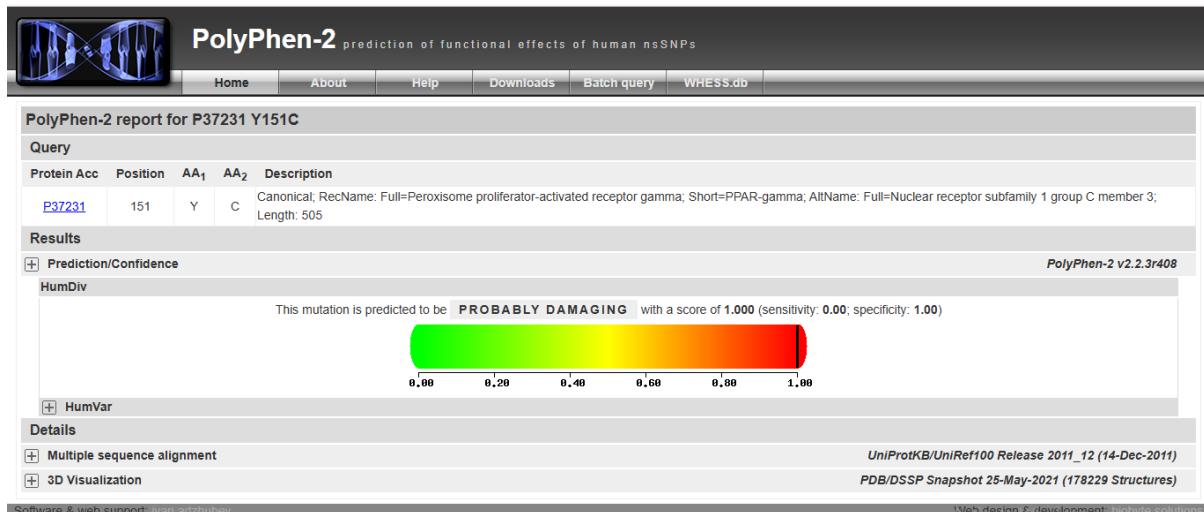
Protein conservation	Species	Match	Gene	AA Alignment
	Human	not conserved		393 EANLGLIEEISGYLKIRRSYALVS
Ptroglobutes	all identical	ENSPTRG0000010386		393 EANLGLIEEISGYLKIRRSYALV
Mmulatta	all identical	ENSMULG00000028907		359 EANLGLIEEISGYLK-RRSYALV
Fcatus	all identical	ENSFAG00000003030		393 EANLGLIEEISGYLKIRRSYALV
Mmusculus	no alignment	ENSMUSG00000005534		n/a
Ggalus	all identical	ENS GAL000000040758		354 EANLGLIEEISGYLK
Trubripes	all identical	ENSTRUG00000007600		393 EASLQLEETGYLTVRSSYALV
Drerio	all identical	ENS DARG000000071524		382 ESSLQLEETGYLTIRRAYALV
Dmelanogaster	no homologue			
Celegans	no homologue			
Xtropicalis	all identical	ENSXETG00000013518	402	SGYLKIRRSYALV

Original gDNA sequence snippet	GCCTCATTAAGAAAATTTCAGGGTATCTAAAAATCCCGCGA
Altered gDNA sequence snippet	GCCTCATTAAGAAAATTTCAGGTATCTAAAAATCCCGCGA
Original cDNA sequence snippet	GCCTCATTAAGAAAATTTCAGGGTATCTAAAAATCCCGCGA
Altered cDNA sequence snippet	GCCTCATTAAGAAAATTCAAGGTATCTAAAAATCCCGCGA
Wildtype AA sequence	<p>MTGGRGAA AAPLLVAVAA LLLGAAGHLY PGEVCPGMID RNNNTRLHEL ENCSVIEGHL QILLMPKTRP EDPRDLSPKI LIMDTDYL RIVYGELESK DLFFNLTIVR GSRLPNFVAL VIFEMMHILKE LGLYNLNITI RGSVRLIEKKN ELCYLATIDM SRLILSVDEN YIVLNKNDNE ECGDICPGTA KGTNCRATV INGOFVERICL THSHQCKVCP TICKSHGCTA EGLCCHSCL QCNSQDQPT KVCACRNHYL DGRCEVTCPP PYHYPQDWRV VNFSFCQDLH HKKCNSHHQG CHQVHLHNK CPECPGSGTY MWSNLLCCTP CLGICPKVCH LLEGEKTIDS VTSQJELRG CTVNGSLLN IRGONNLAE LEANLGLIEE ISRYLKIRHS YALVLSLFFR KKLHLRGETL EIGYISFYAL DNQNLRQWD WSKHMLTTIQ GKLFPHYNPK LCLSEHHOKW EVSGTGKQDE RNDALKTKE DOASCENELL KFSYIRTSDF KILLNWEPMY PIOPFRDLLGF MLF-YEAHQ IWTFEGQDA CGSNSNTVVO IDPLPLSNPD KSONNPQHNL RIGLKPQTOYA IPVKTLVTF5 DEHRTYGAKS DIZYVQDADT NPSVPVLDPIS VSNSQSIU11 KWKPPSDPMH NITHYLVN6 ROADESELPE LDYCLKGKLLK PSRMTWSPPIP SEOSDNKHO5 EYESDAGECC SCPKTD5QL KELEESFSHPF TFEQYLHNVN FVPKRISSTGT GAEDHRPSRK RRSLGQDVGN TVAPVTVAAF PNTSISVPTP SPEEHMPEK VVNEKESLVS GLRHHTGYZ ELQACNQDTP EECRSVAAYV SARTMPEKA DOLIVGPVTHE IFENNVVHML WOEPEKINGL IVLYVEYSR YQEEELHLCV SRKHFALERG CRILKLSQHNG YSVRIRATEL AGQGNTATF YPYFTYUFLD PSNHLAKLIG PLIPVFLFSV VIGSISYFLR KIOPQGPGLP LYASNSNPEYL SASDNPVPCSV YVPOHEVSR EKILLTRELG QGSFGRMYES NAROIIKXGA ETIVAKVTKW ESASLIERIE FLNEASVWKG FCTHMVRLL GVVSKGQPTI VVNELNHAHDG LKSYLRLSLRP EAENNPGRPP PTLEHOMQPA AEIADDNMYL NAKKKVVHRL AABRNCHVARD FTVKLGQDFH TROVETDYDYY IKGGKGLLPV RMWAPESLKD QVFTTSSQHNA SGFGLVMEIT SLAEPOYGLG SNEQVLFKVN DGGYQDOPDN CPERVTLUDK MCQWNPKNH PTFLEVNKA KDOLMPSFPE VFSSHSEEEEM APSEELEME FEDHENVPLD RSSHCQREEA GGRDGSSSLG FKRSYEEHIDP YTHNNGKKN GRLLTLPNS PS*</p>
Mutated AA sequence	<p>MTGGRGAA AAPLLVAVAA LLLGAAGHLY PGEVCPGMID RNNNTRLHEL ENCSVIEGHL QILLMPKTRP EDPRDLSPKI LIMDTDYL RIVYGELESK DLFFNLTIVR GSRLPNFVAL VIFEMMHILKE LGLYNLNITI RGSVRLIEKKN ELCYLATIDM SRLILSVDEN YIVLNKNDNE ECGDICPGTA KGTNCRATV INGOFVERICL THSHQCKVCP TICKSHGCTA EGLCCHSCL QCNSQDQPT KVCACRNHYL DGRCEVTCPP PYHYPQDWRV VNFSFCQDLH HKKCNSHHQG CHQVHLHNK CPECPGSGTY MWSNLLCCTP CLGICPKVCH LLEGEKTIDS VTSQJELRG CTVNGSLLN IRGONNLAE LEANLGLIEE ISRYLKIRHS YALVLSLFFR KKLHLRGETL EIGYISFYAL DNQNLRQWD WSKHMLTTIQ GKLFPHYNPK LCLSEHHOKW EVSGTGKQDE RNDALKTKE DOASCENELL KFSYIRTSDF KILLNWEPMY PIOPFRDLLGF MLF-YEAHQ IWTFEGQDA CGSNSNTVVO IDPLPLSNPD KSONNPQHNL RIGLKPQTOYA IPVKTLVTF5 DEHRTYGAKS DIZYVQDADT NPSVPVLDPIS VSNSQSIU11 KWKPPSDPMH NITHYLVN6 ROADESELPE LDYCLKGKLLK PSRMTWSPPIP SEOSDNKHO5 EYESDAGECC SCPKTD5QL KELEESFSHPF TFEQYLHNVN FVPKRISSTGT GAEDHRPSRK RRSLGQDVGN TVAPVTVAAF PNTSISVPTP SPEEHMPEK VVNEKESLVS GLRHHTGYZ ELQACNQDTP EECRSVAAYV SARTMPEKA DOLIVGPVTHE IFENNVVHML WOEPEKINGL IVLYVEYSR YQEEELHLCV SRKHFALERG CRILKLSQHNG YSVRIRATEL AGQGNTATF YPYFTYUFLD PSNHLAKLIG PLIPVFLFSV VIGSISYFLR KIOPQGPGLP LYASNSNPEYL SASDNPVPCSV YVPOHEVSR EKILLTRELG QGSFGRMYES NAROIIKXGA ETIVAKVTKW ESASLIERIE FLNEASVWKG FCTHMVRLL GVVSKGQPTI VVNELNHAHDG LKSYLRLSLRP EAENNPGRPP PTLEHOMQPA AEIADDNMYL NAKKKVVHRL AABRNCHVARD FTVKLGQDFH TROVETDYDYY IKGGKGLLPV RMWAPESLKD QVFTTSSQHNA SGFGLVMEIT SLAEPOYGLG SNEQVLFKVN DGGYQDOPDN CPERVTLUDK MCQWNPKNH PTFLEVNKA KDOLMPSFPE VFSSHSEEEEM APSEELEME FEDHENVPLD RSSHCQREEA GGRDGSSSLG FKRSYEEHIDP YTHNNGKKN GRLLTLPNS PS*</p>

PPARG

VARIANT c.362A>G (p. Tyr121Cys)

POLY PHEN 2 – Predicted the variant to be a probably damaging kind with a score of 1.000.



SIFT- Sorting Intolerant From Tolerant

Predicted the variant under tolerated

Predict Not Tolerated	Position	Seq	Rep	Predict Tolerated
w h y f m i	101S	0.33	r q c l e d k n v G P T A S	
w	102N	0.33	m f i y C h l v r p Q t k e d S G N A	
w y f c m h i l	103S	0.40	p n r d t V e G K Q A S	
	104L	0.38	c w p d m e q k n r G T S I v h A F L Y	
d h g n c s w r y k p q t	105M	0.45	a E f v M I L	
w h y	106A	0.50	f r q m d c e k N p l g s i t V A	
h d w n e c p g q r s k y t a f	107I	0.52	V M L I	
c w f m y i h v r g n t	108E	0.52	s L k a P q D E	
y w v t s r q p n m l k i h g f e d a	109C	0.98	C	
c w d m i y v	110R	0.98	g p s h n a l t e q F K R	
h w q d p n e r c g k s y m f a l	111V	0.98	I V	
y w v t s r q p n m l k i h g f e d a	112C	0.98	C	
m w i f v l c y r p h t k e a	113G	0.98	d S N Q G	
y w v t s r q p n m l k i h g f e c a	114D	0.98	D	
c w d f m y i g p s h n l a t e q	115K	0.98	V R K	
w h y f m i q r n d e l c k v t p g	116A	0.98	S A	
y w v t r q p n m l k i h g f e d c a	117S	0.98	S	
y w v t s r q p n m l k i h g f e d c a	118G	0.98	G	
n k q r h d g e p c t s a m v w l	119F	0.98	I Y F	
y w v t s r q p n m l k i h g f e d c a	120H	0.98	H	
w v t s r q p n m l k i h g f e d c a	121Y	0.98	Y	
w h y f m i q r n e l c d k v t p s A	122G	1.00	G	
y w t s r q p n m l k i h g f e d c a	123V	1.00	V	
w f m y c e r q	124H	1.00	g l d p k v n a I T S H	
w h y f m i r q c e l d k n v p g	125A	1.00	T S A	
m w i f y h q l k r v e p t s d a g N	126C	1.00	C	
y w v t s r q p n m l k i h g f e d c a	127E	1.00	E	
m w i f c v y l r h q t k p e a s n D	128G	1.00	G	

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.

