ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: https://doi.org/10.29121/granthaalayah.v5.i7.2017.2118

INTERNATIONAL JOURNAL OF RESEARCH – GRANTHAALAYAH
A knowledge Repository

Science

BIOINFORMATICS ANALYSIS OF GENES ASSOCITED WITH TYPE 2 DIABETES MELLITUS

Shruti Mishra[*1], Sunita Singh[2], Rajeev Mishra[3], Prashant Ankur Jain[4], Raghvendra Raman Mishra[5], Ved Kumar Mishra[6]

[1&3]Department of bioinformatics, Mahila Maha Vidalaya Banaras Hindu University, Varanasi, U.P., India

[2]Department of zoology, Mahila Maha Vidalaya, Banaras Hindu University, Varanasi, U.P., India

[4&6]Department of Computational Biology & Bioinformatics, Jacob Institute of Biotechnology & Bioengineering, Sam Higginbottom University of Agriculture, Technology & Sciences (Deemed to be University) Allahabad (U.P.), India

[5]Department of Medical Lab Technology, Institute of Medical Sciences, Banaras Hindu University (BHU), Varanasi, Uttar Pradesh, India

Abstract

Type 2 Diabetes mellitus is a multi-factorial disease caused due to gene defect as well as environmental factor. GWAS have played a primary role in demonstrating that genetic variation in a number of loci, SNPs, affects the risk of T2DM. there are our objective is to find out Disease pathway map by taking all genes of T2DM which are 35 in numbers and but in all there are 10 mostly involve in T2Dm from all over world population and it is find out by GWAS method then after we analyzed the KEGG pathway by analyzing T2DM pathway, Insulin signaling pathway, and WNT signalling pathway to find out common protein then after by bioinformatics analysis combined and expend these pathways toward common protein for understanding the Diseases mechanism. We do Protein-protein interaction and find out their complete target hub protein and target prediction for network hub. so for all these analysis I collect the total genes involve in T2DM and take those gene which are common for all population and their SNPs ,chromosome location in these all genes and by using string database I tried to find out the target protein hub which are found in this disease so there I have taken 5 most frequent genes and doing PPI in human so there are all have their own target protein hub-KCNJ11 have target protein hub PPKACA & TCF7L2 have complete target protein hub TLEI & PPARG have a target protein hub EP300 & CDKL1 have compete target protein hub UCB & HHEX complete target protein SOX2.

Keywords: GWAS; KEGG; T2DM; WNT; KCNJ11.

Cite This Article: Shruti Mishra, Sunita Singh, Rajeev Mishra, Prashant Ankur Jain, Raghvendra Raman Mishra, and Ved Kumar Mishra. (2017). "BIOINFORMATICS ANALYSIS OF GENES

ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379

ASSOCITED WITH TYPE 2 DIABETES MELLITUS." *International Journal of Research - Granthaalayah*, 5(7), 159-178. https://doi.org/10.29121/granthaalayah.v5.i7.2017.2118.

1. Introduction

Diabetes Mellitus (DM) is a group of metabolic diseases characterized by hyperglycaemia resulting from defects in insulin secretion, insulin action, or both. The basis of the abnormalities in carbohydrate, fat, and protein metabolism in diabetes is deficient action of insulin on target tissues. Deficient insulin action results from inadequate insulin secretion and/or diminished tissue responses to insulin at one or more points in the complex pathways of hormone action (1). Impairment of insulin secretion and defects in insulin action frequently coexist in the same patient, and it is often unclear which abnormality is the primary cause of the hyperglycaemia.

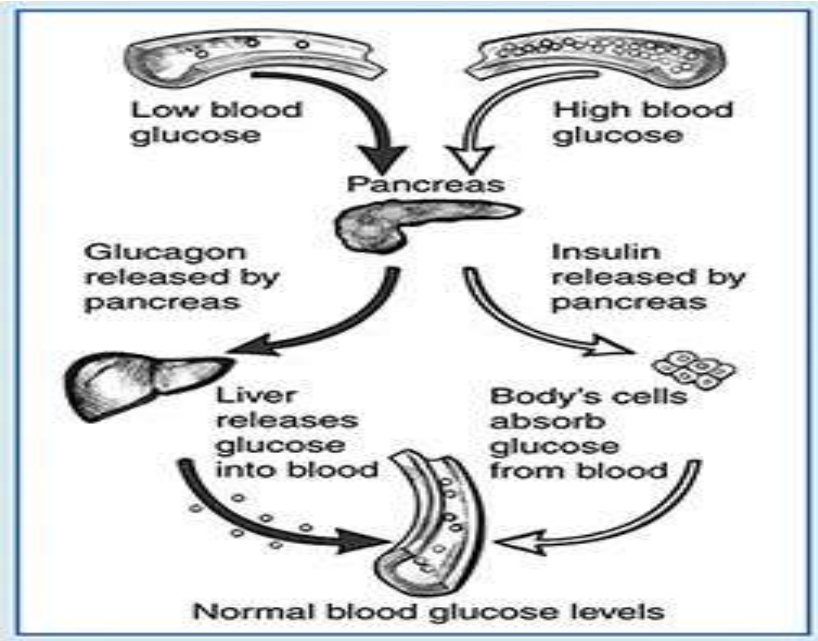


Figure 1: Blood glucose level

Diagnosis: If a diagnosis of diabetes is made, the clinician must feel confident that the diagnosis is fully established since the consequences for the individual are considerable and lifelong. The requirements for diagnostic confirmation for a person presenting with severe symptoms and gross hyperglycaemia differ from those for the asymptomatic person with blood glucose values found to be just above the diagnostic cut–off value. Severe hyperglycaemia detected under conditions of acute infective, traumatic, circulatory or other stress may be transitory and should not in itself be regarded as diagnostic of diabetes. The diagnosis of diabetes in an asymptomatic subject should never be made on the basis of a single abnormal blood glucose value. For the asymptomatic person, at least one additional plasma/blood glucose test result with a value in the diabetic range is essential, either fasting, from a random (casual) sample, or from the oral glucose tolerance test (OGTT) (2).

Type 1 diabetes mellitus is characterized by loss of the insulin-producing beta cells of the islets of Langerhans in the pancreas, leading to insulin deficiency. The majority of type 1 diabetes is of

ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379

the immune-mediated nature, in which a T-cell-mediated autoimmune attack leads to the loss of beta cells and thus insulin. This is Insulin Depandenent Diabetes Mellitus (IDDM). Type 2 diabetes mellitus (T2D) is characterized by persistent high blood glucose in the context of insulin resistance and relative insulin deficiency, due to pancreatic beta-cell dysfunction. Cardiovascular diseases, chronic renal failure, retinal, and nerve damage are usual complications of this illness. This is Non Insulin Depandenent Diabetes Mellitus(NIDDM) T2DM is a complex multifactorial disease which is caused due to genetic as well as environmental factor Type 2 diabetes mellitus (T2DM), a metabolic disorder characterized by insulin resistance and relative insulin deficiency, is a complex disease of major public health importance. Its incidence is rapidly increasing in the Developed countries. Complex diseases are caused by interactions between multiple genes and environmental factors. Most association studies aim to identify individual susceptibility single markers using a simple disease model. Recent studies are trying to estimate the effects of multiple genes and multi-locus in genome-wide association. However, estimating the effects of association is very difficult (1, 3).

Gestational diabetes mellitus (GDM) resembles type 2 diabetes in several respects, involving a combination of relatively inadequate insulin secretion and responsiveness. It occurs in about 2–10% of all pregnancies and may improve or disappear after delivery. Maturity onset type diabetes of the young (MODY) was previously considered to be a third form of type 2 diabetes. However, with the discovery of specific mutations leading to MODY, it is now classified under secondary or other specific types of diabetes. MODY is characterized by onset prior to age 25. All cases to date have shown impaired β-cell function (2).

Autoimmune Diabetes Mellitus
This form of diabetes, previously encompassed by the terms insulin-dependent diabetes, Type 1 diabetes, or juvenile-onset diabetes, results from autoimmune mediated destruction of the beta cells of the pancreas. The rate of destruction is quite variable, being rapid in some individuals and slow in others. The rapidly progressive form is commonly observed in children, but also may occur in adults (3).

1.1. Other genetic syndromes

Genetic defects of beta-cell function
Several forms of the diabetic state may be associated with monogenic defects in beta-cell function, frequently characterized by onset of mild hyperglycemia at an early age (generally before age 25 years). They are usually inherited in an autosomal dominant pattern. Patients with these forms of diabetes, formerly referred to as maturity-onset diabetes of the young (MODY), have impaired insulin secretion with minimal or no defect in insulin action (4-6). Abnormalities at three genetic loci on different chromosomes have now been characterized. The most common form is associated with mutations on chromosome 12 in a hepatic nuclear transcription factor referred to as HNF1alpha (4). A second form is associated with mutations in the glucokinase gene on chromosome 7p (5, 8).

Genetic defects in insulin action
There are some unusual causes of diabetes which result from genetically determined abnormalities of insulin action. The metabolic abnormalities associated with mutations of the

ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379

insulin receptor may range from hyper-insulinaemia and modest hyperglycemia to symptomatic diabetes (7).

Drug- or Chemical-induced Diabetes

- Nicotinic acid
- Glucocorticoids
- Thyroid hormone
- Alpha-adrenergic agonists
- Beta-adrenergic agonists
- Interferon-alpha therapy
- Others

Type 2 Diabetes Mellitus-Genome-wide association studies (GWASs) have discovered association of several loci with Type 2 diabetes (T2D), a common complex disease characterized by impaired insulin secretion by pancreatic beta cells and insulin signalling in target tissues (9-11). The genes influencing common complex or multifactorial diseases T2D has a complex pathogenesis that was classically characterized by pancreatic beta-cell dysfunction (with diminished insulin secretion) followed by decline of the beta cell mass Ethnic variation of T2D represents strong evidence for the genetic basis of this disease. The maximum prevalence is recorded in Pima Indians from USA and South Sea Island populations where it now reaches ~50%. A low prevalence (~3%) is recorded in some African populations1 while the lowest (~1%) is recorded in some isolated rural populations from South America. Type 2 diabetes mellitus (T2DM) is a complex disease; both environmental and genetic factors are involved in the development of the disease. Environmental factors of particular importance include overweight/obesity, increased amount of body fat, hypertension, lack of physical exercise. Regarding the genetic factors, there is evidence that T2DM has a strong genetic component, as long appreciated by physicians because of the role as risk indicators of family history of diabetes and of ethnicity (12-15).

Genetic Predisposition

The fact that type 2 diabetes is a genetic disease is well known to clinicians by how it occurs in families, and by there being ethnic populations who are particularly high risk. the genetic basis for many monogenic forms of diabetes has been discovered such as mitochondrial genome defects and the association with diabetes there may be gene mutation, or dislocation of chromosome and SNP Environmental factor- The second factor is environmental aspects (16, 17). An important concept is the diabetes genotype typically causes only a predisposition for glucose intolerance Whether one develops the diabetes phenotype depends on environmental factors, Numerous multifactorial mechanisms that include genetic and environmental factors related to obesity are involved in the development of insulin resistance and impaired insulin secretion. Insulin resistance is associated with inactivity, obesity and ageing. The insulin secreting pancreatic islet b cells respond to insulin resistance by enhancing their mass and metabolic function. T2D however develops when increase in insulin secretion by b cells is not able to keep pace with the increase in insulin resistance (18).

SNP-A **Single Nucleotide Polymorphism**, DNA sequence variation occurring commonly within a population (e.g. 1%) in which a single nucleotide — A, T, C or G — in the genome (or other

 ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379

shared sequence) differs between members of a biological species or paired chromosomes. For example, two sequenced DNA fragments from different individuals, AAGCCTA to AAGCTTA, contain a difference in a single nucleotide. In this case we say that there are two *alleles*. Almost all common SNPs have only two alleles. The genomic distribution of SNPs is not homogenous; SNPs occur in non-coding regions more frequently than in coding regions or, in general, where natural selection is acting and 'fixing' the allele (eliminating other variants) of the SNP that constitutes the most favorable genetic adaptation. Other factors, like genetic recombination and mutation rate, can also determine SNP density (19, 20).

2. Materials and Methods

- Data set collection-data set, total no. of genes which are involved in predisposing to T2D from different GWAS studies
- Major genes predisposing to T2D in different ethnic groups
- Single nucleotide polymorphism (SNPs) involved in different genes and their chromosome location.
- Data on protein and their sequences collection
- Analysis of protein which are present in protein interaction which is common. (STRING DATABASE)
- Pathway Analysis (e.g. use of KEGG tool-check common protein in pathways)

(Protein analysis which are found maximum in Pathway).
New target Prediction-Protein which are present on crucial point will suggest new theraptic target for Diseases (20).

3. Result and Discussion

KEGG DATABSE-Kyoto Encyclopedia of Genes and Genomes (KEGG) is a knowledge base for systematic analysis of gene functions in terms of the networks of genes and Molecules. The major component of KEGG is the PATHWAY database that consists of graphical diagrams of biochemical pathways including most of the known metabolic pathways and some of the known regulatory pathways. The pathway information is also represented by the ortholog group tables summarizing orthologous and analogous gene groups among different organisms. KEGG maintains the GENES database for the gene catalogs of all organisms with complete genomes and selected organisms with partial genomes, which are continuously re-annotated, As well as the LIGAND database for chemical compounds and enzymes.

String Database
Information on protein–protein interactions is still mostly limited to a small number of model organisms, and originates from a wide variety of Experimental and computational techniques. The database and online resource STRING generalizes access to protein interaction data, by integrating known and predicted interactions from a variety of sources. The underlying infrastructure includes a consistent body of completely sequenced genomes and exhaustive orthology classifications, based on which interaction evidence is transferred between Organisms. Although primarily developed for protein interaction analysis, the resource has also been successfully applied to comparative genomics, phylogenetics and network studies, T2D is a complex disease, the genetic risk being influenced by the conjoint effects of variation at an

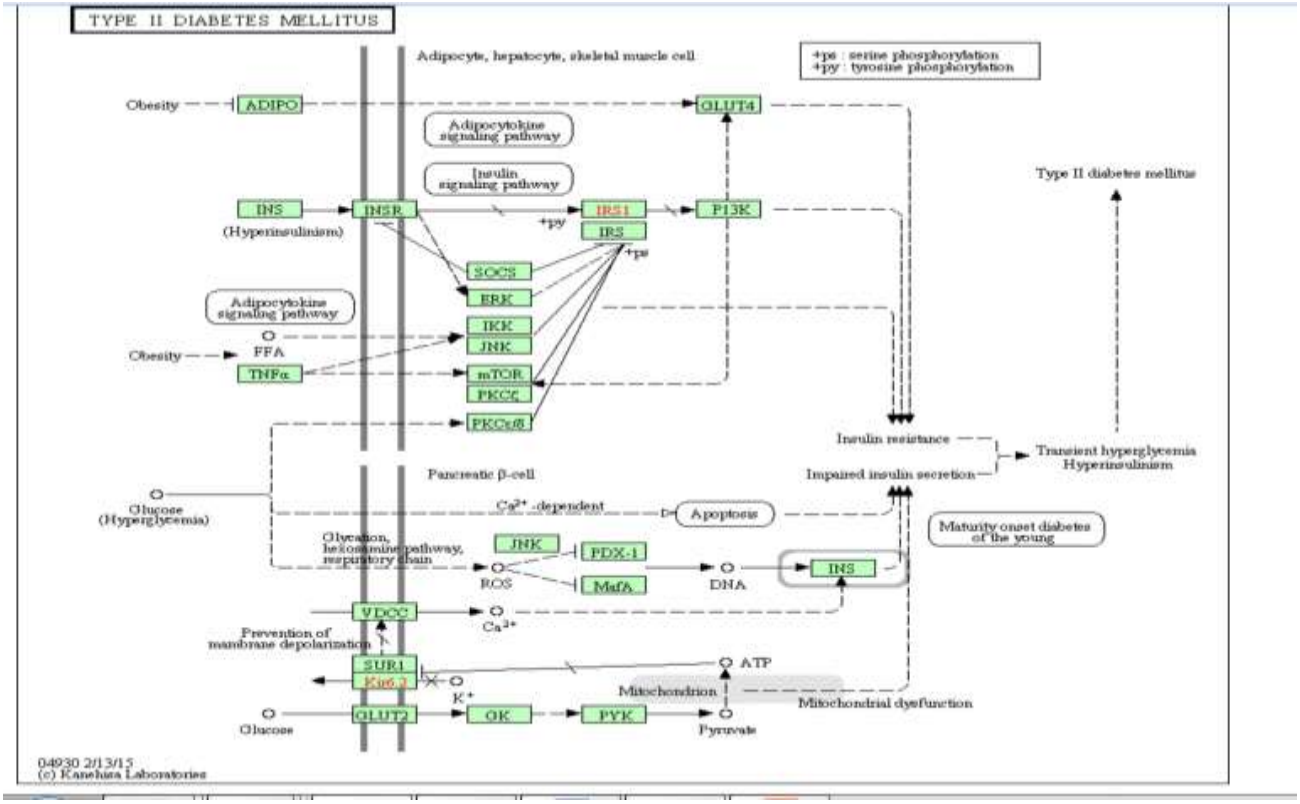
ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379

undetermined number of genomic sites. The main methods for mapping the T2D genes were the hypothesis driven candidate gene analysis and the hypothesis free genome-wide scanning studies. The candidate gene approach led to the identification of two T2D genes now considered widely replicated: PPARG and the β-cell potassium channel gene, KCNJ11. The genome-wide linkage approach led to the identification of several loci, the most prominent being the TCF7L2 (Transcription Factor 7 Like 2) gene on chromosome 10q25.3. TCF7L2 has been replicated in almost every population examined and, with an OR of about 1.4, represents the strongest T2D gene identified so far. Finally, during the last 5 years, the genome-wide association approach led to the identification of almost 40 T2D genes. The majority of these appear to affect beta cell function. Deciphering the genetic background of T2D will contribute to the prediction of the disease in high risk subjects, with possible benefits for its Recently, genes discovered to be significantly associated with developing type 2 DM, include *TCF7L2*, *PPARG*, *FTO*, *KCNJ11*, *NOTCH2*, *WFS1*, *CDKAL1*, *IGF2BP2*, *SLC30A8*, *JAZF1*, and *HHEX*. *KCNJ11* (potassium inwardly rectifying channel, subfamily J, member 11), encodes the islet ATP-sensitive potassium channel Kir6.2, and *TCF7L2* (transcription factor 7-like 2) regulates proglucagon gene expression and thus the production of glucagon-like peptide-1.