Mutation Taster

mutation t@sting

Prediction: Deleterious [Permalink](#)

Summary:

- Amino acid sequence changed
- Known disease mutation at this position (HGMD CM114674)

• Model: simple_aae
 • Tree vote: 100|0 (del | benign)

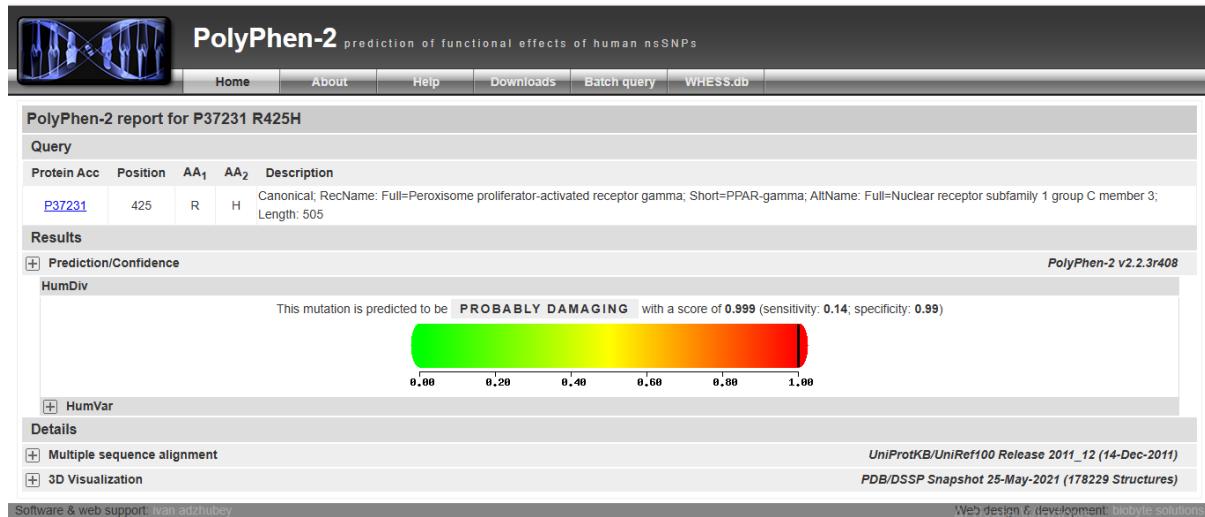
Analysed issue	Analysis result
Phys. location	chr3:12422962A>G show variant in all transcripts IGV
Gene symbol	PPARG
ExAC LOF metrics	LOF: 0.67, missense: 2.28, synonymous: -0.49
Ensembl transcript ID	ENST00000539812.1
Genbank transcript ID	
UniProt peptide	P37231
Variant type	Single base exchange
Gene region	CDS
DNA changes	c.362A>G g.94096A>G

Protein conservation	Species	Match	Gene	AA	Alignment
	Human	not conserved		121	RVCGD[KASGFHYGVHACEGCKGF
	mutated			121	CGVHACEGCKGF
	Ptroglodytes	all identical	ENSPTRG00000014632	151	YGVHACEGCKGF
	Mmutatta	all identical	ENSMMUG00000007191	121	YGVHACEGCKGF
	Fcatus	all identical	ENSFCAG00000006231	151	YGVHACEGCKGF
	Mmusculus	all identical	ENSMSU00000000440	121	YGVHACEGCKGF
	Ggallus	all identical	ENS GALG00000004974	121	YGVHACEGCKGF
	Trubripes	all identical	ENSTRUG00000015522	147	RVCGD[KASGFYGVHACEGCKGF
	Drerio	no homologue			
	Dmelanogaster	no homologue			
	Celegans	no homologue			
	Xtropicalis	no alignment	ENSXETG00000017422	n/a	

Original gDNA sequence snippet	TAAAGCTTCTGGATTTCACT A TGGAGTTCATGCTTGAAAG
Altered gDNA sequence snippet	TAAAGCTTCTGGATTTCACT G TGGAGTTCATGCTTGAAAG
Original cDNA sequence snippet	TAAAGCTTCTGGATTTCACT A TGGAGTTCATGCTTGAAAG
Altered cDNA sequence snippet	TAAAGCTTCTGGATTTCACT G TGGAGTTCATGCTTGAAAG
Wildtype AA sequence	MVDTEMPPWP TNFGISSLVL SVNEDHSHSF DIKPFTTVDF SSISIPTHYED IPFTRTDPVV ADYKYDLKLQ EYOSAIKVEP ASPPYSEKT QLYNPKHEEP SNSLMAIECR VCGDKASGFH YGVHACEGCK GFRRRTIRLK LIYDRCNLNC RIHKKSRSNKC QCRCFQKCLLA VGMSHNAIRF GRMPQAEEK LLAEISSSDID QLNIPESADLR ALAKHLYDYSY IKSFPPLTKAK ARAZLTGKTT DKSTAQVC*
Mutated AA sequence	MVDTEMPPWP TNFGISSLVL SVNEDHSHSF DIKPFTTVDF SSISIPTHYED IPFTRTDPVV ADYKYDLKLQ EYOSAIKVEP ASPPYSEKT QLYNPKHEEP SNSLMAIECR VCGDKASGFH C GVHACEGCK GFRRRTIRLK LIYDRCNLNC RIHKKSRSNKC QCRCFQKCLLA VGMSHNAIRF GRMPQAEEK LLAEISSSDID QLNIPESADLR ALAKHLYDYSY IKSFPPLTKAK ARAZLTGKTT DKSTAQVC*

VARIANT c.1184G>A (p. Arg395His)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 0.999.



SIFT- Sorting Intolerant From Tolerant

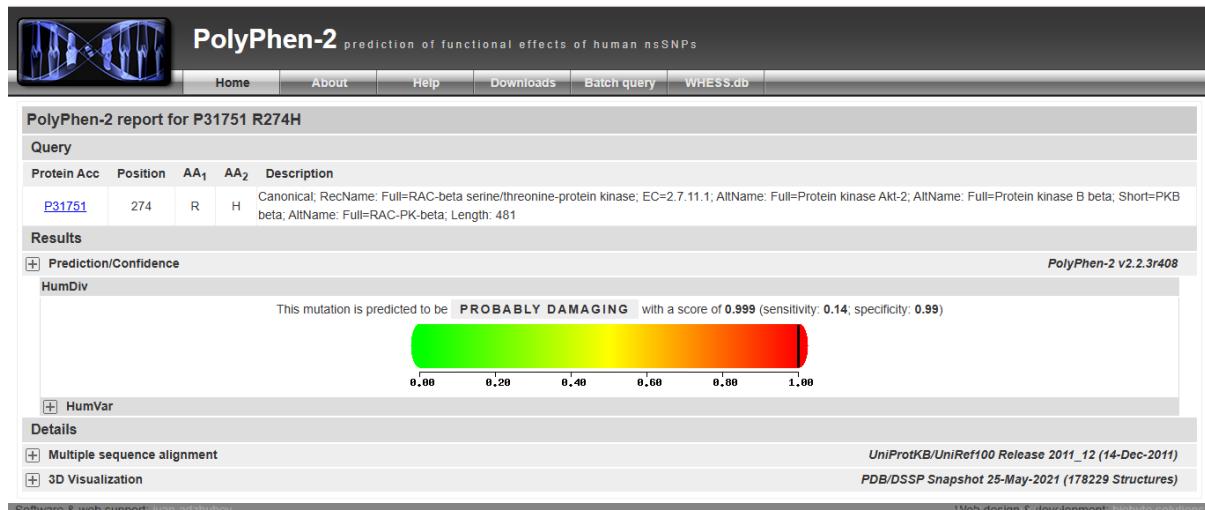
Predicted the variant under tolerated

dhgncswrkyppqtafmv	384I	1.00	IL
hdngkrecqsptawvmiy	385F	1.00	LF
wghydrfqne	386I	1.00	kmpClasTIV
whyfmiqrndelkcvtsg	387A	1.00	SA
whydngrqf	388V	1.00	ekcpmsl1AIv
hwqdnpregskyfmta	389I	1.00	LCVI
hdwneecgpqrskytafm	390I	1.00	VLI
hwdnqgerpskytfam	391L	1.00	CVLI
whfymlrieqldkvpng	392S	1.00	ATSC
whyfimqrndelkcvt	393G	1.00	sPAG
ywvtsrqpnmlkihgfea	394D	1.00	D
ywvtsrqpnmlkihgfedca	395R	1.00	R
whyfincrediekvt	396P	1.00	MgSAQP
mifvclyrqhptkeas	397G	1.00	nWDG
dghnecswyrkpqtafvM	398L	1.00	IL
dghwnccsypkqR	399L	1.00	EtafMVIL
wcmifvlyhrtpaskg	400N	1.00	dQEN

AKT2

VARIANT c.821G>A (p. Arg274His)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 0.999.



SIFT- Sorting Intolerant From Tolerated

Predicted the variant under tolerated

d h g n e c k s r p q t w a y v m i	266L	1.00	F L
m w c f i y v l t a p s e n d r g k	267H	1.00	Q H
w y f c	268S	1.00	m i H p v g L N D T Q R a K E S
w	269R	1.00	f m y C I p v H L G t d a N S Q e R K
mi w v f l c y r p q a t h s	270D	1.00	E K D N G
h q w p d n e c r k s g y a t m f l	271V	1.00	V I
h q w p d n e c r k s g y a t m f	272V	1.00	L V I
w v t s r q p n m l k i h g f e d c a	273Y	1.00	Y
w v t s r q p n m l k i h g f e d c a	274R	1.00	R
y w v t s r q p n m l k i l g f e c a	275D	1.00	D
d g h n e c s w y r k p q t a f v	276I	1.00	I M L
y w v t s r q p n m l i h g f e d c a	277K	1.00	K
h e a w d e y r k q s i g a f v m i	278L	1.00	P L
c w m f i i y l v r h p g s a k q N	279E	1.00	D E
y w v t s r q p n m l k i h g f e d c a	280N	1.00	N
h w d q p n e c r s g k y a t f m	281L	1.00	I V L
d g h n e c s w r k y p q t a f i V	282M	1.00	L M

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.



mutation t@sting

Prediction:

Deleterious

[Permalink](#)

Summary:

- Amino acid sequence changed
- Known disease mutation at this position (HGMD CM041234)
- Known disease mutation: ClinVar ID 13982 (pathogenic)
- Protein features (might be) affected
- Model: simple_aae
- Tree vote: 95|5 (del | benign)
- Automatic classification due to ClinVar

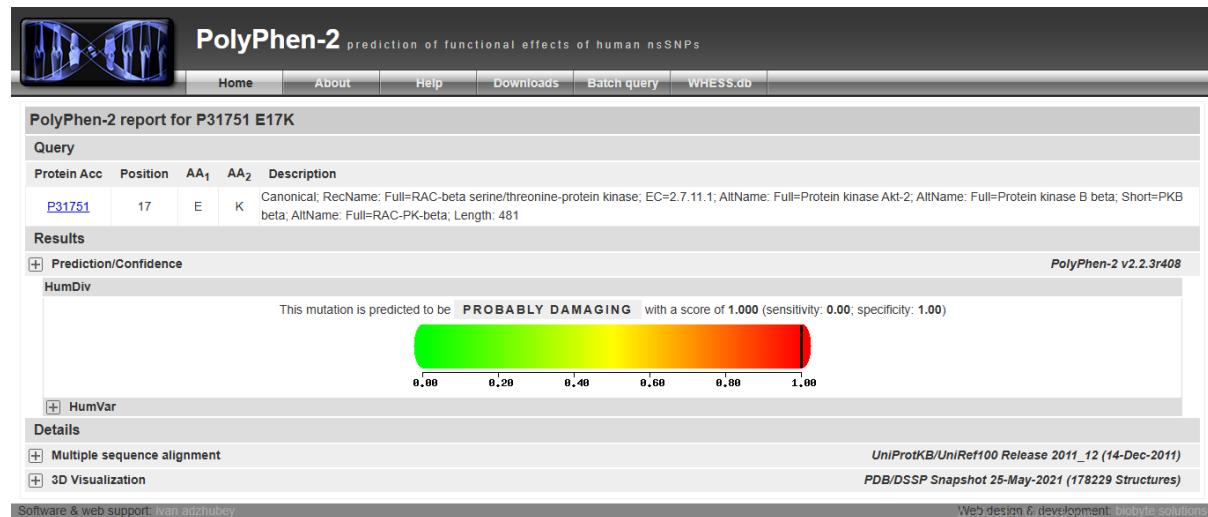
Analysed issue	Analysis result
Phys. location	chr19:40743886C>T show variant in all transcripts IGV
Gene symbol	AKT2
ExAC LOF metrics	LOF: 1.00, missense: 3.48, synonymous: 0.56
Ensembl transcript ID	ENST00000392038.2
Genbank transcript ID	NM_001626 (exact from MANE) , NM_001243027 (by similarity) , NM_001243028 (by similarity)
UniProt peptide	P31751
Variant type	Single base exchange
Gene region	CDS
DNA changes	c.821G>A g.47558G>A

Protein conservation				
	Species	Match	Gene	AA Alignment
Human	Human	not conserved		274 LEYLHSRDNVVYRDIKLENHLDK
mutated				274 LEYLHSRDNVVYRDIKLENHLDK
Ptroglydotes	all identical	ENSPTRG00000010992		212 LEYLHSRDNVV O IKLENHLDK
Mmulatta	all identical	ENSMUUG00000003621		274 LEYLHSRDNVV O IKLENHLDK
Fcatu	all identical	ENSFCAG000000012180		274 LEYLHSRDNVV O IKLENHLDK
Mmusculus	all identical	ENSMUSG00000004056		274 LEYLHSRDNVV O IKLENHLDK
Ggallus	all identical	ENSGAL000000054458		275 LEYLHSRDNVV O IKLENHLDK
Tribripes	all identical	ENSTRUG00000008633		211 LEYLHSRDNVV O IKLENHLDK
Dreno	all identical	ENSDARG000000011219		272 LEYLHSRDNVV O IKLENHLDK
Dmelanogaster	no homologue			
Celegans	no homologue			
Xtrropicalis	all identical	ENSXETG00000016986		274 LEYLHSRDNVV O IK-----