Table 1: SNP associated with type-2 diabetes identified by GWAS

SNP ID	Candidate gene	Risk allele	Chr.	SNP position (build 37.1)	SNP location	Reference
pG-SNPs where risk alleles remove a CpG site ($n=8$)						
rs5945326	*DUSP9*[a]	A	X	152899922	Intergenic	Voight et al (2010) [20]
rs564398	*CDKN2A*[a]	A	9	22029547	Intergenic	Zeggini et al (2007) [21]
rs11708067	*ADCY5*[a]	A	3	123065778	Intron	Dupuis et al (2010) [22]
rs1801214	*WFS1*[a]	T	4	6303022	Cds-synon	Voight et al (2010) [20]
rs2334499	*DUSP8*[a]	A	11	1696849	Intergenic	Kong et al (2009) [23]
rs7578326	*IRS1*[a]	A	2	227020653	Intergenic	Voight et al (2010) [20]
rs5219	*KCNJ11*[b]	T	11	17409572	Cds-nonsynon	Scott et al (2007) [24]
rs1801282	*PPARG*[a]	C	3	12393125	Cds-nonsynon	Scott et al (2007) [24]
pG-SNPs where risk alleles introduce a CpG site ($n=11$)						
rs13292136	*CHCHD9*[a]	C	9	81952128	Intergenic	Voight et al (2010) [20]
rs7901695	*TCF7L2*[a]	G	10	114754088	Intron	Zeggini et al (2007) [21]
rs7754840	*CDKAL1*[a]	C	6	20661250	Intron	Scott et al (2007) [24]
rs391300	*SRR*[a]	G	17	2216258	Intron	Tsai et al (2010) [25]
rs5015480	*HHEX*[a]	G	10	94465559	Intergenic	Voight et al (2010) [20]
rs13266634	*SLC30A8*[a]	G	8	118184783	Cds-nonsynon	Sladek et al (2007) [26]
rs4457053	*ZBED3*[a]	G	5	76424949	Intergenic	Voight et al (2010) [20]
rs7961581	*TSPAN8*[a]	C	12	71663102	Intergenic	Zeggini et al (2008) [27]
rs1531343	*HMGA2*[a]	C	12	66174894	Intergenic	Voight et al (2010) [20]
rs2237895	*KCNQ1*[a]	C	11	2857194	Intron	Yasuda et al (2008) [28]
rs12779790	*CDC123*[a]	G	10	12328010	Intergenic	Zeggini et al (2008) [27]
on-CpG SNPs ($n=21$)						
rs7961581	*TSPAN8*[a]	C	12	71663102	Intergenic	Zeggini et al (2008) [27]
rs1531343	*HMGA2*[a]	C	12	66174894	Intergenic	Voight et al (2010) [20]
rs2237895	*KCNQ1*[a]	C	11	2857194	Intron	Yasuda et al (2008) [28]
rs12779790	*CDC123*[a]	G	10	12328010	Intergenic	Zeggini et al (2008) [27]
Non-CpG SNPs ($n=21$)						
rs243021	*BCL11A*	A	2	60584819	Intergenic	Voight et al (2010) [20]
rs10830963	*MTNR1B*	G	11	92708710	Intron	Dupuis et al (2010) [22]
rs7578597	*THADA*	T	2	43732823	Missense	Zeggini et al (2008) [27]
rs17584499	*PTPRD*	T	9	8879118	Intron	Tsai et al (2010) [25]
rs10923931	*NOTCH2*	T	1	120317959	Intron	Zeggini et al (2008) [27]
rs7593730	*RBMS1*	C	2	161171454	Intron	Qi et al (2010) [29]
rs4607103	*ADAMTS9*	C	3	64711904	Intergenic	Zeggini et al (2008) [27]
rs1470579	*IGF2BP2*	C	3	185529080	Intron	Voight et al (2010) [20]
rs864745	*JAZF1*	T	7	28180556	Intron	Zeggini et al (2008) [27]
rs972283	*KLF14*	G	7	130466854	Intergenic	Voight et al (2010) [20]
rs896854	*TP53INP1*	T	8	95960511	Intron	Voight et al (2010) [20]
rs1552224	*CENTD2*	A	11	72433098	Intron	Voight et al (2010) [20]
rs7957197	*HNF1A*	T	12	121460686	Intron	Voight et al (2010) [20]
rs11634397	*ZFAND6*	G	15	80432222	Intergenic	Voight et al (2010) [20]
rs8042680	*PRC1*	A	15	91521337	Intron	Voight et al (2010) [20]
rs9939609	*FTO*	A	16	53820527	Intron	Frayling et al (2007) [30]
rs4607517	*GCK*	A	7	44235668	Intergenic	Dupuis et al (2010) [22]
rs2191349	*DGKB*	T	7	15064309	Intergenic	Dupuis et al (2010) [22]
rs780094	*GCKR*	C	2	27741237	Intron	Dupuis et al (2010) [22]
rs340874	*PROX1*	C	1	214159256	Intergenic	Dupuis et al (2010) [22]
rs4430796	*HNF1B*	G	17	36098040	Intron	Voight et al (2010) [20]

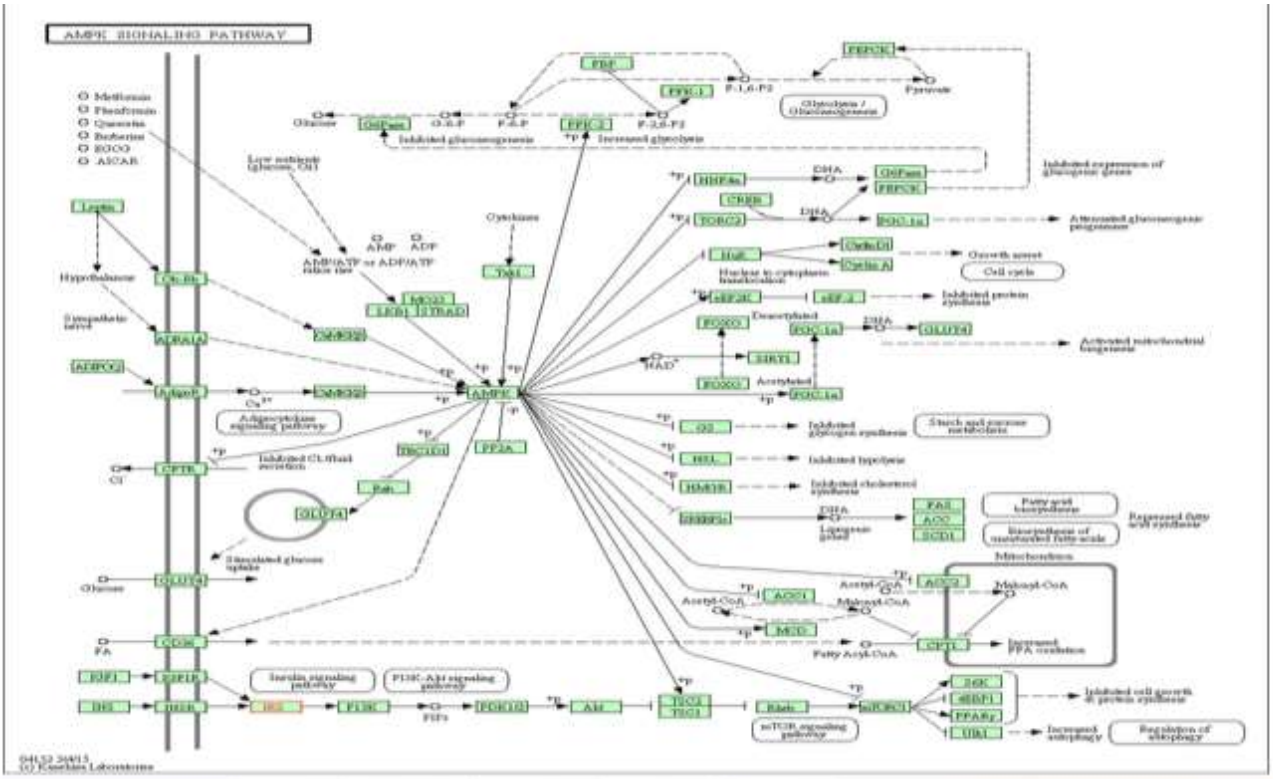
ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379

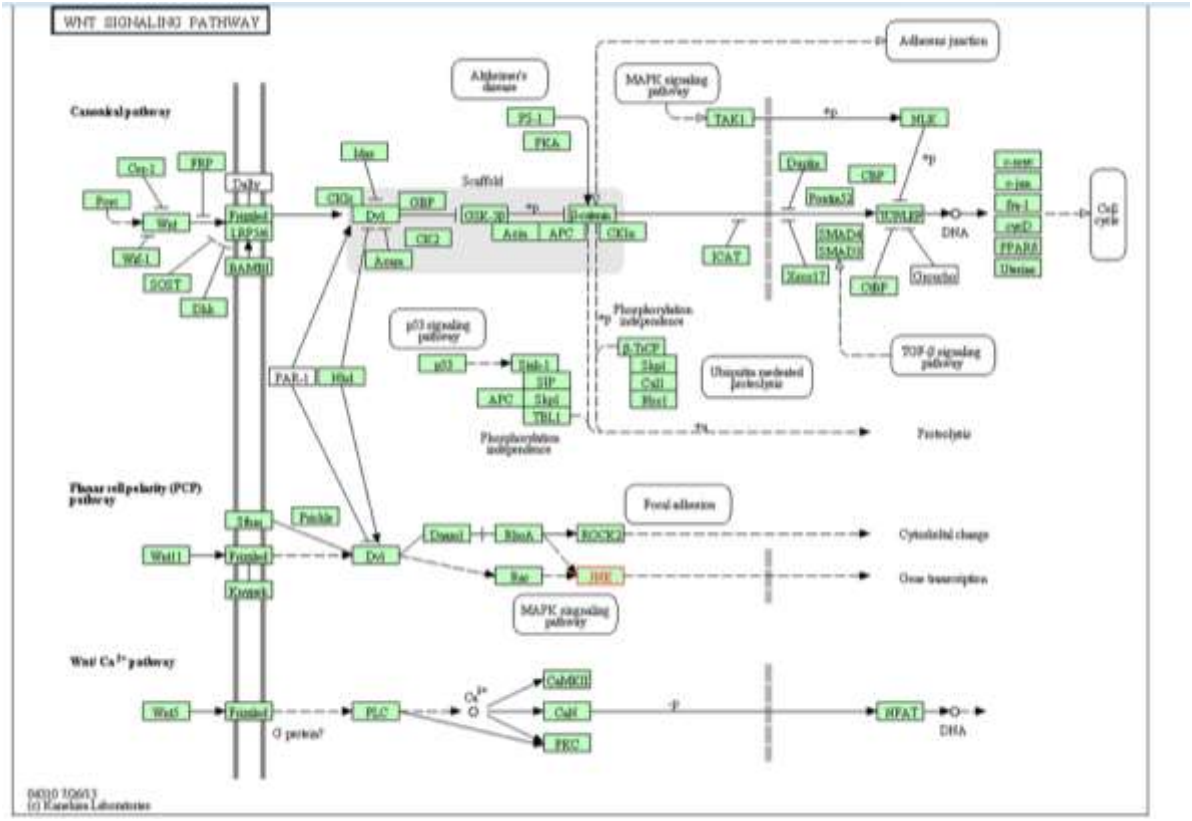


Insulin resistance is strongly associated with type II diabetes. "Diabetogenic" factors including FFA, TNFalpha and cellular stress induce insulin resistance through inhibition of IRS1 functions. Serine/threonine phosphorylation, interaction with SOCS, regulation of the expression, modification of the cellular localization, and degradation represent the molecular mechanisms stimulated by them. Various kinases (ERK, JNK, IKKbeta, PKCzeta, PKCtheta and mTOR) are involved in this process. The development of type II diabetes requires impaired beta-cell function. Chronic hyperglycemia has been shown to induce multiple defects in beta-cells.