Original gDNA sequence snippet	CTCGGGGACGTGGTATACCG C GACATCAAGGTTAGGGCA
Altered gDNA sequence snippet	CTCGGGGACGTGGTATACCA C GACATCAAGGTTAGGGCA
Original cDNA sequence snippet	CTCGGGGACGTGGTATACCG C GACATCAAGCTGGAAAACC
Altered cDNA sequence snippet	CTCGGGGACGTGGTATACCA C GACATCAAGCTGGAAAACC
Wildtype AA sequence	MNEVSVIKEG WLHKRGEYIK TWRPRYFLLK SDGSFIGYKE RPEAPDQTLPLNNFSVAEC QLMKTERRP NTFVIRCLQW TTVIERTFHV DSPDREEMM RAIQMVANSL KQRAPGEDPM DYKCGSPSDS STTEMEAV SKARAKVTMN DFDFYLKLLGK GTFGKVILVR EKATGRYYAM KILRKEVIIIA KDEVAHTVTE SRVLQNTRHQ FLTALKYAFQ THDRLCFVME YANGGELFFH LSRERVFTEE RARFYGAEV SALEYLHSRD VVYRDIKLEN LMLDKDGHIK ITDFGLCKEG ISDGATMKTG CGTPEYLAPE VLEDNDYGRA VDWNGLGVM YEMMCGRLPF YNQDHERLFE LILMEEIRFP RTLSPAKSL LAGLLKKDPK QRLLGGPSDA KEVMEHRFFL SINWQDVVK KLPPFKPKQV TSEVDTRYFD DEFTAQSITI TPPDRYDSLQ LLELDQRTHF PQFSYSASIR E*
Mutated AA sequence	MNEVSVIKEG WLHKRGEYIK TWRPRYFLLK SDGSFIGYKE RPEAPDQTLPLNNFSVAEC QLMKTERRP NTFVIRCLQW TTVIERTFHV DSPDREEMM RAIQMVANSL KQRAPGEDPM DYKCGSPSDS STTEMEAV SKARAKVTMN DFDFYLKLLGK GTFGKVILVR EKATGRYYAM KILRKEVIIIA KDEVAHTVTE SRVLQNTRHQ FLTALKYAFQ THDRLCFVME YANGGELFFH LSRERVFTEE RARFYGAEV SALEYLHSRD VVYD IKLEN LMLDKDGHIK ITDFGLCKEG ISDGATMKTG CGTPEYLAPE VLEDNDYGRA VDWNGLGVM YEMMCGRLPF YNQDHERLFE LILMEEIRFP RTLSPAKSL LAGLLKKDPK QRLLGGPSDA KEVMEHRFFL SINWQDVVK KLPPFKPKQV TSEVDTRYFD DEFTAQSITI TPPDRYDSLQ LLELDQRTHF PQFSYSASIR

VARIANT c.49G>A (p. Glu17Lys)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 1.000.



SIFT- Sorting Intolerant From Tolerant

Predicted the variant under tolerated

Predict Not Tolerated	Position	Seq Rep	Predict Tolerated
y w v t s r q p n l k i h g f e d c a	1M	0.25 M	
w m f i c l v y r h q p e k a	2N	0.25 t g d S N	
c w f m y i h l v r g	3E	0.27 n p s a k Q T E D	
w y h d n g r f q e k c m p s	4V	0.27 l i T A V	
w	5S	0.27 f y c h i M l v r g n d P Q k e A S T	
h q w d p n e r c g k s y a t m f l	6V	0.31 V I	
h w d n g q r e y p s c f m t a	7I	0.31 l K I V	
c w f d m	8K	0.31 y g v p s n I a t l H e q r K	
y w v t s r q p n m l k i h g f d c a	9E	0.31 E	
w m i f l c v y r h q p t k e a n	10G	0.31 D S G	
h q k n r d g e p c t s a m v i l	11W	0.31 y F W	
d h g n w e c s r k y p q t a f m i	12L	0.31 V L	
13H	0.31 w c d p f i g y M v n e l a s r Q k T H		
y w v t s r q p n m l i h g f e d c a	14K	0.31 K	
c w f m d i y v g p s h n l a t q	15R	0.31 E K R	
y w v t s r q p n m l k i h f e d c a	16G	0.31 G	
w m c f i y l h v r t p k	17E	0.31 n q s a d G E	
18Y	0.31 w m c i p v q l r f t e k a s d H G n Y		
h d n r e k q g c p s t a w m y v	19I	0.31 l F I	
c w d f m i y v g p s h n l a t e q	20K	0.31 R K	
w f y m h c l r i e q v g k	21T	0.31 p d a S N T	
y v t s r q p n m l k i h g f e d c a	22W	0.31 W	
c w f d m i v v g p s h n l a t e q	23R	0.31 K R	

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.

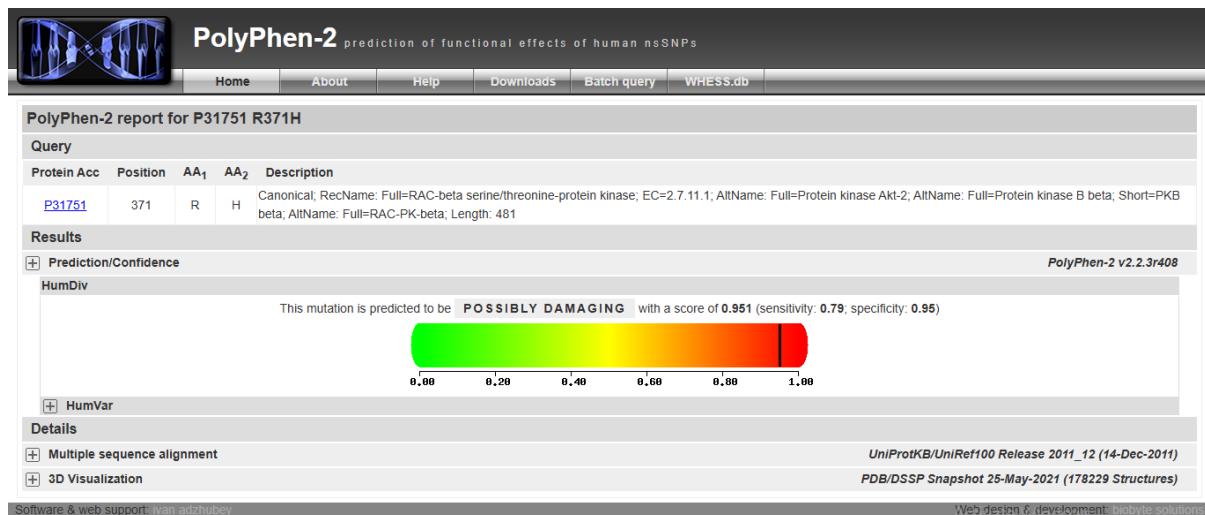
Mutation Taster

mutation t@sting

Prediction:	Deleterious	Permalink			
Summary:	<ul style="list-style-type: none"> Amino acid sequence changed Known disease mutation at this position (HGMD CM118819) Known disease mutation: ClinVar ID 29804 (pathogenic) Protein features (might be) affected Splice site changes 	<ul style="list-style-type: none"> Model: simple_aae Tree vote: 98 2 (del benign) Automatic classification due to ClinVar 			
Analysed issue	Analysis result				
Phys. location	chr19:40762959C>T show variant in all transcripts IGV				
Gene symbol	AKT2				
ExAC LOF metrics	LOF: 1.00, missense: 3.48, synonymous: 0.56				
Ensembl transcript ID	ENST00000392038_2				
Genbank transcript ID	NM_001626 (exact from MANE), NM_001243027 (by similarity), NM_001243028 (by similarity)				
UniProt peptide	P31751				
Variant type	Single base exchange				
Gene region	CDS				
DNA changes	c.49G>A g.28485G>A				
Activating mutations					
Protein conservation	Species	Match	Gene	AA	Alignment
	Human	all conserved		17	VIKEGWLHKRG EYIKTWRPRYFL
	mutated	all conserved		17	VIKEGWLHKRG EYIKTWRPRYFL
	Ptroglodytes	no alignment	ENSPTRG00000010992	n/a	
	Mmulatta	all identical	ENSMUUG00000003621	17	VIKEGWLHKRG EYIKTWRPRYFL
	Fcatus	all identical	ENSFAG00000012180	17	VIKEGWLHKRG EYIKTWRPRYFL
	Mmusculus	all identical	ENSMUSG00000004056	17	VIKEGWLHKRG EYIKTWRPRYFL
	Ggallus	all identical	ENSGALG00000054458	17	VIKEGWLHKRG EYIKTWRPRYFL
	Trubripes	not conserved	ENSTRUG0000008633	208	YLHSRDVVYRDLKLENLM
	Drerio	all identical	ENSDARG0000011219	17	VVREGWLHKRG EYIKTWRPRYFI
	Dmelanogaster	no homologue			
	Celebensis	no homologue			
	Xtropicalis	all identical	ENSXETG00000016986	13	VIKEGWLQKRG EYIKTWRPRYFL
Sequence details					
Original gDNA sequence snippet	CTTCCTGCCTCATTTCAAGGT GAATAACATCAAGACCTGGAGG				
Altered gDNA sequence snippet	CTTCCTGCCTCATTTCAAGGT AAATAACATCAAGACCTGGAGG				
Original cDNA sequence snippet	GCTGGCTTCCACAAGCGTGG GAATAACATCAAGACCTGGAGG				
Altered cDNA sequence snippet	GCTGGCTTCCACAAGCGTGG AAATAACATCAAGACCTGGAGG				
Wildtype AA sequence	MNEVSVIKEG WLHKRG EYIK TWRPRYFLK SDGSFIGYKE RPEAPDQTL PNNFSVAEC QLMKTERPPR NTFVIRCLQW TTVIERTFH DSPDEREEMW RAIQMVANSL KQRAPGEDPM DYKCGSPSDS STTEEMEVAV SKARAKVTMN DFODYLKLKG GTFGKVILVR EKATGRYYAM KILRKEVIIIA KDEVAHTVTE SRVLQNTRHP FLTALKYAFQ THDRLCFVME YANGGELFFH LSRERVFTEE RARFYGAEV SALEYLHSRD VVYRDIKLEN LMLDKDGHIK ITDFGLCKEG ISDGATHKTF CGTPEYLAPE VLEDDNYGRA VDWINGLGVM YEMMCGRLPF YNQDHERLFE LILMEEIRFP RTLSPPEAKSL LAGLLKKDPK QRLGGGPSDA KEVMEHRFFL SINWQDVVKQ KLLPPFKPKQV TSEVDTRYFD DEFTAQSITI TPPDRYDSLQ LLELDQRTHF PQFSYSASIR E*				
Mutated AA sequence	MNEVSVIKEG WLHKRG KYIK TWRPRYFLK SDGSFIGYKE RPEAPDQTL PNNFSVAEC QLMKTERPPR NTFVIRCLQW TTVIERTFH DSPDEREEMW RAIQMVANSL KQRAPGEDPM DYKCGSPSDS STTEEMEVAV SKARAKVTMN DFODYLKLKG GTFGKVILVR EKATGRYYAM KILRKEVIIIA KDEVAHTVTE SRVLQNTRHP FLTALKYAFQ THDRLCFVME YANGGELFFH LSRERVFTEE RARFYGAEV SALEYLHSRD VVYRDIKLEN LMLDKDGHIK ITDFGLCKEG ISDGATHKTF CGTPEYLAPE VLEDDNYGRA VDWINGLGVM YEMMCGRLPF YNQDHERLFE LILMEEIRFP RTLSPPEAKSL LAGLLKKDPK QRLGGGPSDA KEVMEHRFFL SINWQDVVKQ KLLPPFKPKQV TSEVDTRYFD DEFTAQSITI TPPDRYDSLQ LLELDQRTHF PQFSYSASIR E*				

VARIANT c.1112G>A (p. Arg371His)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 0.951.