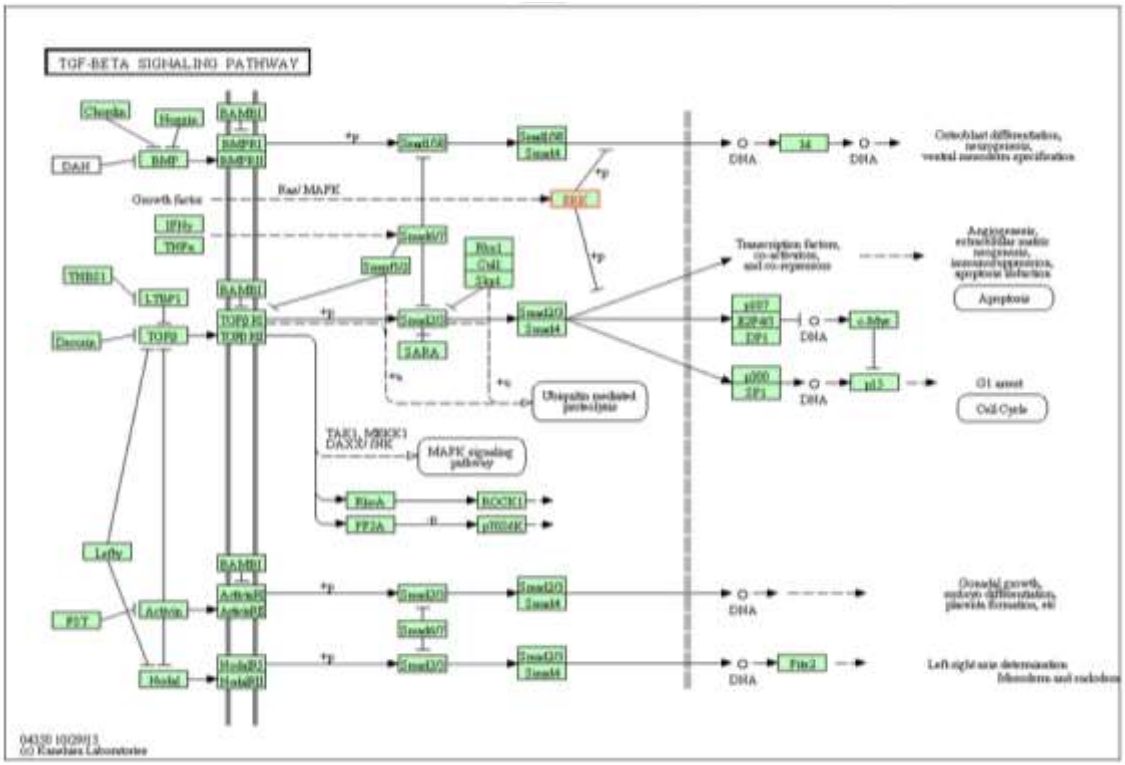
Hyperglycemia has been proposed to lead to large amounts of reactive oxygen species (ROS) in beta-cells, with subsequent damage to cellular components including PDX-1. Loss of PDX-1, a critical regulator of insulin promoter activity, has also been proposed as an important mechanism leading to beta-cell dysfunction. Although there is little doubt as to the importance of genetic factors in type II diabetes, genetic analysis is difficult due to complex interaction among multiple susceptibility genes and between genetic and environmental factors. Genetic studies have therefore given very diverse results. Kir6.2 and IRS are two of the candidate genes. It is known that Kir6.2 and IRS play central roles in insulin secretion and insulin signal transmission, respectively.IRS (insulin resistance substrate) regulated by gene by SOCS,ERK, IKK, JNK mTOR, PKC family. and IRS1 regulate the activity of PI3K

ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379





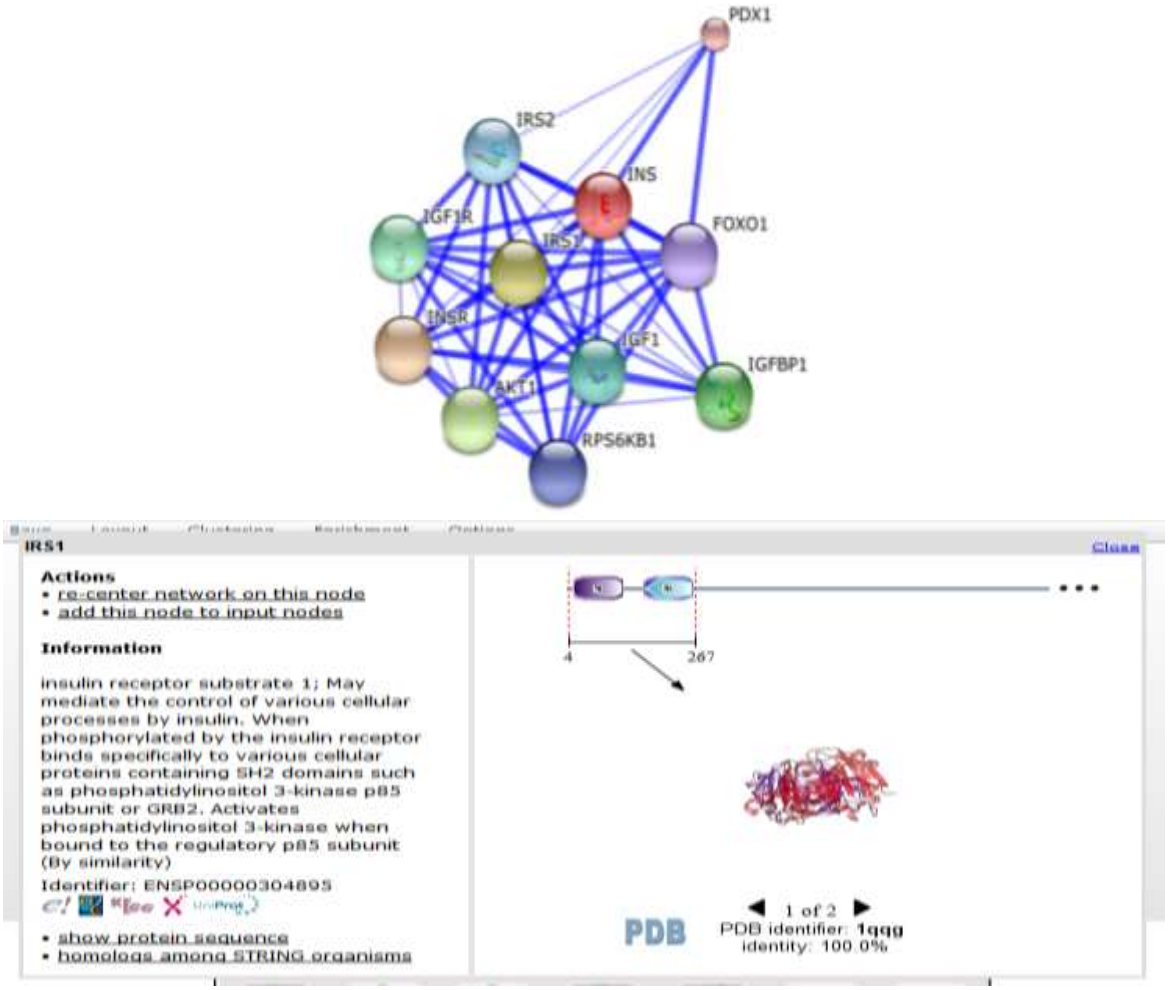
ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379



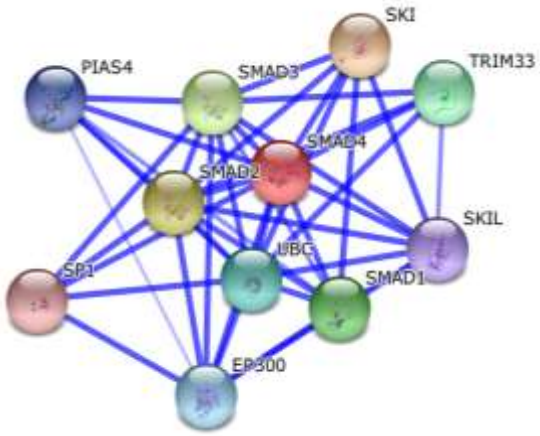
The transforming growth factor-beta (TGF-beta) family members, which include TGF-betas, activins and bone morphogenetic proteins (BMPs), are structurally related secreted cytokines found in species ranging from worms and insects to mammals. A wide spectrum of cellular functions such as proliferation, apoptosis, differentiation and migration are regulated by TGF-beta family members. TGF-beta family member binds to the Type II receptor and recruits Type I, whereby Type II receptor phosphorylates and activates Type I. The Type I receptor, in turn, phosphorylates receptor-activated Smads (R-Smads: Smad1, Smad2, Smad3, Smad5, and Smad8). Once phosphorylated, R-Smads associate with the co-mediator Smad, Smad4, and the heteromeric complex then translocates into the nucleus. In the nucleus, Smad complexes activate specific genes through cooperative interactions with other DNA-binding and coactivator (or co-repressor) proteins.