SIFT- Sorting Intolerant From Tolerant

Predicted the variant under tolerated

w	361L	1.00	f	m	y	c	h	i	v	q	p	d	L	e	N	r	g	K	t	S	A																
y	w	v	t	s	r	q	p	n	m	l	k	h	g	f	e	d	c	a	362I	1.00	I																
d	h	g	n	e	c	w	s	y	r	k	p	q	a	f	v	i	363L	1.00	T	ML																	
w	364M	1.00	c	f	y	M	H	i	p	V	l	g	N	r	q	T	D	S	A	k	E																
wf	my	i	v	l	365E	1.00	p	r	C	t	a	s	n	G	Q	K	H	D	E																		
wy	c	f	m	366E	1.00	i	H	l	g	V	r	P	q	s	a	T	K	N	D	E																	
wh	g	d	n	y	r	q	e	s	k	c	f	m	367I	1.00	a	T	P	L	I	V																	
wc	f	368R	1.00	y	m	h	i	v	p	g	n	L	d	T	Q	A	S	e	R	K																	
nk	h	r	q	d	g	e	p	c	t	s	a	m	v	i	w	369F	1.00	L	Y	F																	
mc	wf	y	i	h	v	n	l	t	q	d	e	s	r	a	g	K	370P	1.00	P																		
wc	371R	1.00	m	f	y	P	g	I	d	H	l	V	N	t	a	S	q	e	R	K																	
dh	g	n	e	c	s	w	y	r	k	p	q	t	a	f	v	372T	1.00	c	W	m	p	d	q	e	g	N	K	i	R	v	T	S	A	H	l	F	j
wh	y	f	m	i	r	c	q	e	d	k	v	n	L	374S	1.00	I	ML																				
wc	f	v	m	375P	1.00	a	G	P	T	S																											
wc	f	v	m	375P	1.00	i	h	v	l	g	t	a	N	O	D	e	S	R	P	K																	

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.



mutation t@sting

Prediction:

Deleterious

[Permalink](#)

Summary:

- Amino acid sequence changed
- Known as potential disease variant: ClinVar ID 376038 (likely pathogenic)
- Protein features (might be) affected
- Model: simple_aae
- Tree vote: 67|33 (del | benign)
- Automatic classification due to ClinVar

Analysed issue

Analysis result

Phys. location	chr19:40741860C>T show variant in all transcripts IGV
Gene symbol	AKT2
ExAC LOF metrics	LOF: 1.00, missense: 3.48, synonymous: 0.56
Ensembl transcript ID	ENST00000392038.2
Genbank transcript ID	NM_001626 (exact from MANE), NM_001243027 (by similarity), NM_001243028 (by similarity)
UniProt peptide	P31751
Variant type	Single base exchange
Gene region	CDS
DNA changes	c.111G>A g.49584G>A

Protein conservation

Species Match Gene AA Alignment

Human	not conserved	371	LILMEIERFPRTLSPEAKSLLA
Pygmytides	all identical	371	LHEEERFPRTLSPEAKSLLA
Mmulatta	all identical	371	LILMEIERFPRTLSPEAKSLLA
Fcatus	all identical	371	LILMEIERFPRTLSPEAKSLLA
Mmusculus	all identical	371	LILMEIERFPRTLSPEAKSLLA
Ggalus	all identical	372	LILMEIERFPRTLSPEAKALLA
Tribripes	all identical	308	LILMEIERFPRTLSPEAKSLLA
Dreno	all identical	369	LILMEIERFPRTLSPEAKALLA
Dmelanogaster	no homologue		
Celegans	no homologue		
Xtrropicalis	all identical	328	LILMEIERFPRTLSPEAKSLLA

.....

Original gDNA sequence snippet

GGAAAGAGATCCGCTTCCGCCACCGCTCACGCCAGCCCCGAGGCCA

Altered gDNA sequence snippet

GGAAAGAGATCCGCTTCCGCCACCGCTCACGCCAGCCCCGAGGCCA

Original cDNA sequence snippet

GGAAAGAGATCCGCTTCCGCCACCGCTCACGCCAGCCCCGAGGCCA

Altered cDNA sequence snippet

GGAAAGAGATCCGCTTCCGCCACCGCTCACGCCAGCCCCGAGGCCA

Wildtype AA sequence

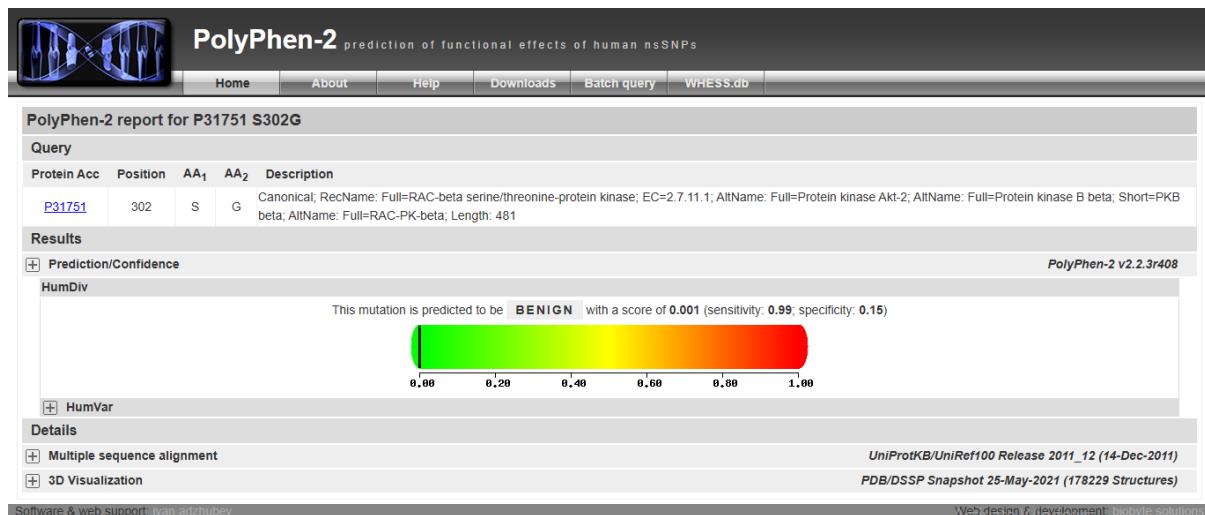
MNEVSIVIKEG WLHKRGEYIK TIRPRYFLLK SDSFSFYKE RPEAPDQTLP PLNNFSVAEC QLWKTTERPP NTFVVERLQLW TTIVERTFHV DSDPEREEWM RAJQWVANSL KQRAPGEDPM DYKCGSPDS STTEEMEVAV SKARAKVTHN PDYLYKLKG GTFGKVILVR EKATGRYAM KILRKVEVIA KDEVAHVHTV SRLVQNTTRP FTALKYAFQ THDRLCFVME YANGELFFH LSRERVFTE RARFYGAEIV SALEYHSRD VVYRDXKLEN LMLDIOGHIX ITDFGLCKEG ISDGATMKTIF CGTPYEYLPE VLENDYGRV VDINGLGVMV YEMMGRLPF YNQDHERLFE LIUHMEIRFP RTLSPEAKSL LAGLKKDKPQ QRLGGPSDA KEVHEHRFLF S1NWQDVQVK KLIPPKFKPVQ TSEVOTRYFD DEFTAQSIITI TPPDRYOSLG LLELDQRTHF PQFSYSASIR E*

Mutated AA sequence

MNEVSIVIKEG WLHKRGEYIK TIRPRYFLLK SDSFSFYKE RPEAPDQTLP PLNNFSVAEC QLWKTTERPP NTFVVERLQLW TTIVERTFHV DSDPEREEWM RAJQWVANSL KQRAPGEDPM DYKCGSPDS STTEEMEVAV SKARAKVTHN PDYLYKLKG GTFGKVILVR EKATGRYAM KILRKVEVIA KDEVAHVHTV SRLVQNTTRP FTALKYAFQ THDRLCFVME YANGELFFH LSRERVFTE RARFYGAEIV SALEYHSRD VVYRDXKLEN LMLDIOGHIX ITDFGLCKEG ISDGATMKTIF CGTPYEYLPE VLENDYGRV VDINGLGVMV YEMMGRLPF YNQDHERLFE LIUHMEIRFP RTLSPEAKSL LAGLKKDKPQ QRLGGPSDA KEVHEHRFLF S1NWQDVQVK KLIPPKFKPVQ TSEVOTRYFD DEFTAQSIITI TPPDRYOSLG LLELDQRTHF PQFSYSASIR E*

VARIANT c.904A>G (p. Ser302Gly)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 0.001.



SIFT- Sorting Intolerant From Tolerant

Predicted the variant under tolerated

Predict Not Tolerated	Position	Seq Rep	Predict Tolerated
h d w n e p q c r g s k y t a f	301I	1.00	V L M I
	302S	1.00	c W M p d q N g i r E K T S V h a L F v
m w i f v l c y r p q h k e s	303D	1.00	w c m i q P r K E T v a H l s G F N D Y
w y	304G	1.00	A N T D G
w c f y m d h g p i e n v q a r s	305A	1.00	f c M h I p l V G N R T q D S A K e
w h y c e r f q d k g p n i v l a s	306T	1.00	L K T
w c f y m i h	307M	1.00	M T
y w v s r q p n m l k i h g f e d c a	308K	1.00	v p l d n Q e A G T R S K
y w v t s r q p n m l k i h g f e d c a	309T	1.00	T
y w v t s r q p n m l k i h g f e d c a	310F	1.00	F
y w v t s r q p n m l k i h g f e d c a	311C	1.00	C
y w v t s r q p n m l k i h g f e d c a	312G	1.00	G
y w v s r q p n m l k i h g f e d c a	313T	1.00	T
y w v t s r q p n m l k i h g f e d c a	314P	1.00	P
c w m f i y l v h r g t n s p a k q	315E	1.00	D E
k q h n r d g e p c t s a m v i w l	316Y	1.00	F Y
d h g n e c s w r k y p q t a f v	317L	1.00	M I L
y w v t s r q p n m l k i h g f e d c	318A	1.00	A
y w v t s r q p n m l k i h g f e d c a	319P	1.00	P
w m c f i y h v l r t n p s k q a d G	320E	1.00	E

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.



mutation t@sting

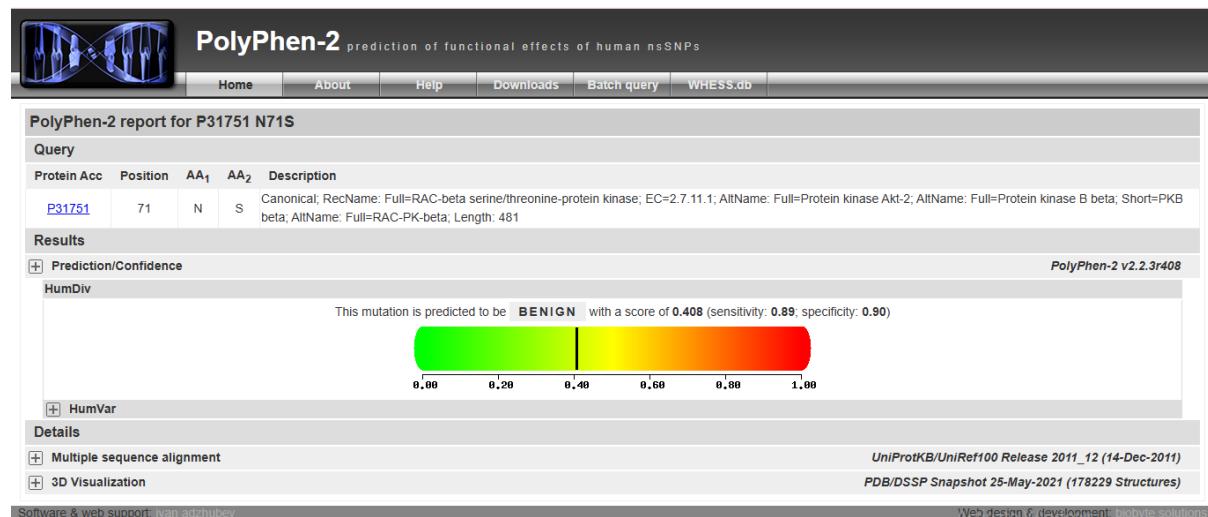
Prediction:	Deleterious	Permalink
Summary:	<ul style="list-style-type: none"> Amino acid sequence changed Known as potential disease variant: ClinVar ID 376039 (likely pathogenic) Protein features (might be) affected 	<ul style="list-style-type: none"> Model: simple_aae Tree vote: 36 64 (del benign) Automatic classification due to ClinVar, tree vote conflicts with the automatic classification
Analysed issue	Analysis result	
Phys. location	chr19:40742220T>C show variant in all transcripts IGV	
Gene symbol	AKT2	
ExAC LOF metrics	LOF: 1.00, missense: 3.48, synonymous: 0.56	
Ensembl transcript ID	ENST00000392038.2	
Genbank transcript ID	NM_001626 (exact from MANE) , NM_001243027 (by similarity) , NM_001243028 (by similarity)	
UniProt peptide	P31751	
Variant type	Single base exchange	
Gene region	CDS	
DNA changes	c.904A>G g.49224A>G	