Here from all pathway analysis I have found that INS,PI3K,RAS,GLUT4,SMAD,I NRS genes are common for all pathways that is T2DM pathway, AMPK pathway, TGF-signaling pathway, Insulin signaling pathway and for WNT sinaling pathway further analysis and protein protein interaction can b done by string database. Here after seeing the T2DM pathway and involve pathway like insulin signaling pathway, TGF-signaling pathway and MPAK pathway I found that there are # gene are common for all pathway and is most frequent in whole population like SMAD gene

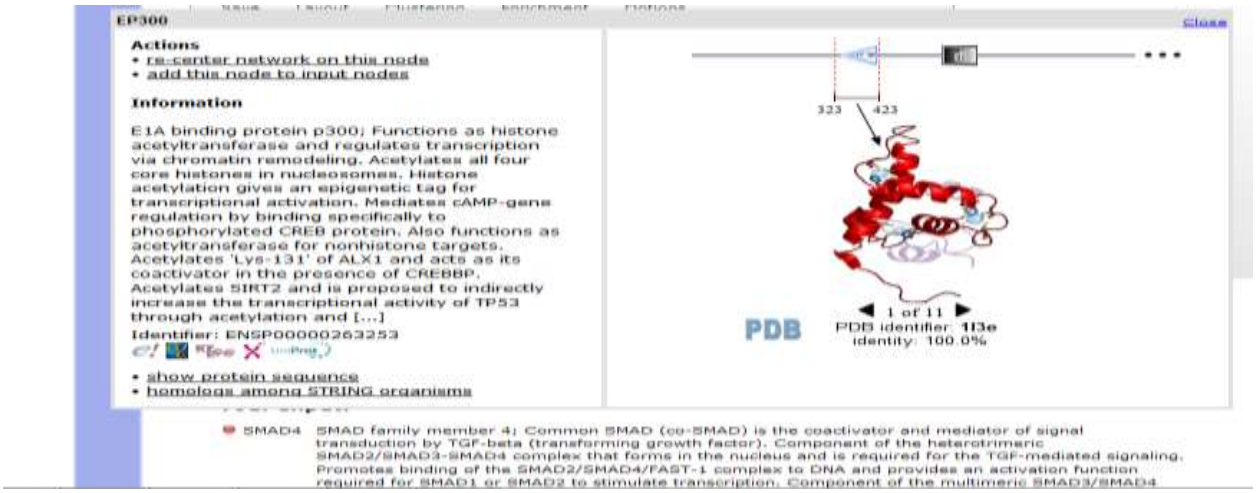
ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379



INS(insulin) protein of all over pathway that are common for all pathway having complete interaction with IRS(insulin receptor substrate)



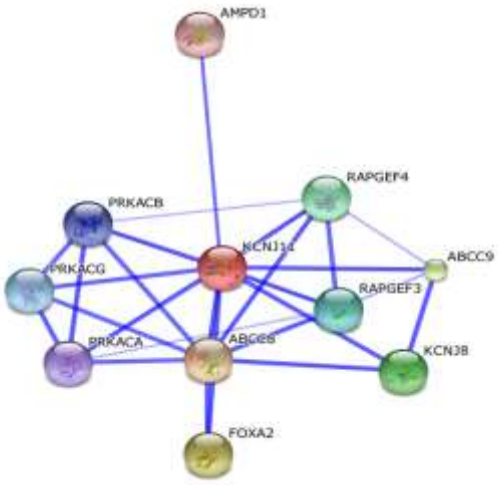
ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379

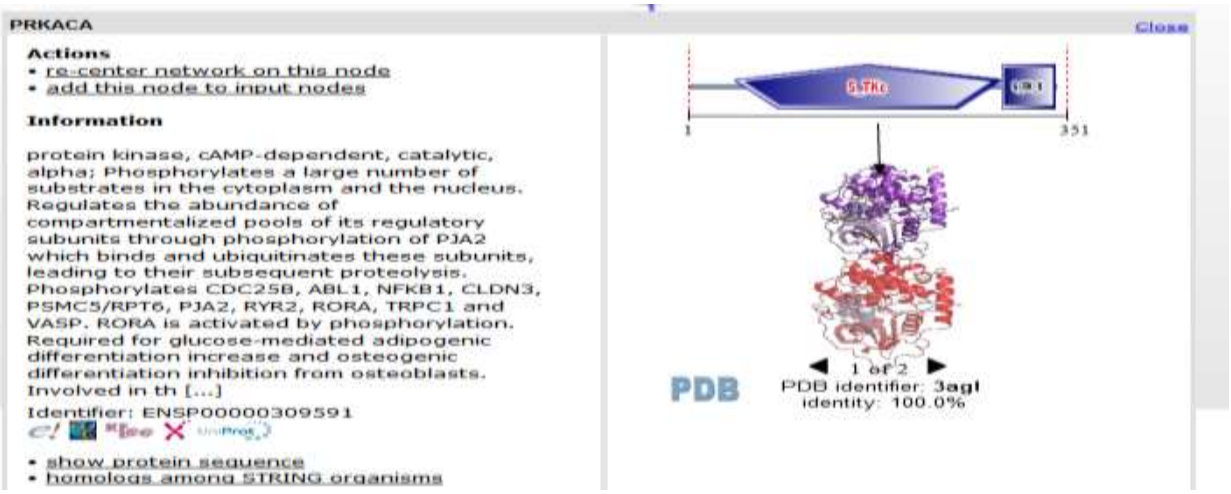


SMAD having complete interaction with E 300 and UBC both proteins. SMAD and INS ,IRS,GLUT4 are common genes which are constructed from all the pathways which are involve in type 2Diabetes mellitus *PPARG KCNJ11/ ABCC8 TCF7L2 IGF2BP2 CDKAL1 SLC30A8 CDKN2A/B HHEX FTO HNF1 NOTCH2 THADA ADAMTS9 JAZF1 CDC123/CAMK1D KCNQ1 TSPAN8/ LGR5 IRS1 DUSP9 PROX1 BCL11A GCKR ADCY5 WFS1 ZBED3 DGKB/TMEM1 GCK 7 KLF14 7 TP53INP1 TLE4/CHCHD9 CENTD2 MTNR1B HMGA2*

HNF1A PRC ZFAND6 .SMAD There are total 35 to 40 genes are involve in type 2 Diabetes Mellitus in which 5 to 10 are most frequent in all over world population and find out by GWAS genome wide analysis All these gene are very important genes which are involve in T2DM frequently but there are some gene which are more frequent in all population like India ,south Africa, china, USA ,Japan ,jarmen,itely,brazil that are Recently, genes discovered to be significantly associated with developing type 2 DM, include TCF7L2, PPARG, FTO KCNJ11, NOTCH2, WFS1, CDKAL1, IGF2BP2, SLC30A8, JAZF1, and HHEX. KCNJ11 (potassium inwardly rectifying channel, subfamily J, member 11), encodes the islet ATP-sensitive potassium channel Kir6.2, and TCF7L2 (transcription factor 7-like 2) regulates pro-glucagon gene expression and thus the production of glucagon-like peptide-1 We can analyse it via KEGG PATHWAY MAP and string database and we were found that in type 2diabetes pathways SOCS,ERK,IKK,JNK,mTOR and PKCr18 promote the production of IRS and it regulate PI3K PI3k signalling pathway and WNT signalling pathways1-Through string database most frequent gene KCNJ11 potassium inwardly-rectifying channel, subfamily J, member 11 as a hub protein where other interacting partners are AMDD1, PRKACB, ABC9C9, KCNJ8, RAPGEF3, ABCC8, PRKACA, PRKACG, PRKACB where PRKACAI (protein kinase, cAMP-dependent, catalytic, alpha) gene is 100% identical with KCNJ11 protein kinase, cAMP-dependent, catalytic, alpha; Phosphorylates a large number of substrates in the cytoplasm and the nucleus. Regulates the abundance of compartmentalized pools of its regulatory subunits through phosphorylation of PJA2 which binds and ubiquitinates these subunits, leading to their subsequent proteolysis RAPGEF4 shows 95% identity with KCNJ11 and FOXAL show 91%identity with KCNJ11the least identity shows by ABCC8 gene that is 23.9% so it is a least interacting partner of KCNJ11

ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379





PROTEIN SEQUANCE-

PRKACA *[ENSP00000309591]*, Homo sapiens

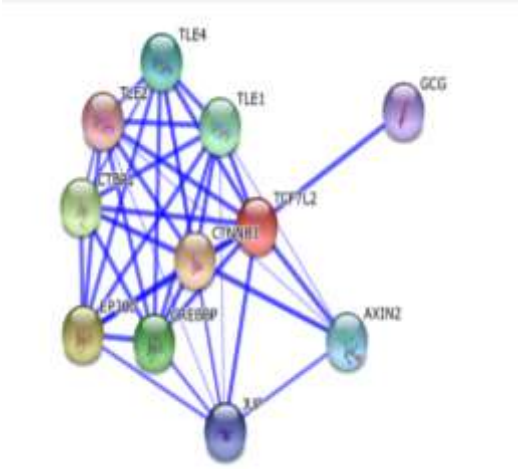
protein kinase, cAMP-dependent, catalytic, alpha; Phosphorylates a large number of substrates in the cytoplasm and the nucleus. Regulates the abundance of compartmentalized pools of its regulatory subunits through phosphorylation of PJA2 which binds and ubiquitinates these subunits, leading to their subsequent proteolysis. Phosphorylates CDC25B, ABL1, NFKB1, CLDN3, PSMC5/RPT6, PJA2, RYR2, RORA, TRPC1 and VASP. RORA is activated by phosphorylation. Required for glucose-mediated adipogenic differentiation increase and osteogenic differentiation inhibition from osteoblasts. Involved in th [...]