Activ

Protein conservation	Species	Match	Gene	AA	Alignment
	Human	not conserved		302	ITDFGLCKEGISDGATMKTFCGT
	mutated			302	IGDGATMKTFCGT
	Ptroglodytes	all identical	ENSPTRG00000010992	240	ITDFGLCKEGISDGATMKTFCGT
	Mmulatta	all identical	ENSMMUG00000003621	302	ISDGATMKTFCGT
	Fcatus	all identical	ENSFCAG00000012180	302	ISDGATMKTFCGT
	Mmusculus	all identical	ENSMUSG0000004056	302	ISDGATMKTFCGT
	Ggallus	all identical	ENSGALG00000054458	303	GISDGATMKTFCGT
	Trubripes	all conserved	ENSTRUG00000008633	239	IPDATMKTFCGT
	Drerio	all conserved	ENSDARG00000011219	300	INNEATMKTFCGT
	Dmelanogaster	no homologue			
	Celegans	no homologue			
	Xtrropicalis	not conserved	ENSXETG00000016986	278	-----
Original gDNA sequence snippet	GCCTCTGCAAAGAGGGCATC	AGT GACGGGCCACCATGAAA			
Altered gDNA sequence snippet	GCCTCTGCAAAGAGGGCATC	GGT GACGGGCCACCATGAAA			
Original cDNA sequence snippet	GCCTCTGCAAAGAGGGCATC	AGT GACGGGCCACCATGAAA			
Altered cDNA sequence snippet	GCCTCTGCAAAGAGGGCATC	GGT GACGGGCCACCATGAAA			
Wildtype AA sequence	MNEVSVIKEG WLHKRGEYIK TWRPRYFLLK SDGSFIGYKE RPEAPDQTL PLNNSVAEC QLMKTERPRP NTFVIRCLW TTVIERTFH DSPDEREEWM RAIQMVANSL KQRAPGEDPM DYKCGSPSDS STTEEMEVAV SKARAKVTMN DFODYLKLKG GTFGKVILVR EKATGRYYAM KILRKVEIIIA KDEVAHTVTE SRVLQNTRHP FLTALKYAFQ THDRLCFVME YANGGELFFH LSRSERVFTEE RARFYGAEV SALEYLHSRD VVYRDIKLEN LMLDKDGHIK ITDFGLCKEG ISDGATMKTFCGT PEYLAPE VLEDNDYGRA VDWNGLGVM YEMMCGRLPF YNQDHERLFE LILMEEIRFP RTLSPPEAKSL LAGLLKKDPK QRLGGGPSDA KEVMEHRFFL SINWQDVVK KLPPPFKPQV TSEVDTRYFD DEFTAQSITI TPPDRYDSLQ LLELDQRTHF PQFSYSASIR E*				
Mutated AA sequence	MNEVSVIKEG WLHKRGEYIK TWRPRYFLLK SDGSFIGYKE RPEAPDQTL PLNNSVAEC QLMKTERPRP NTFVIRCLW TTVIERTFH DSPDEREEWM RAIQMVANSL KQRAPGEDPM DYKCGSPSDS STTEEMEVAV SKARAKVTMN DFODYLKLKG GTFGKVILVR EKATGRYYAM KILRKVEIIIA KDEVAHTVTE SRVLQNTRHP FLTALKYAFQ THDRLCFVME YANGGELFFH LSRSERVFTEE RARFYGAEV SALEYLHSRD VVYRDIKLEN LMLDKDGHIK ITDFGLCKEG ISDGATMKTFCGT PEYLAPE VLEDNDYGRA VDWNGLGVM YEMMCGRLPF YNQDHERLFE LILMEEIRFP RTLSPPEAKSL LAGLLKKDPK QRLGGGPSDA KEVMEHRFFL SINWQDVVK KLPPPFKPQV TSEVDTRYFD DEFTAQSITI TPPDRYDSLQ LLELDQRTHF PQFSYSASIR E*				

VARIANT c.212A>G (p. Asn71Ser)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 0.408.



SIFT- Sorting Intolerant From Tolerated

Predicted the variant under tolerated

n	a	w	g	n	q	r	e	s	c	y	p	k	t	r	a	65M	0.31	i	v	1	M															
	w	64K	0.31	y	d	f	h	m	g	n	C	p	e	q	i	s	a	L	r	v	T	K														
		65T	0.31	w	d	c	g	e	q	p	r	k	h	m	N	s	y	F	a	i	T	V	L													
c	w	m	f	i	y	l	v	h	r	t	g	p	n	a	k	q	66E	0.31	S	D	E															
c	w	d	f	m	y	v	g	p	s	h	n	l	a	t	e	q	67R	0.31	K	R																
w	c	f	m	y	h	v	l	d	n	t	q	s	e	g	a	r	68P	0.31	K	P																
c	w	d	f	m	y	v	g	p	s	h	n	l	a	t	e	q	69R	0.31	K	R																
																	70P	0.31	t	s	g	i	l	A	V	P										
																			g	d	F	N														
																			i	c	q	r	e	d	l	k	v	n	p	M	G	a	s	T		
y	w	v	t	s	r	q	p	n	m	l	k	i	h	g	e	d	c	73F	0.31	F																
																		74V	0.31	t	a	M	E	I	V	I										
h	q	w	d	p	n	e	r	c	g	k	s	y	a	t	m	f	l	75I	0.31	V	I															
c	w	f	d	m	i	y	v	p	g	h	s	l	n	t	a	e	k	76R	0.31	Q	R															
																		77C	0.29	s	a	G	C													
y	w	v	t	s	r	q	p	n	m	k	i	h	g	f	e	d	c	78L	0.29	L																
y	w	v	t	s	r	p	n	m	l	k	i	h	g	f	e	d	c	79Q	0.29	Q																
y	v	t	s	r	q	p	n	m	l	k	i	h	g	f	e	d	c	80W	0.29	W																

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.



mutation t@sting

Prediction:	Deleterious	Permalink
Summary:	<ul style="list-style-type: none"> Amino acid sequence changed Protein features (might be) affected 	<ul style="list-style-type: none"> Model: simple_aae Tree vote: 67 33 (del benign)
Analysed issue		
Phys. location	chr19:40761140T>C	show variant in all transcripts IGV
Gene symbol	AKT2	
ExAC LOF metrics	LOF: 1.00, missense: 3.48, synonymous: 0.56	
Ensembl transcript ID	ENST00000392038.2	
Genbank transcript ID	NM_001626 (exact from MANE) , NM_001243027 (by similarity) , NM_001243028 (by similarity)	
UniProt peptide	P31751	
Variant type	Single base exchange	
Gene region	CDS	
DNA changes	c.212A>G g.30304A>G	

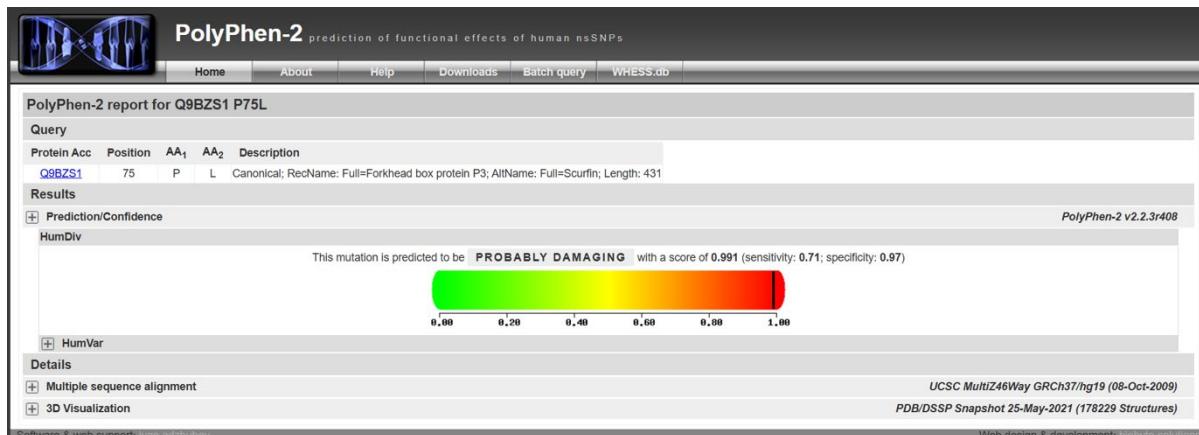
Protein conservation	Species	Match	Gene	AA Alignment
Human	Human	all conserved		71 CQLMKTERPRPNTFVIRCLQWTT
mutated				71 QLMKTERPRPNTFVIRCLQWTT
Ptroglobutes	all identical		ENSPTRG0000010992	9 MKTERPRPNTFVIRCLQWTT
Mmulatta	all identical		ENSMUG000000003621	71 QLMKTERPRPNTFVIRCLQWTT
Fcatus	all identical		ENSFCAG00000012180	71 QLMKTERPRPNTFVIRCLQWTT
Mmusculus	all identical		ENSMUSG0000004056	71 QLMKTERPRPNTFVIRCLQWTT
Ggallus	all identical		ENSGAL000000054458	71 QLMKTERPRPNTFVIRCLQWTT
Trubripes	all identical		ENSTRUG00000008633	9 MKTERPKTNTFVQLQWTT
Dreario	all identical		ENSDARG00000011219	71 QLMKTERPRPNTFVIRCLQWTT
Dmelanogaster	no homologue			
Celegans	no homologue			
Xtropicalis	all identical		ENSXETG00000016988	68 TERPRPNTFVIRCLQWTT

Original gDNA sequence snippet	GACCGAGAGGCCGCAGCCA A CACCTTGTCATACGCTGCC
Altered gDNA sequence snippet	GACCGAGAGGCCGCAG C ACCTTGTCATACGCTGCC
Original cDNA sequence snippet	GACCGAGAGGCCGCAGCCA A CACCTTGTCATACGCTGCC
Altered cDNA sequence snippet	GACCGAGAGGCCGCAG C ACCTTGTCATACGCTGCC
Wildtype AA sequence	MNEVSVIKEG WLHKRGEYIK TWRPRYFLK SDGSFIGYKE RPEAPDQTL P LNNFSVAEC QLMKTERPRP NTFVIRCLQW TTVIERTFHV DSDPDEREEM RAIQMVANSL KQRAPGEDPM DYKCGSPSDS STTEEMEVAV SKARAKVTMM DF DYKL LGK GTFGKVILVR EKATGRYYAM KILRK EVIIIA KDEVAHTVTE SRVLQNTRHP F LTALKYAFQ THDRLCFVME YANGGELFFH LSRSRVFTEE RARFYGAEV SALEYLHSRD VVYRDIKLEN LMLDKDGHIK ITDFGLCKEG ISDGATMKTG CGTPEYLAPE VLEDNDYGRA VDWNGLGVM YEMMCGRLPF YNQDHERLFE LILMEEIRFP RTLSPAKSL LAGLLKKDPK QR LGGGPSDA KEVMEHRFFL SINWQDVVQK KLLPPFKPQV TSEVDTRYFD DEFTAQSITI TPPDRYDSL G LLELDQRTHF PQFSYSASIR E*
Mutated AA sequence	MNEVSVIKEG WLHKRGEYIK TWRPRYFLK SDGSFIGYKE RPEAPDQTL P LNNFSVAEC QLMKTERPRP S TFVIRCLQW TTVIERTFHV DSDPDEREEM RAIQMVANSL KQRAPGEDPM DYKCGSPSDS STTEEMEVAV SKARAKVTMM DF DYKL LGK GTFGKVILVR EKATGRYYAM KILRK EVIIIA KDEVAHTVTE SRVLQNTRHP F LTALKYAFQ THDRLCFVME YANGGELFFH LSRSRVFTEE RARFYGAEV SALEYLHSRD VVYRDIKLEN LMLDKDGHIK ITDFGLCKEG ISDGATMKTG CGTPEYLAPE VLEDNDYGRA VDWNGLGVM YEMMCGRLPF YNQDHERLFE LILMEEIRFP RTLSPAKSL LAGLLKKDPK QR LGGGPSDA KEVMEHRFFL SINWQDVVQK KLLPPFKPQV TSEVDTRYFD DEFTAQSITI TPPDRYDSL G LLELDQRTHF PQFSYSASIR E*

FOXP3

VARIANT c.224C>T (p. Pro75Leu)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 0.991.



SIFT- Sorting Intolerant From Tolerant

Predicted the variant under tolerated

ywvtsrqnmklkihgfedca	65P	0.01	P
ywvtsrqnmklkihgfedca	66P	0.01	P
ywvtrqpnmlkihgfedca	67S	0.01	S
ywvtsrpnmlkihgfedca	68Q	0.01	Q
ywvtsrqpnmkihgfedca	69L	0.01	L
ywvtsrpnmlkihgfedca	70Q	0.01	Q
ywvtsrqpnmkihgfedca	71L	0.01	L
ywvtsrqnmklkihgfedca	72P	0.01	P
ywvsrqpnmlkihgfedca	73T	0.01	T
dhgnwercskyqptafmi	74I	0.01	VL
ywvtsrqnmklkihgfedca	75P	0.01	P
ywvtsrqpnmkihgfedca	76L	0.01	L
ywtsrqpnmlkihgfedca	77V	0.01	V
ywvtsrqpnmlkihgfedca	78M	0.01	M
ywtsrqpnmlkihgfedca	79V	0.01	V
ywvtsrqpnmlkihgfedc	80A	0.01	A
ywvtsrqnmklkihgfedca	81P	0.01	P
ywvtrqpnmlkihgfedca	82S	0.01	S
ywvtsrqpnmlkihgfedca	83G	0.01	G
ywvtsrqpnmlkihgfedc	84A	0.01	A
ywvtsrqpnmlkihgfedca	85R	0.01	R

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a benign variant.



mutation t@sting

Prediction: **Benign** [Permalink](#)

Summary: • Amino acid sequence changed
• Protein features (might be) affected

- Model: simple_aae
- Tree vote: 18/82 (del | benign)