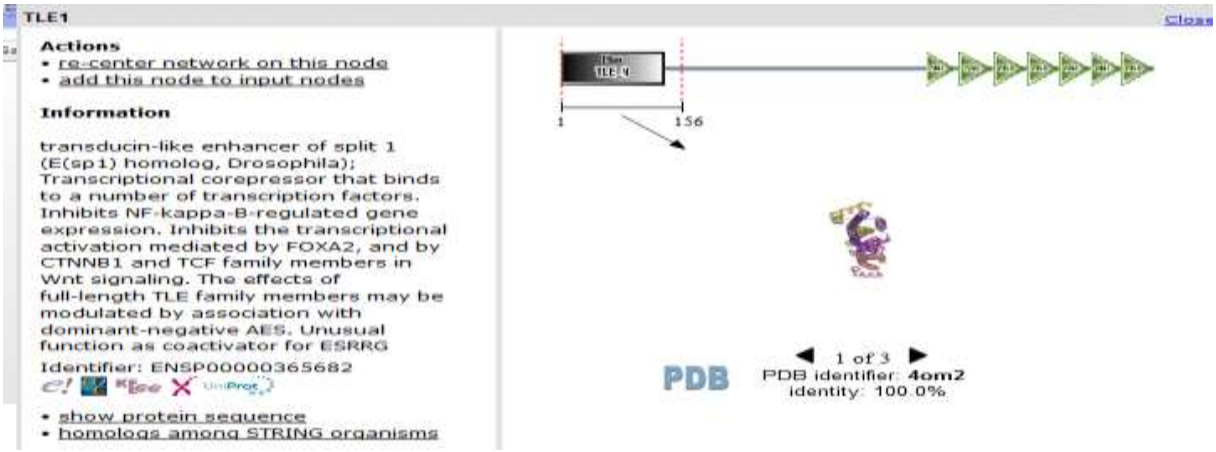
source: e!

```
MGNAAAAKKGSEQESVKEFLAKAKEDFLKKWESPAQNTAHLDQFERIKTLGTGSFGRVMLVKHKETGNHYAMKILDKQKV
VKLKQIEHTLNEKRILQAVNFPFLVKLEFSFKDNSNLYMVMEYVPGGEMFSHLRRIGRFSEPHARFYAAQIVLTFEYLHS
LDLIYRDLKPENLLIDQQGYIQVTDFGFAKRVKGRTWTLCGTPEYLAPEIILSKGYNKAVDWWALGVLIYEMAAGYPPFF
ADQPIQIYEKIVSGKVRFPSHFSSDLKDLLRNLLQVDLTKRFGNLKNGVNDIKNHKWFATTDWIAIYQRKVEAPFIPKFK
GPGDTSNFDDYEEEEIRVSINEKCGKEFSEF
```

2-In other gene like TCF7L2 transcription factor 7-like 2 (T-cell specific, HMG-box) (602 aa) *where GCG & TLE1 shows 100% identity with hub protein TLE1* transducin-like enhancer of split 1 (E(sp1) homolog, Drosophila); Transcriptional corepressor that binds to a number of transcription factors. Inhibits NF-kappa-B-regulated gene expression. Inhibits the transcriptional Activation mediated by FOXA2, and by CTNNB1 and TCF family members in Wnt signaling,*TNNB1 shows 97% identity with hub protein.*

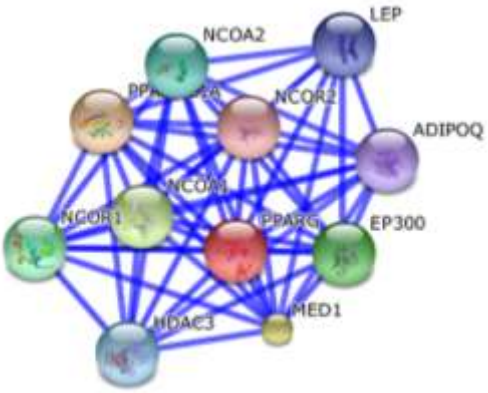
ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379

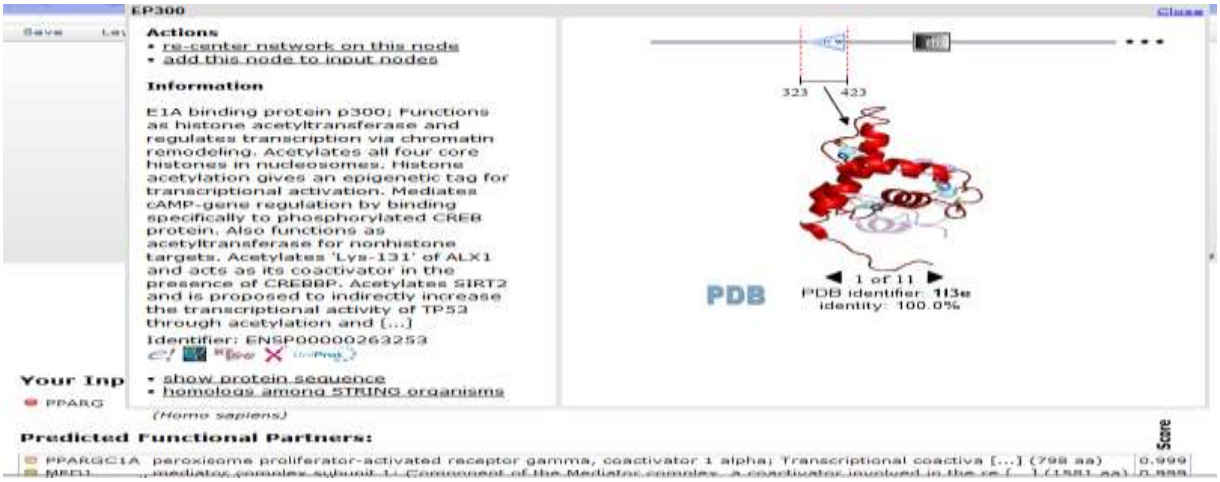




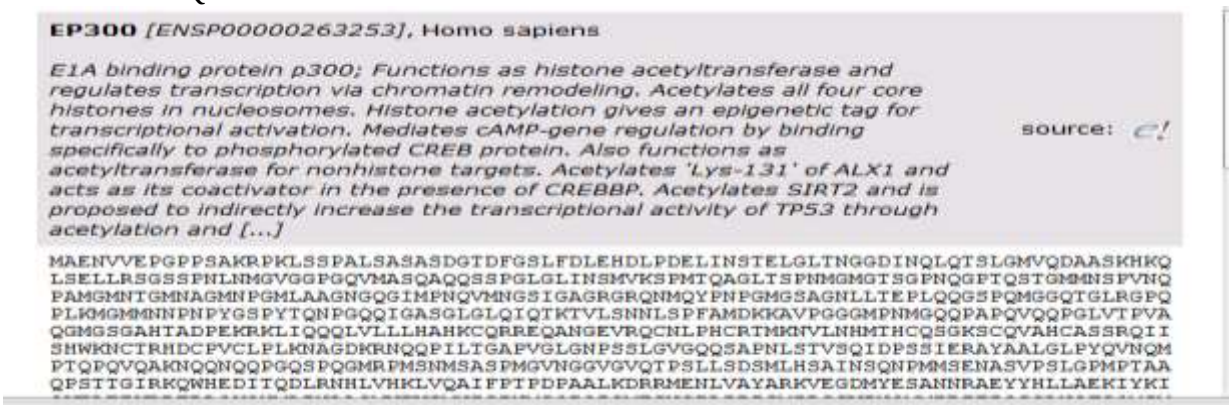
4-In PPARG peroxisome proliferator-activated receptor gamma (505 aa) *gene their interacting partners are E P300,MED1,ADIPOQ,LEP,NCOA1,NCOR2,HDAC3 in all these protein -protein interaction(PPI) EP300 shows complete 100% interaction with PPARG But LEP shows 99% identity with hub protein NCOR2 shows 97% identity while PRARGC1A shows only 28% identity with hub protein.*

ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379



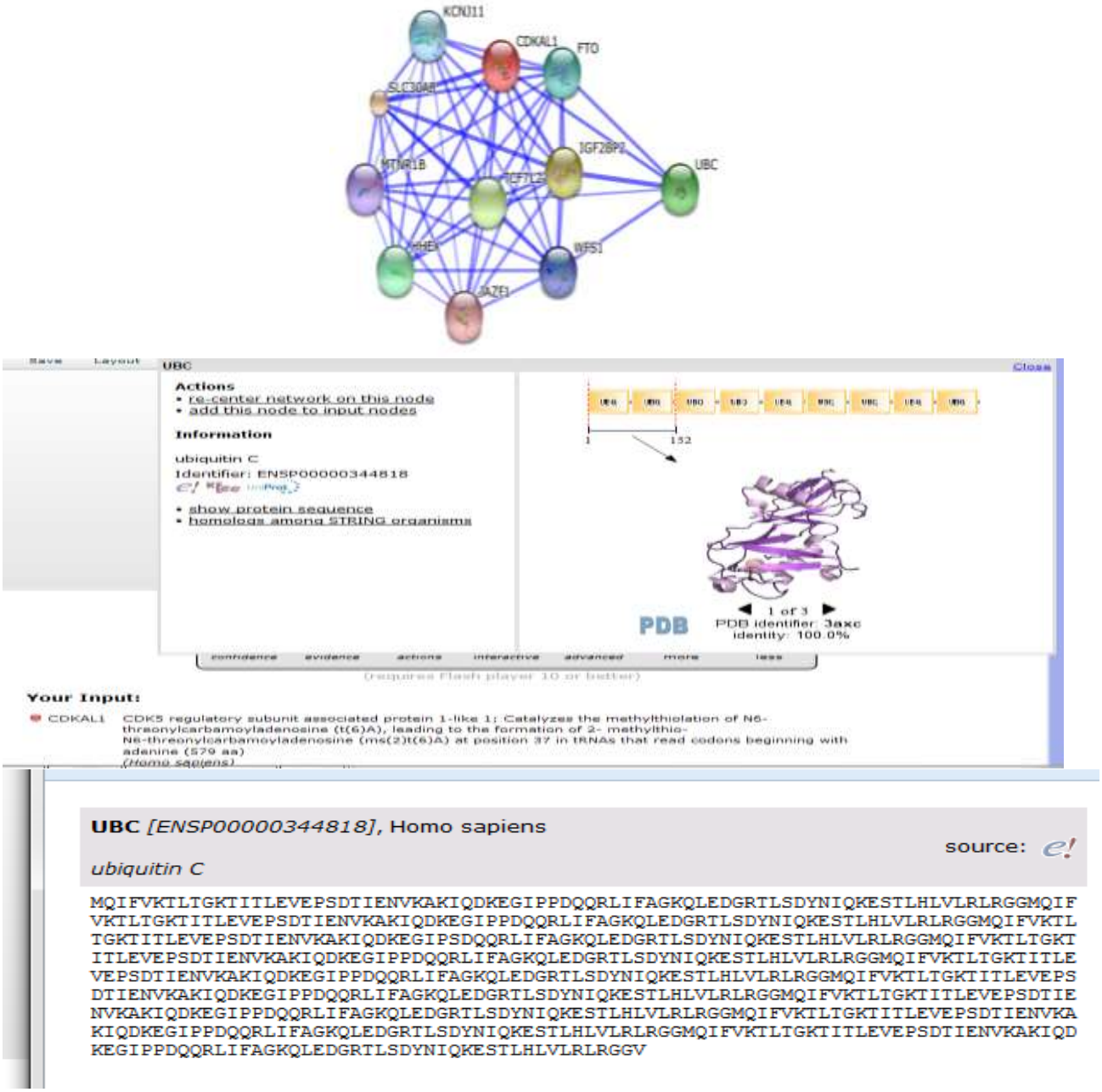


PROTEIN SEQUANCE-



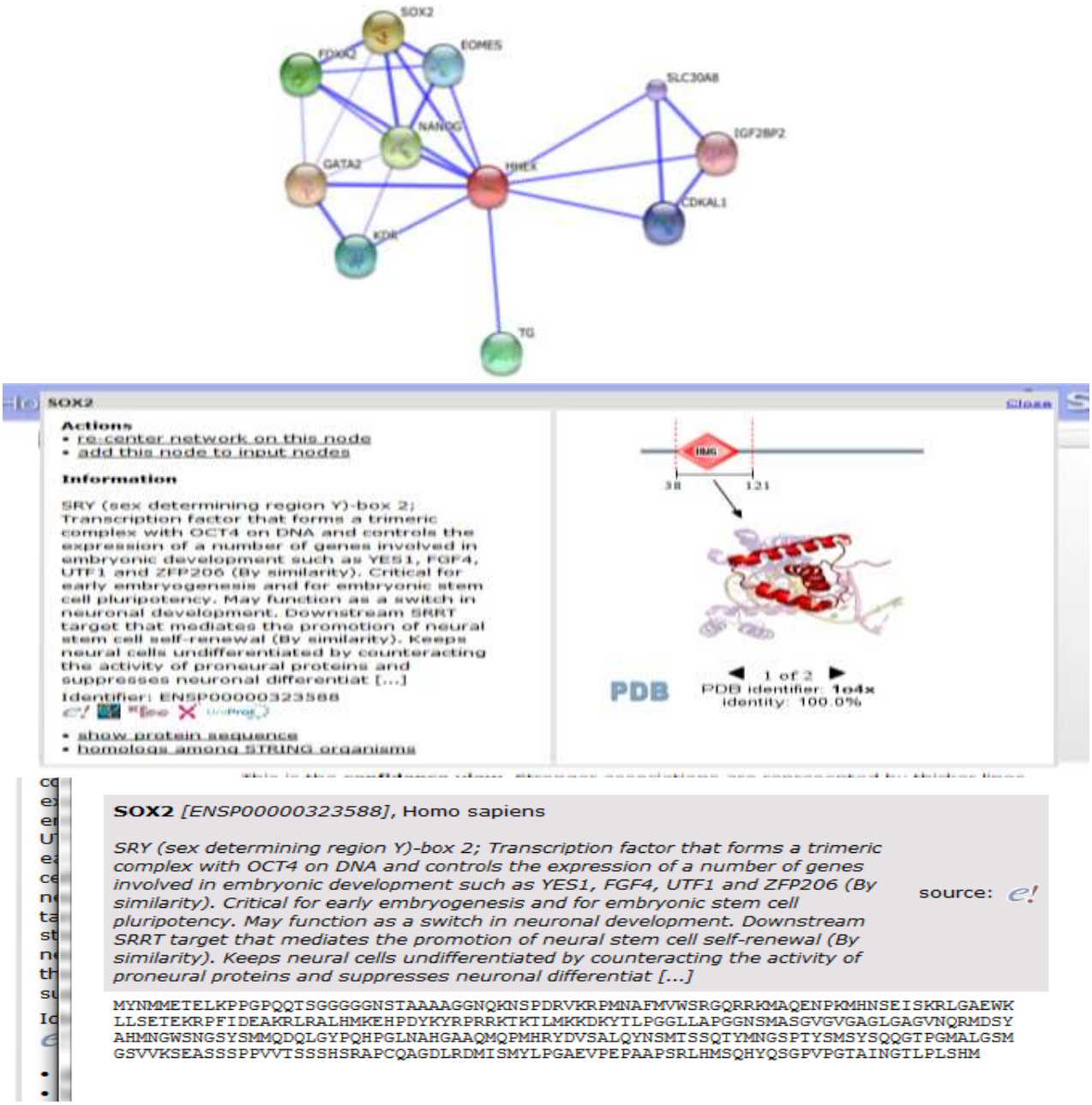
4-Another gene CDK L1 shows there are some many interacting partners are present but UCB protein shows highest intaction with CDKL1 and consider as a Hub protein and CDKL1 shows 98% identity with TCF7L2 but KCNJ11 shows only 49% identity

ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379



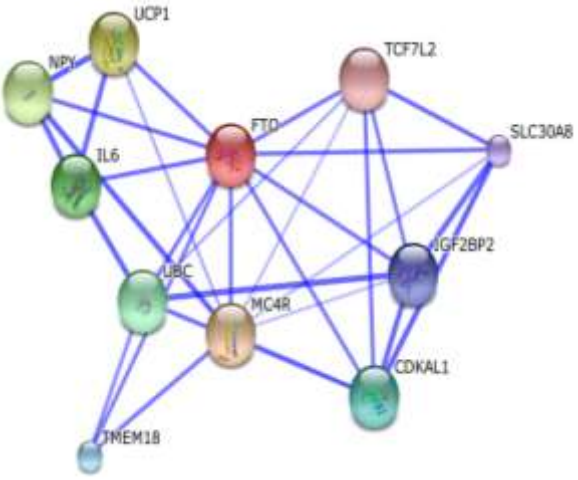
4-HHEXgene shows intraction with CDKL1,IGF2BP2,FOXA2,SOX2,TG,GATA2 and many other genee but it shows complete interaction with SOX2 Protein.

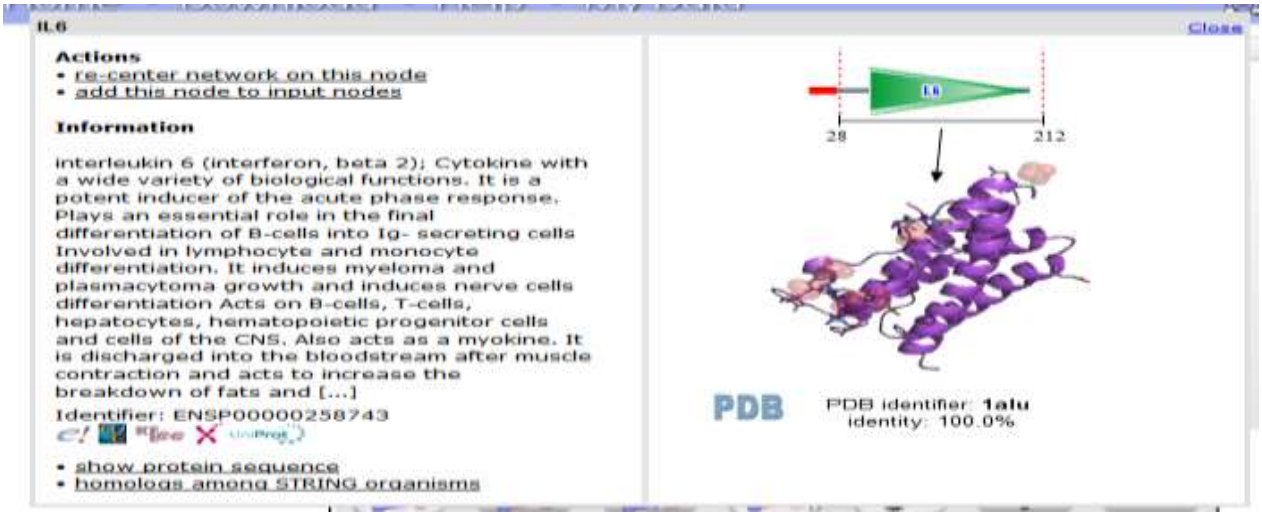
ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379



6-FTO gene(fat mass and obesity associated) show association with TCF7L2,CDKL1T2DM gene and also with SLC3018,IGF2BP2,MC4R,UBC,TMEM18,IL6 and others but there are strongest interaction with IL6 protein

ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379





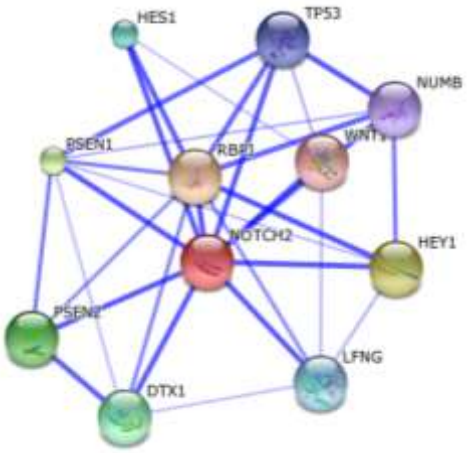
IL6 *[ENSP00000258743]*, Homo sapiens

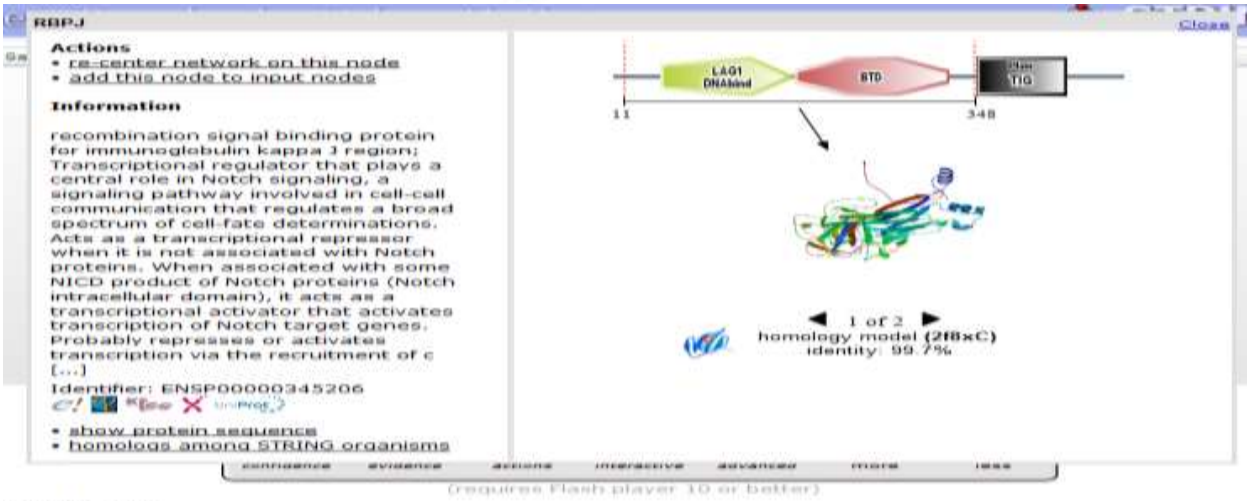
interleukin 6 (interferon, beta 2); Cytokine with a wide variety of biological functions. It is a potent inducer of the acute phase response. Plays an essential role in the final differentiation of B-cells into Ig- secreting cells Involved in lymphocyte and monocyte differentiation. It induces myeloma and plasmacytoma growth and induces nerve cells differentiation Acts on B-cells, T-cells, hepatocytes, hematopoietic progenitor cells and cells of the CNS. Also acts as a myokine. It is discharged into the bloodstream after muscle contraction and acts to increase the breakdown of fats and [...]