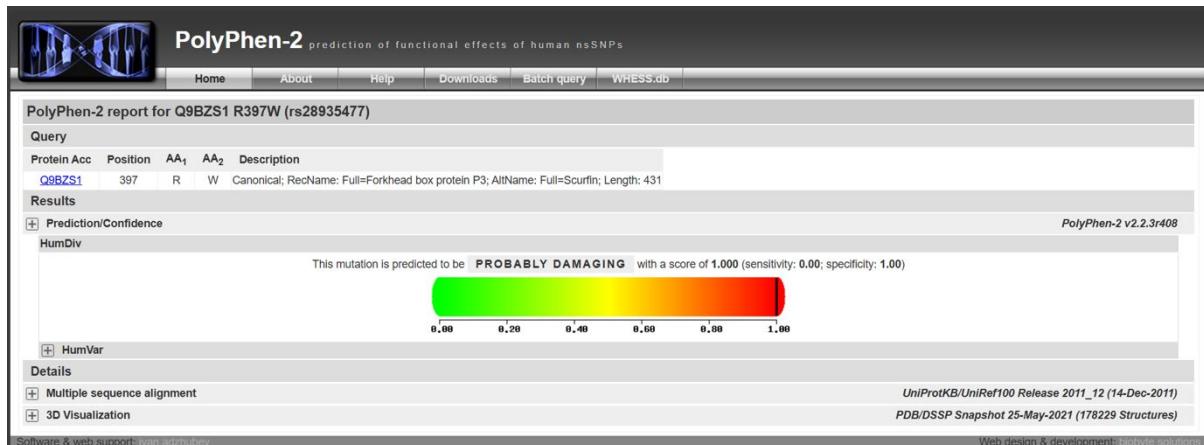
Analysed issue	Analysis result
Phys. location	chr23:49114212G>A show variant in all transcripts IGV
Gene symbol	FOXP3
ExAC LOF metrics	LOF: 0.95, missense: 0.86, synonymous: 0.11
Ensembl transcript ID	ENST00000376207.4
Genbank transcript ID	NM_014009 (exact from MANE)
UniProt peptide	Q9BZS1
Variant type	Single base exchange
Gene region	CDS
DNA changes	c.224C>T g.7077C>T

Protein conservation	Species	Match	Gene	AA	Alignment
	Human	not conserved		75	HPPSQLQLPTPLLVNVAPSGARLG
	Pygmytarsier	all identical	ENSPTRG00000021890	75	HPPSQLQLPTL L VNVAPSGARL
	Mmullata	all identical	ENSMML000000008624	183	HPPSQLQLPTL L VNVAPSGARL
	Fcatus	all identical	ENSFCA00000003823	74	HPPSQLQLPTL L VNVAPSGARL
	Mmusculus	no alignment	ENSMUSG000000039521	n/a	
	Ggaalus	no homologue			
	Tribripes	no homologue			
	Dreino	no homologue			
	Dmelanogaster	no homologue			
	Celegans	no homologue			
	Xtropicalis	all identical	ENSXETG000000031498	64	QLOQWGVIVVSPS-ASF

Original gDNA sequence snippet	TGCCAGCTGCCAACACTGCC C CCTAGTCATGGTGGCACCCCT
Altered gDNA sequence snippet	TGCCAGCTGCCAACACTGC T CCTAGTCATGGTGGCACCCCT
Original cDNA sequence snippet	GCTGCAGCTGCCAACACTGC C CCTAGTCATGGTGGCACCCCT
Altered cDNA sequence snippet	GCTGCAGCTGCCAACACTGC T CCTAGTCATGGTGGCACCCCT
Wildtype AA sequence	MPNPRPGKPS AP\$ALALGPSP GASPSWRAAP KASDLLGARG PGGTFOQGRDL RGGAHASSSS LNPMPPSQLQ LPTL L VMVA PSGARLGPLP HLQALLQDRP HFMHQQLSTVD AHARTPVVLQV HPLESPAMIS LTPPTTATGV FSLKARPGLP PGINVASLEW VSREPALLCT FPNPSAPRKD STLSAVPQSS YPLLNGVCK WPGCEKVFEF PEDFLKHQCQ DHLLDEKGRA QCLLQREMVQ SLEQQVLKEK EKLSAMQAHL AGKMALTAKAS SVASSDKGS CIVAAGSQGP VVPAWSGP APDSLFAVRN HLWGSHGNST FPEFLHNMDY FKFHNMRRPFF TYATLIRWAI LEAPEKQRTL NEIYHNFTRM FAFFRNHPAT WKNAIRHNL LHKCFVRVES EKGAVWTVDE LEFRKKRSQR PSRCSNPTPG P*
Mutated AA sequence	MPNPRPGKPS AP\$ALALGPSP GASPSWRAAP KASDLLGARG PGGTFOQGRDL RGGAHASSSS LNPMPPSQLQ LPTL L VMVA PSGARLGPLP HLQALLQDRP HFMHQQLSTVD AHARTPVVLQV HPLESPAMIS LTPPTTATGV FSLKARPGLP PGINVASLEW VSREPALLCT FPNPSAPRKD STLSAVPQSS YPLLNGVCK WPGCEKVFEF PEDFLKHQCQ DHLLDEKGRA QCLLQREMVQ SLEQQVLKEK EKLSAMQAHL AGKMALTAKAS SVASSDKGS CIVAAGSQGP VVPAWSGP APDSLFAVRN HLWGSHGNST FPEFLHNMDY FKFHNMRRPFF TYATLIRWAI LEAPEKQRTL NEIYHNFTRM FAFFRNHPAT WKNAIRHNL LHKCFVRVES EKGAVWTVDE LEFRKKRSQR PSRCSNPTPG P*

VARIANT c.1189C>T (p. Arg397Trp)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 1.000.



SIFT- Sorting Intolerant From Tolerant

Predicted the variant under tolerated

miwvfl	cyrpqahatkegSD	383N	1.00	N
whyfmirqdeclnkvpgT		384A	1.00	SA
hqpwdencrksgyatmfl		385I	1.00	VI
ywvtsqpnmklkihgfedca		386R	1.00	R
ywvtsrqpnmlkigfedca		387H	1.00	H
mwifvclyrqhptkeagds		388N	1.00	N
ywvtsrqpnmkihgfedca		389L	1.00	L
whyfmirqelkdvnpgatC		390S	1.00	S
dhnegcrkpqtwayvmisF		391L	1.00	L
wmicfvpatqekgLdYRS		392H	1.00	NH
wcmfivylhpqgANTRES		393K	1.00	DK
wfhidnleqvtpgMRYKSA		394C	1.00	C
ywvtsrqpnmlkihgedca		395F	1.00	F
wdcfygspnaHTQIFR		396V	0.99	MKIV
cwfcdmiyvsgphnalteq		397R	0.99	KR
hwgdnqyrskskpcfaMETL		398V	0.99	IV
wycfmivgnHTRSKAQLD		399E	0.98	PE
cwfmiyvdhlgtAe		400S	0.98	QKPRNS

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.



mutation t@sting

Prediction:

Deleterious

[Permalink](#)

Summary:

- Amino acid sequence changed
- Known disease mutation at this position (HGMD CM010059)
- Known disease mutation: ClinVar ID 11407 (pathogenic)
- Protein features (might be) affected
- Model: simple_aae
- Tree vote: 77|23 (del | benign)
- Automatic classification due to ClinVar

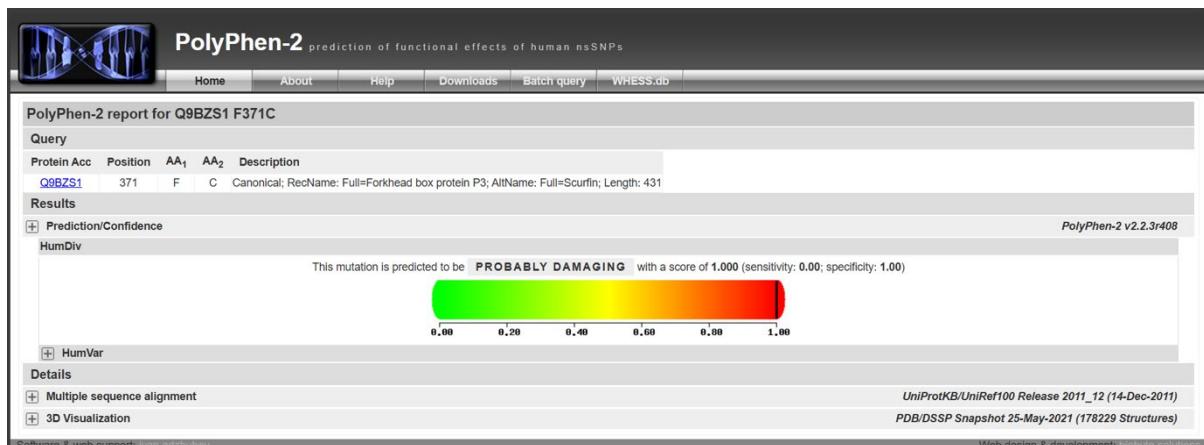
Analysed issue	Analysis result
Phys. location	chr23:49107902G>A show variant in all transcripts IGV
Gene symbol	FOXP3
ExAC LOF metrics	LOF: 0.95, missense: 0.86, synonymous: 0.11
Ensembl transcript ID	ENST00000376207.4
Genbank transcript ID	NM_014009 (exact from MANE)
UniProt peptide	Q9BZS1
Variant type	Single base exchange
Gene region	CDS
DNA changes	c.1189C>T g.13387C>T

Protein conservation	Species	Match	Gene	AA Alignment
Human	not conserved			397 RHNL S HKKCFVRESEKGA W TVD
Ptroglobutes	all identical	ENSPTRG00000021890		397 RHNL S HKKCFV W ESEKGA W TV
Mmulatta	all identical	ENSMUG00000008624		505 RHNL S HKKCFV W ESEKGA W TV
Fcatus	all identical		ENSFCAG00000003823	369 RHNL S HKKCFV W ESEKGA W TV
Mmusculus	no alignment		ENSMUSG00000039521	n/a
Ggallus	no homologue			
Trubripes	no homologue			
Drerio	no homologue			
Dmelanogaster	no homologue			
Celegans	no homologue			
Xtropicalis	all identical	ENSXETG00000031498		386 RHNL S HKKCFV W ENIKGA W MV

Original gDNA sequence snippet	GTCTGCACAAGTGT T GT G GGGTGGAGAGCGAGAAGGG
Altered gDNA sequence snippet	GTCTGCACAAGTGT T GT G GGGTGGAGAGCGAGAAGGG
Original cDNA sequence snippet	GTCTGCACAAGTGT T GT G GGGTGGAGAGCGAGAAGGG
Altered cDNA sequence snippet	GTCTGCACAAGTGT T GT G GGGTGGAGAGCGAGAAGGG
Wildtype AA sequence	MPNPRPGKPS APSLALGPSP GASPSWRAAP KASDLLGARG PGGTFQGRDL RGGAHASSSS LNPMPPSQLQ LPTLPLVMVA PSGARLGPLP HLQALLQDRP HFMHQLSLTV AHARTPVLVQ HPLESPAMIS LTPPTTATGV FSLKARPGLP PGINVASLEW VSREPALLCT FPNPSAPRKD STLSAVPQSS YPFLANGVCK WPGCEKVFFEE PEDFLKHQA DHLLDEKGRA QCLLQREMVQ SLEQQQLVLER EKLSAMQ AHL AGKMALTAKS SVASSDKGS CIVAAGSQGP VVPAWSGPRE APDSLFAVRR HLWGSHGNST FPEFLHNMDY FKFHNMRRPPF TYATLIRWAI LEAPEKQRTL NEIYHWIFTRM FAFFRNHPAT WKNAIRHNLS LHKCFVRVES EKGAVWTDE LEFRKKRSQR PSRCSNPTPG P*
Mutated AA sequence	MPNPRPGKPS APSLALGPSP GASPSWRAAP KASDLLGARG PGGTFQGRDL RGGAHASSSS LNPMPPSQLQ LPTLPLVMVA PSGARLGPLP HLQALLQDRP HFMHQLSLTV AHARTPVLVQ HPLESPAMIS LTPPTTATGV FSLKARPGLP PGINVASLEW VSREPALLCT FPNPSAPRKD STLSAVPQSS YPFLANGVCK WPGCEKVFFEE PEDFLKHQA DHLLDEKGRA QCLLQREMVQ SLEQQQLVLER EKLSAMQ AHL AGKMALTAKS SVASSDKGS CIVAAGSQGP VVPAWSGPRE APDSLFAVRR HLWGSHGNST FPEFLHNMDY FKFHNMRRPPF TYATLIRWAI LEAPEKQRTL NEIYHWIFTRM FAFFRNHPAT WKNAIRHNLS LHKCFVRVES EKGAVWTDE LEFRKKRSQR PSRCSNPTPG P*

VARIANT c.1112T>G (p. Phe371Cys)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 1.000.



SIFT- Sorting Intolerant From Tolerant

Predicted the variant under tolerated

cwmfiylvrtpskHN	362E	1.00	DAGQE
hqwnekdrscpygtmafIV	363I	1.00	I
dwhegrkpqstafmvNICL	364Y	1.00	Y
wyfcmipvlglT	365H	1.00	DRAEQKSNH
cmpkrirtvsnalDQHGE	366W	1.00	YFW
hdwnnegcrpsqkyatLM	367F	1.00	VIF
wpfngndyHIEKLQCRVA	368T	1.00	MST
wyfcmhipiLVAGNSKT	369R	1.00	QEDR
wCdpegi	370M	1.00	aVQsfKYRHLNTM
dcpenqkgrstamiHlWVY	371F	1.00	F
whyfminqdekRvtLgs	372A	1.00	CPA
kqhnrdrgepcstsamviwl	373F	1.00	FY
kqhnrdrgepcstsamviwl	374F	1.00	YF
cwdfmiyvgptsLEQHAN	375R	1.00	KR
wcmpIV	376N	0.98	IFYSAQDGTHKERN
wmifvlyrpCeQTg	377H	1.00	KDSANH
wcmh1VINeRFG	378P	1.00	QSYDKTAP
wyfch1VGRINTMESKD	379A	1.00	QPA
wfvhmicleadvnDKR	380T	1.00	ASGT

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.

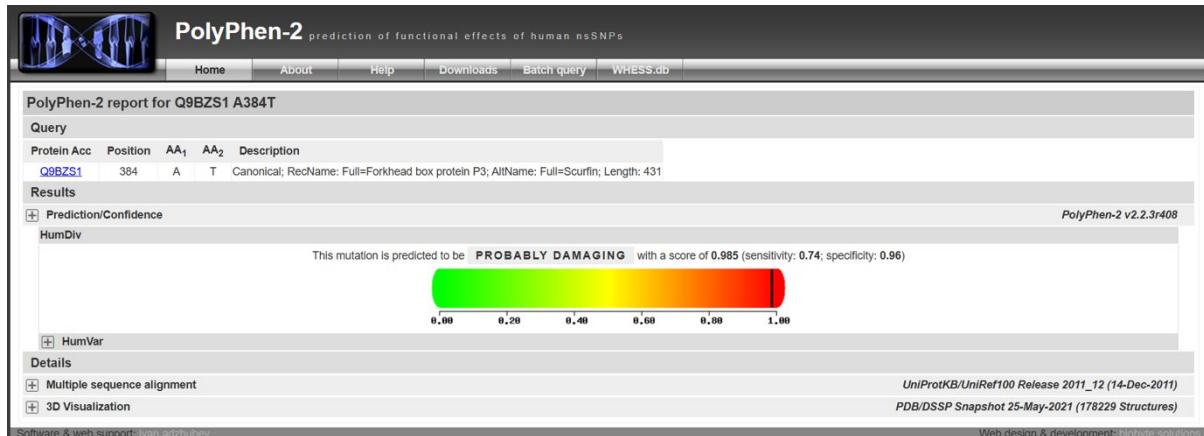


mutation t@sting

Prediction:	Deleterious	Permalink			
Summary:	<ul style="list-style-type: none"> Amino acid sequence changed Known disease mutation at this position (HGMD CM010057) Known disease mutation: ClinVar ID 11409 (pathogenic) Protein features (might be) affected 	<ul style="list-style-type: none"> Model: simple_aae Tree vote: 100 0 (del benign) Automatic classification due to ClinVar 			
Analysed issue					
Phys. location	chr23:49108159A>C show variant in all transcripts IGV				
Gene symbol	FOXP3				
ExAC LOF metrics	LOF: 0.95, missense: 0.86, synonymous: 0.11				
Ensembl transcript ID	ENST00000376207.4				
Genbank transcript ID	NM_014009 (exact from MANE)				
UniProt peptide	Q9BZS1				
Variant type	Single base exchange				
Gene region	CDS				
DNA changes	c.1112T>G g.13130T>G				
Protein conservation					
	Species	Match	Gene	AA	Alignment
	Human			371	LNEIYHWFTRMFAFFRNHPATWK
	mutated	not conserved		371	NEIYHWFTRMCAFFRNHPATWK
	Ptroglobutes	all identical	ENSPTRG00000021890	371	NEIYHWFTRM CA FFRNHPATWK
	Mmulatta	all identical	ENSMUG00000008624	479	NEIYHWFTRM CA FFRNHPATWK
	Fcatus	all identical	ENSFCAG0000003823	343	NEIYHWFTRM CA FFRNHPATWK
	Mmusculus	no alignment	ENSMUSG00000039521	n/a	
	Ggalus	no homologue			
	Trubripes	no homologue			
	Drerio	no homologue			
	Dmelanogaster	no homologue			
	Celegans	no homologue			
	Xtrropicalis	all identical	ENSXETG00000031498	360	LNEIYHWFTRM CA FFRYNTATWK
Original gDNA sequence snippet					
	CCACTGGTTCACACGCATGT T TGCTCTTCAGAACCATC				
Altered gDNA sequence snippet					
	CCACTGGTTCACACGCATGT G TGCTCTTCAGAACCATC				
Original cDNA sequence snippet					
	CCACTGGTTCACACGCATGT T TGCTCTTCAGAACCATC				
Altered cDNA sequence snippet					
	CCACTGGTTCACACGCATGT G TGCTCTTCAGAACCATC				
Wildtype AA sequence					
	MPNPRPGKPS APSALGPSP GASPSWRAAP KASDLLGARG PGGTFQGRDL RGGAHASSSS LNPPMPPSQLQ LPTPLVIVMA PSGARGLPLP HLQALLQDRP HFMHQQLSTVD AHARTPVQLV HPLESPAMIS LTPPTTATGV FSLKARPGLP PGIVVASLW VSREPALLCT FPNPSAPRKD STLSAVPOSS YPLLANGVCK WPGCEKVFFEE PEDFLKHQCQA DHLLDEKGRA QCLLQRENVQ SLEQQQLVLEK EKLSAMQAHL AGKMALTAKS SVASSDKGSC CIVAAGSGQP VVPWAWSGPRE APDSLFAVRR HLHSHGHNST FPEFLHNNDY FKFHNNRPPF TYATLIRWAI LEAPEKQRTL NEIYHWFTRM FAFFRNHPAT WKNAIRHNLs LHKCFVRVES EKGAWITVDE LEFRKKRSQR PSRCSNPTPG P*				
Mutated AA sequence					
	MPNPRPGKPS APSALGPSP GASPSWRAAP KASDLLGARG PGGTFQGRDL RGGAHASSSS LNPPMPPSQLQ LPTPLVIVMA PSGARGLPLP HLQALLQDRP HFMHQQLSTVD AHARTPVQLV HPLESPAMIS LTPPTTATGV FSLKARPGLP PGIVVASLW VSREPALLCT FPNPSAPRKD STLSAVPOSS YPLLANGVCK WPGCEKVFFEE PEDFLKHQCQA DHLLDEKGRA QCLLQRENVQ SLEQQQLVLEK EKLSAMQAHL AGKMALTAKS SVASSDKGSC CIVAAGSGQP VVPWAWSGPRE APDSLFAVRR HLHSHGHNST FPEFLHNNDY FKFHNNRPPF TYATLIRWAI LEAPEKQRTL NEIYHWFTRM CA FFRNHPAT WKNAIRHNLs LHKCFVRVES EKGAWITVDE LEFRKKRSQR PSRCSNPTPG P*				

VARIANT c.1150G>A (p. Ala384Thr)

POLY PHEN 2– Predicted the variant to be a probably damaging kind with a score of 0.985.



SIFT- Sorting Intolerant From Tolerant

Predicted the variant under tolerated

c w d f m i y v g p t S L E Q H A N	375 R	1.00	K R
w c m p I v	376 N	0.98	I F Y S A Q D G T H K E R N
w m i f v l y r p C e Q T g	377 H	1.00	K D S A N H
w c m h l V I N e R F G	378 P	1.00	Q S Y D K T A P
w y f c h l V G R I N T M E S K D	379 A	1.00	Q P A
w f y h m i c l e q d v n p K R	380 T	1.00	A S G T
y v t s r q p n m l k i h g f e d c a	381 W	1.00	W
c w d f m i y v g p s n a l t e H R	382 K	1.00	Q K
m i w v f l c y r p q a h t k e g S D	383 N	1.00	N
w h y f m i r q d e c l n k v p g T	384 A	1.00	S A
h q p w d e n c r k s g y a t m f l	385 I	1.00	V I
y w v t s q p n m l k i h g f e d c a	386 R	1.00	R
y w v t s r q p n m l k i g f e d c a	387 H	1.00	H
m w i f v c l y r q h p t k e a g d S	388 N	1.00	N
y w v t s r q p n m k i h g f e d c a	389 L	1.00	L
w h y f m i r q e l k d v n p g a T C	390 S	1.00	S
d h n e g c r k p q t w a y v m i S F	391 L	1.00	L
w m i c f v p a t q e k g L d Y R S	392 H	1.00	N H
w c m f i v y l h p q g A N T R E S	393 K	1.00	D K
w f h i d n l e q v t p g M R Y K S A	394 C	1.00	C
y w v t s r q p n m l k i h g e d c a	395 F	1.00	F

MUTATION TASTER-This software predicts the pathogenicity of the variant. It predicted the variant as a deleterious variant.



mutation t@sting

Prediction:	Deleterious	Permalink
Summary:		
	<ul style="list-style-type: none"> Amino acid sequence changed Known disease mutation at this position (HGMD CM010058) Known disease mutation: ClinVar ID 11410 (pathogenic) Protein features (might be) affected 	<ul style="list-style-type: none"> Model: simple_aae Tree vote: 99 1 (del benign) Automatic classification due to ClinVar
Analysed issue		
Phys. location	chr23:49107941C>T show variant in all transcripts GOV	
Gene symbol	FOXP3	
ExAC LOF metrics	LOF: 0.95, missense: 0.86, synonymous: 0.11	
Ensembl transcript ID	ENST00000376207_4	
Genbank transcript ID	NM_014009 (exact from MANE)	
UniProt peptide	Q9BZS1	
Variant type	Single base exchange	
Gene region	CDS	
DNA changes	c.1150G>A g.1348G>A	

Protein conservation		Species	Match	Gene	AA	Alignment
Human	mutated	not conserved			384	FFRNHPATWKNAIRHNLSLHKCFV
Ptroglydotes	all identical		ENSPTRG00000021890		384	FFRNHPATWKNTIRHNLSLHKCF
Mmulatta	all identical		ENSMRMUG000000008624		492	FFRNHPATWKNTIRHNLSLHKCF
Fcatus	all identical		ENSFCAG00000003823		356	FFRNHPATWKNTIRHNLSLHKCF
Mmusculus	no alignment		ENSMUSG00000039521		n/a	
Ggallus	no homologue					
Trubripes	no homologue					
Drerio	no homologue					
Dmelanogaster	no homologue					
Celegans	no homologue					
Xtropicalis	all identical		ENSXETG00000031498		373	FFRYNTATWKNTIRHNLSLHKCF

Original gDNA sequence snippet	TCCCCGGCTTCCACAGAAC G CCATCCGCCACAACCTGAGT
Altered gDNA sequence snippet	TCCCCGGCTTCCACAGAAC A CCATCCGCCACAACCTGAGT
Original cDNA sequence snippet	ATCCTGCCACCTGGAGAAC G CCATCCGCCACAACCTGAGT
Altered cDNA sequence snippet	ATCCTGCCACCTGGAGAAC A CCATCCGCCACAACCTGAGT
Wildtype AA sequence	MPNPRPGKPS APSALGPSP GASPSWRAAP KASDLLGARG PGGTQGRDL RGGAHASSSS LNPNMPPSQLQ LPTPLVWVA PSGARLGPLP HLQALLQDRP HFMHQLSLTV AHARTPVLQV HPLESPAMIS LTPPTTATGV FSLKARPGLP PGIVASLEW VSREPALLCT FPNSPAPRKD STLSAVPQSS YPLLANGVCK WPGCEKVFEF PEDFLKHQCQ A DHLLEDEKGRA QCLLQREMVQ SLEQQQLVLEK EKLSAMQAHL AGKMALTAKAS SVASSDKGSC CIVAAGSQGP VVPAWSGPRE APDSDLFAVRR HLNGSHGINST FPEFLHNNDY FKFHNMNRPF TYATLIRWAI LEAPEKQRTL NEIYHWTRM FAFFRNHPAT WKNTIRHNLNS LHKCFVRVES EKGAWVTVD LEFRKKRSQR PSRCSNPTPG P*
Mutated AA sequence	MPNPRPGKPS APSALGPSP GASPSWRAAP KASDLLGARG PGGTQGRDL RGGAHASSSS LNPNMPPSQLQ LPTPLVWVA PSGARLGPLP HLQALLQDRP HFMHQLSLTV AHARTPVLQV HPLESPAMIS LTPPTTATGV FSLKARPGLP PGIVASLEW VSREPALLCT FPNSPAPRKD STLSAVPQSS YPLLANGVCK WPGCEKVFEF PEDFLKHQCQ A DHLLEDEKGRA QCLLQREMVQ SLEQQQLVLEK EKLSAMQAHL AGKMALTAKAS SVASSDKGSC CIVAAGSQGP VVPAWSGPRE APDSDLFAVRR HLNGSHGINST FPEFLHNNDY FKFHNMNRPF TYATLIRWAI LEAPEKQRTL NEIYHWTRM FAFFRNHPAT WKNTIRHNLNS LHKCFVRVES EKGAWVTVD LEFRKKRSQR PSRCSNPTPG P*