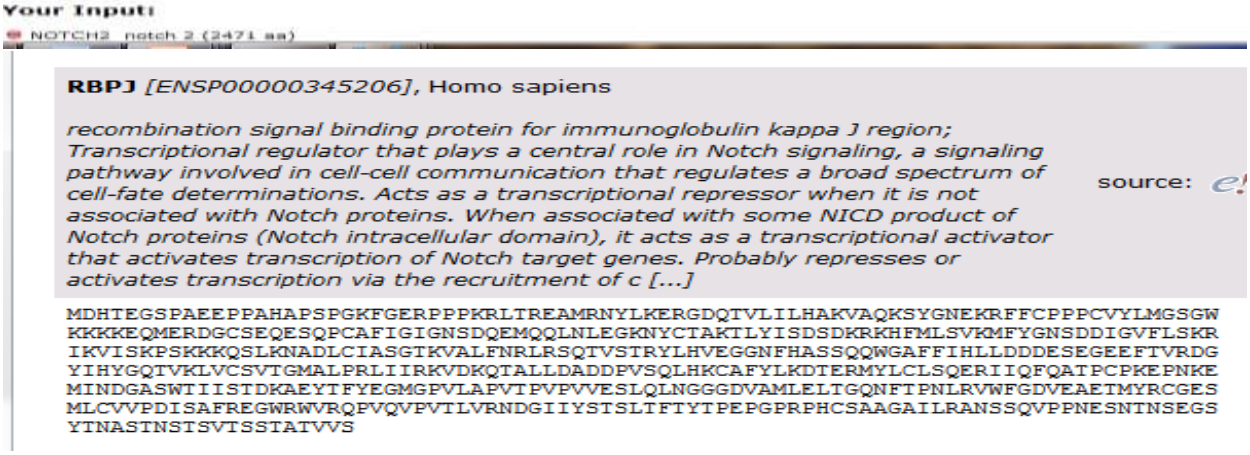
source:

```
MNSFSTSAFGPVAFSLGLLLVLPAAFPAPVPPGEDSKDVAAPHRQPLTSSERIDKQIRYILDGISALRKETCNKSNMCES
SKEALAENNLNLPKMAEKDGCFQSGFNEETCLVKIITGLLEFEVYLEYLQNRFESSEEQARAVQMSTKVLIQFLQKKAKN
LDAITTPDPTTNASLLTKLQAQNQWLQDMTTHLILRSFKEFLQSSLRALRQM
```

ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379







4. Conclusion

By analyzing each and every data her I have observed that that there are T2DM is a multifactorial and complex disease where several genes are involver that is near about 30 to 40 but here 10 to 15 are most frequent and found in most of the World population here I take 7 most frequent and recently studies genes and By pathways analysis I observed that there Are Gene are common in all frequently occurring pathways(by KEGG PATHWAY) like T2DM

ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379

pathways,WNT signaling Pathway and insulin signaling pathway. that are upstream and downstream regulated and involve in protein protein interaction so by taking that all important gene I have used STRING DATABASE and done protein protein interaction (PPI) of all 7 common genes which are KCNJ11, TCF7L2, PPARG, CDKL1, HHEX, NOTCH2, FTO After giving gene name I found that there are many more interacting partners of that single gene like (KCNJ11, TCF7L2, PPARG, CDKL1, HHEX, NOTCH, FTO) but there some shows least interaction some shows medium interaction but a single gene shows completely interaction with that protein and it act as a hub protein in PPIso there are showing 100% interaction with that single gene and act as a interacting partner of that gene and also showing full description of hub protein and their protein sequence also.So there are I have observed that KCNJ11. The future prospect is that we can discover and design drug by targeting the tolerant region of gene of T2DM and further analysis of genes are required and also their drug are required for diagnosis and prevention of disease(T2DM).

Acknowledgement

I am highly thankful to my supervisor Dr. Sunita Singh Department of Zoology MMV, BHU guidance, support and encouragement, inspiration always with us from initial stage to final stage of our project I find to myself blessed to work under her supervision. I would like to thanks Dr. Rajeev Mishra Coordinator whose always support me and encouraged me for doing better in life and in academics I am also thankful to our principle Dr. Sandhya kushik to giving opportunity and facility to the bioinformatics students. At last but not least I would like to express my gratitude to my parents for their never ending patience, everlasting love and ultimate support all of which have been guiding me for academic excellence and encouraging me to achieve higher goals. It's a privilege being supervised by for write paper on my thesis topic Honorable PhD adviser Dr. Prashant Ankur Jain, Assistant professor and In-charge, and Mr Ved Kumar Mishra (PhD Scholar) Department of Computational Biology and Bioinformatics, Jacob Institute of Biotechnology and Bioengineering, Sam Higginbottom Institute of Agriculture, Technology and Sciences (Deemed-to-be-university) Allahabad, and PhD co-adviser Dr. Raghvendra Raman Mishra, Assistant Professor, Department of Medical Lab Technology, Institute of Medical Sciences, Banaras Hindu University (BHU), Varanasi, for his selfless help, precise guidance, keen interest, uncased encouragement during the course of this research paper work and making desired suggestions to enable me to accomplish this task well in time. Working under him is a great experience in my part for which I shall be always grateful.

References

[1] Alessandro Brito dos Santos et al.''Basic bioinformatics tools for genotyping studies. A practical exercise with type 2 diabetes mellitus". International Research Journal of Biochemistry and Bioinformatics. .3 3.6 (2013) 115-124
[2] Cécile M Povel et al. "Single nucleotide polymorphisms (SNPs) involved in insulin resistance, weight regulation, lipid metabolism and inflammation in relation to metabolic syndrome: an epidemiological study". Cardiovascular Diabetology21.1(2012)11-33
[3] Christian von Mering1,2 et al."STRING 7—recent developments in the integration and prediction of protein interactions". Nucleic Acids Research.35.5(2010)13-19

ISSN- 2350-0530(O), ISSN- 2394-3629(P)
DOI: 10.5281/zenodo.836379

[4] Cristen J. Willer et al". Screening of 134 Single Nucleotide Polymorphisms (SNPs) Previously Associated With Type 2 Diabetes Replicates Association With 12 SNPs in Nine Genes". American Diabetes Association..56.2(2013)56-2712
[5] Cristian guja et al.(2012). "Ganetic factors involved in pathogenesis of type 2Diabetes mellitus".medicin.20.5(2012)44-61
[6] Ehab M. Abdella et al". Bioinformatics analysis of some functional genes and protein involved in obesity induced type 2 diabetes mellitus". American Journal of Bioinformatics, 1 .2(2012)87-98.
[7] Hyo-Jeong Ban et al". Identification of Type 2 Diabetes-associated combination of SNPs using Support Vector Machine". BMC Genetics, 11.2(2012)6-30.
[8] Hao Yu et al." Protein-interaction-network-based analysis for genome-wide association analysis of schizophrenia in Han Chinese population". Journal of Psychiatric Research.50.1 (2014) 73-78.
[9] Hiroyuki Ogata et al KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research,.27.1(2014)1-9
[10] Ine' s Barroso et al."Candidate Gene Association Study in Type 2 Diabetes Indicates a Role for Genes Involved in b-Cell Function as Well as Insulin Action".plos Biology.21.1(2005) 41-51
[11] Kyle J. Gaulton et al." Comprehensive Association Study of Type 2 Diabetes and Related Quantitative Traits With 222 Candidate Genes". .diabetes journals. 57.31(2010)36–3144.
[12] Priyanka Jain,E et al. " Systems Biology Approach Reveals Genome to Phenome Correlation in Type 2 Diabetes". A journal of clinical Investigation 8.1(2011)2-8.
[13] QiongWang et al." Targets and Candidate Agents for Type 2 Diabetes Treatment with Computational Bioinformatics Approach". Journal of Diabetes Research 3.1(2013)451-489
[14] Robert Wagner et al."Untangling the interplay of genetic and metabolic influences on beta-cell function Examplesof potential therapeutic implications involving TCF7L2 andFFAR1".molecular metabolism. 3.1(2014) 261–267.
[15] Simon D Rees et al ."Common variants of the TCF7L2 gene are associated with increased risk of type 2 diabetes mellitus in a UK-resident South Asian Population". BMC Medical Genetics.4.3(2010)41-58
[16] Srinubabu Gedela et al.".Identification of Biomarkers for Type 2 Diabetes and Its Complications.A Bioinformatic Approach". International journal of Biomedical science . 11.3(2007)55-78
[17] Suresh KG Shettigar et al ."Genetic Analysis of Type 2 Diabetes"Diabetes Care. 33.5 (2010) 62-69.
[18] Yanli Liu Zezhi et al.."Exploring the pathogenetic association between schizophrenia and type 2 diabetes mellitus diseases based on pathway analysis" Medical Genomics.64.1(2010)19-28.
[19] Uma Makheswari M et al."A Review on Bio Informatics for Diabetic Mellitus". International Journal of Pharma Sciences and Research.3.3(2008)89-395
[20] Wilfred Ip et al." The involvement of the wnt signaling pathway and TCF7L2 in diabetes mellitus The current understanding, dispute, and perspective". Cell & Bioscience. 2.2(2009) 8-40

*Corresponding author.
E-mail address: shrutimishra4january@ gmail.com