GENE	VARIANTS	POLYPHEN 2	SIFT	MUTATIONTATSER
INS	c.147C>G (p.Phe49Leu) c.143T>C (p.Phe48Ser) c.266G>A (p.Arg89His) c.274G>T (p.Val92Leu) c.266G>T (p.Arg89Leu)	Damaging – 0.994 Damaging – 1 Damaging – 1 Damaging – 1 Possibly Damaging – 0.865	Not Tolerated Not Tolerated Not Tolerated Not Tolerated Not Tolerated	Deleterious Deleterious Deleterious Deleterious Deleterious
INSR	c.3104G>T (p.Gly1035Val) c.1459A>G (p.Lys487Glu) c.1177G>A (p.Gly393Arg)	Damaging – 1 Benign – 0.024 Damaging – 1	Not Tolerated Tolerated Not Tolerated	Deleterious Deleterious Deleterious
PPARG	c.362A>G (p.Tyr121Cys) c.1184G>A (p.Arg395His)	Damaging – 1 Damaging – 1	Not Tolerated Not Tolerated	Deleterious Not Found
AKT2	c.821G>A (p.Arg274His) c.49G>A (p.Glu17Lys) c.1112G>A (p.Arg371His) c.904A>G(p.Ser302Gly) c.212A>G (p.Asn71Ser)	Damaging – 0.999 Damaging – 1 Damaging – 0.951 Benign –0.001 Benign –0.408	Not Tolerated Not Tolerated Tolerated Tolerated Not Tolerated	Deleterious Deleterious Deleterious Deleterious Deleterious
FOXP3	c.224C>T (p.Pro75Leu) c.1189C>T (p.Arg397Trp) c.1112T>G (p.Phe371Cys) c.1150G>A (p.Ala384Thr)	Damaging – 0.991 Damaging – 1 Damaging – 1 Damaging – 0.985	Not Tolerated Not Tolerated Not Tolerated Not Tolerated	Benign Deleterious Deleterious Deleterious

DISCUSSION

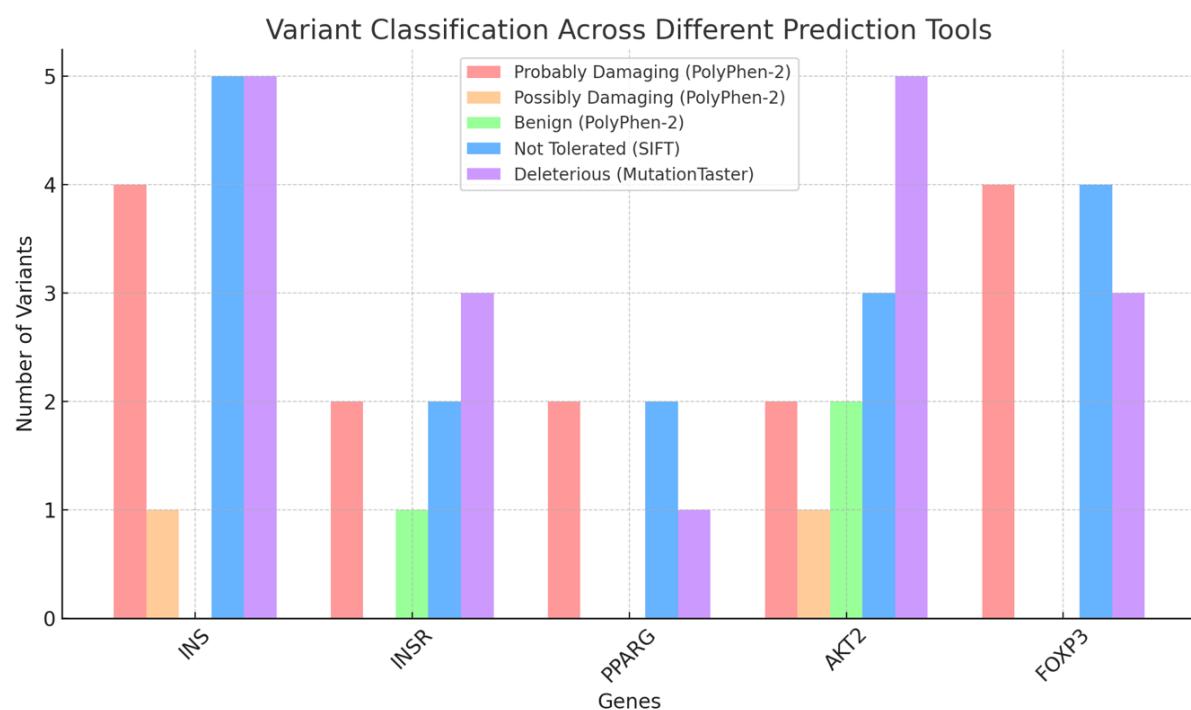


Figure 13 summarising the percentage of variants classified as harmful

The chart illustrates the number of variants classified as ‘Probably Damaging,’ ‘Possibly Damaging,’ ‘Benign’ (PolyPhen-2), ‘Not Tolerated’ (SIFT), and ‘Deleterious’ (MutationTaster) for the INS, INSR, PPARG, AKT2, and FOXP3 genes. These classifications help predict the potential pathogenicity of mutations affecting insulin-related functions.

For the INS gene, which plays a crucial role in insulin synthesis and secretion, and is also associated with autism, four variants were predicted to be "probably damaging" and one as "possibly damaging" in PolyPhen-2. SIFT classified all variants as "not tolerated," while MutationTaster identified them as "deleterious." These findings indicate that these variants could disrupt insulin function, potentially leading to metabolic disorders like diabetes, and contribute to the development of autism.

In the INSR gene, responsible for insulin signalling, and linked to autism, two variants were predicted as "probably damaging" and one as "benign" by PolyPhen-2. SIFT showed similar results, while MutationTaster classified all three variants as "deleterious," suggesting a possible link to insulin resistance, metabolic syndrome, endocrine disruption, and potentially, autism.

The PPARG gene, which regulates glucose metabolism and lipid storage, and has also been associated with autism, had two variants classified as "probably damaging" in PolyPhen-2 and "not tolerated" in SIFT. MutationTaster identified one variant as "deleterious," but no data was found for the other. This suggests that at least one variant could be involved in metabolic disorders like diabetes and obesity, and possibly contribute to autism, while further research is needed for the other.

In the AKT2 gene, which plays a role in insulin signalling and glucose homeostasis, and is also implicated in autism, two variants were "probably damaging," one was "possibly damaging," and two were "benign" in PolyPhen-2. SIFT classified three as "not tolerated" and two as "tolerated," while MutationTaster labelled all as "deleterious." This suggests that some of these variants might contribute to insulin resistance, related diseases, and potentially, the development of autism.

For the FOXP3 gene, which is crucial for immune system regulation, all variants were classified as "probably damaging" in PolyPhen-2 and "not tolerated" in SIFT. MutationTaster predicted three variants as "deleterious," while one was "benign." These results indicate that most of these variants could potentially impair immune function, leading to autoimmune diseases such as type 1 diabetes, inflammatory bowel disease (IBS).

Overall, the discrepancies between different prediction tools emphasize the importance of integrating multiple bioinformatics approaches to assess variant pathogenicity accurately. While many variants were classified as potentially harmful, affecting insulin-related pathways

and potentially contributing to disorders like diabetes, insulin resistance, and autism, further functional studies are required to confirm their actual impact on protein function and disease progression. These findings contribute to a better understanding of the genetic basis of insulin-related disorders, including their possible links to autism, and may help improve risk assessment, early diagnosis, and targeted therapeutic strategies.

CONCLUSION

The current study provides valuable insights into the potential pathogenicity of genetic variants in insulin-related genes using in-silico analysis. The findings suggest that several variants in INS, INSR, PPARG, AKT2, and FOXP3 genes may impact metabolic, neurological, immune, and reproductive functions, with most being classified as deleterious or not tolerated by predictive tools. While computational methods offer a strong foundation for identifying potentially harmful mutations, further experimental validation is necessary to confirm their biological significance. These results contribute to a better understanding of insulin-related genetic variants and their role in disease susceptibility, potentially aiding in improved diagnostic and therapeutic strategies.

REFERENCE

1. Wilcox, Gisela. "Insulin and Insulin Resistance." *Clinical Biochemist Reviews*, vol. 26, no. 2, May 2005, pp. 19–39. PubMed Central, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1204764/>.
2. Lebovitz, H. E. "Insulin Resistance: Definition and Consequences." *Experimental and Clinical Endocrinology & Diabetes: Official Journal, German Society of Endocrinology [and] German Diabetes Association*, vol. 109 Suppl 2, 2001, pp. S135-148. PubMed, <https://doi.org/10.1055/s-2001-18576>.
3. Pedersen, O. "[Insulin resistance--a physiopathological condition with numerous sequelae: non-insulin-dependent diabetes mellitus (NIDDM), android obesity, essential hypertension, dyslipidemia and atherosclerosis]." *Ugeskrift for Laeger*, vol. 154, no. 20, May 1992, pp. 1411–18.
4. Freeman, Andrew M., et al. "Insulin Resistance." StatPearls, StatPearls Publishing, 2025. PubMed, <http://www.ncbi.nlm.nih.gov/books/NBK507839/>.
5. About Polycystic Ovary Syndrome (PCOS) | NICHD - Eunice Kennedy Shriver National Institute of Child Health and Human Development. 21 Aug. 2024, <https://www.nichd.nih.gov/health/topics/pcos/conditioninfo>.
6. About Polycystic Ovary Syndrome (PCOS) | NICHD - Eunice Kennedy Shriver National Institute of Child Health and Human Development. 21 Aug. 2024, <https://www.nichd.nih.gov/health/topics/pcos/conditioninfo>.
7. Purwar, Ananya, and Shailesh Nagpure. "Insulin Resistance in Polycystic Ovarian Syndrome." *Cureus*, vol. 14, no. 10, p. e30351. PubMed Central, <https://doi.org/10.7759/cureus.30351>. Accessed 26 Feb. 2025.
8. Ataie-Ashtiani, Sara, and Briony Forbes. "A Review of the Biosynthesis and Structural Implications of Insulin Gene Mutations Linked to Human Disease." *Cells*, vol. 12, no. 7, Jan. 2023, p. 1008. www.mdpi.com, <https://doi.org/10.3390/cells12071008>.

9. Lee, Seung-Hwan, et al. “Insulin in the Nervous System and the Mind: Functions in Metabolism, Memory, and Mood.” *Molecular Metabolism*, vol. 5, no. 8, June 2016, pp. 589–601. PubMed Central, <https://doi.org/10.1016/j.molmet.2016.06.011>.
10. Sliwowska, Joanna H., et al. “Insulin: Its Role in the Central Control of Reproduction.” *Physiology & Behavior*, vol. 0, June 2014, pp. 197–206. PubMed Central, <https://doi.org/10.1016/j.physbeh.2014.05.021>.
11. Nandi, Anindita, et al. “The Effect of Insulin Signaling on Female Reproductive Function Independent of Adiposity and Hyperglycemia.” *Endocrinology*, vol. 151, no. 4, Apr. 2010, pp. 1863–71. DOI.org (Crossref), <https://doi.org/10.1210/en.2009-0788>.
12. Makhijani, Priya, et al. “Regulation of the Immune System by the Insulin Receptor in Health and Disease.” *Frontiers in Endocrinology*, vol. 14, Mar. 2023, p. 1128622. PubMed Central, <https://doi.org/10.3389/fendo.2023.1128622>.
13. UniProt. <https://www.uniprot.org/uniprotkb/P31751/entry>. Accessed 28 Feb. 2025.
14. Costa, Valerio, et al. “PPARG: Gene Expression Regulation and Next-Generation Sequencing for Unsolved Issues.” *PPAR Research*, vol. 2010, 2010, pp. 1–17. DOI.org (Crossref), <https://doi.org/10.1155/2010/409168>.
15. Mutskov, Vesco, and Gary Felsenfeld. “The Human Insulin Gene Is Part of a Large Open Chromatin Domain Specific for Human Islets.” *Proceedings of the National Academy of Sciences*, vol. 106, no. 41, Oct. 2009, pp. 17419–24. DOI.org (Crossref), <https://doi.org/10.1073/pnas.0909288106>.
16. UniProt. <https://www.uniprot.org/uniprotkb/P31751/entry>. Accessed 28 Feb. 2025.
17. UniProt. <https://www.uniprot.org/uniprotkb/Q99JB6/entry>. Accessed 3 Mar. 2025.

18. Bjørnvold, Marit, et al. “FOXP3 Polymorphisms in Type 1 Diabetes and Coeliac Disease.” *Journal of Autoimmunity*, vol. 27, no. 2, Sept. 2006, pp. 140–44. ScienceDirect, <https://doi.org/10.1016/j.jaut.2006.06.007>.
19. Bioinformatics. <https://www.pnnl.gov/explainer-articles/bioinformatics>. Accessed 16 Mar. 2025.
20. Heredity - Genes, DNA, Inheritance | Britannica. 27 Jan. 2025, <https://www.britannica.com/science/heredity-genetics/The-genetic-code>.
21. Medicine, Manchester Centre for Genomic. What Is Genetics? | Manchester Centre for Genomic Medicine. <https://www.mangen.co.uk/patients-visitors/what-is-genetics/>. Accessed 16 Mar. 2025.
22. Pruess, Manuela, and Rolf Apweiler. “Bioinformatics Resources for In Silico Proteome Analysis.” *Journal of Biomedicine and Biotechnology*, vol. 2003, no. 4, Oct. 2003, pp. 231–36. PubMed Central, <https://doi.org/10.1155/S1110724303209219>.
23. Mattaini, Katherine. Chapter 16. The Central Dogma: Genes to Traits. July 2020. [rwu.pressbooks.pub](https://rwu.pressbooks.pub/bio103/chapter/the-central-dogma-genes-to-trait/), <https://rwu.pressbooks.pub/bio103/chapter/the-central-dogma-genes-to-trait/>.