International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2019): 7.583

Analysis of Neuronal Differentiation 1 (NEUROD1) Gene and Mutations in Subphylum Vertebrata by using Bioinformatics Tools

Rodhan Patke

Department of Biotechnology, Thakur College of Science and Commerce, University of Mumbai, Maharashtra, India
rodhanp[at]rocketmail.com

Abstract: ***Chemically Neuronal differentiation 1 (NEUROD1) (C1741H2688N484O561S17) is a 356 amino acids protein sequence of NeuroD family that acts as an transcription factor, which forms heterodimers with basic helix-loop-helix protein and activates transcription of genes that contain an E-box (specific DNA sequence). It regulates expression of the insulin gene and mutations in this gene can result in Maturity-onset diabetes of the young (truncated NEUROD1) or Type II diabetes mellitus (arg111-to-leu). In this examination, bioinformatics tools were used for multiple sequence alignment of an aggregate of two hundred and ninety eight NEUROD1 protein sequences from vertebrata subphylum was done and they were studied by free to use bioinformatics softwares available on internet. The protein properties of the sequences (molecular mass, pI, signal peptide, transmembrane helices and conserved domains, secondary and 3D structures) were studied. The study of NEUROD1 protein sequences uncovered that there is high identity between the NEUROD1 present in vertebrata subphylum. PROCHECK tool was used to draw Ramachandran plot of the NEUROD1 protein (Gene ID: 4760) and the structure of the protein was studied. The study demonstrated that most of the residues of the protein sequence were situated in the most preferred areas in Ramachandran plot, showing that the simulated three-dimensional structure was authentic. Gene ontology and local synteny conservation of the NEUROD1 protein was studied. Evolutionary investigation demonstrated that there is a relationship between the NEUROD1 proteins in species of vertebrata subphylum under study. As indicated by the analysis, NEUROD1 ought to be derived from a common predecessor. For genomic analysis of mutations using RFLP Restriction mapping of enzymes were done and primer designing for NEUROD1 protein sequence was done. These predictions, however, need further work to validate reliability.***

Keywords: NEUROD1, bioinformatics, MODY, Type II diabetes mellitus

1. Introduction

Neuronal differentiation 1 (NEUROD1) (Gene ID- 4760), also termed as BETA2, plays an role in regulation of insulin by encoding a basic helix-loop-helix (bHLH) transcription factor that heterodimerizes with the ubiquitous bHLH protein E47 and binds to the E-box motif on insulin promoter. NEUROD1 is present on the long (q) arm of *Homo sapiens* chromosome 2 at position 31.3, between 181,668,295 base pairs to 181,680,665 base pairs.

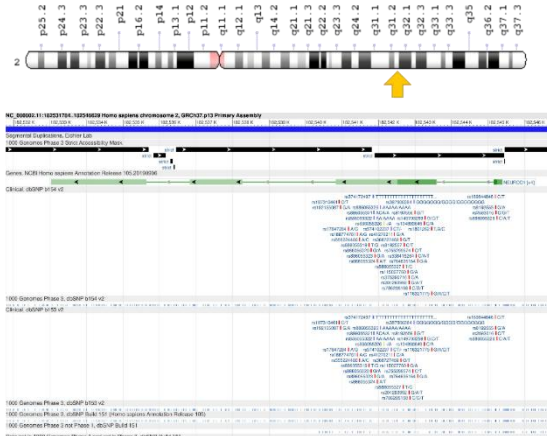


Figure 1: The Human NEUROD1 protein sequence was visualised on chromosome using Genome decoration page of NCBI (https://www.ncbi.nlm.nih.gov/genome/tools/gdp) and the Human NEUROD1 protein sequence was further analysed using 1000 genome browser of NCBI (https://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/)

In 1995, Naya et al studied NEUROD1 and discovered that NEUROD1 is transcriptional activator that binds to the E-box promoter sequence (5'-CANNTG-3') and initiates transcriptional activation. NEUROD1 plays an important role in regulation of various cellular pathways, such as initiation of pathway for formation of early retinal ganglion cells and inner ear sensory neurons. Regulation of amacrine cell fate specification is also regulated by NEUROD1, cerebellar cortex maintenance and morphogenesis of dendrites also involves important role of NEUROD1 protein.

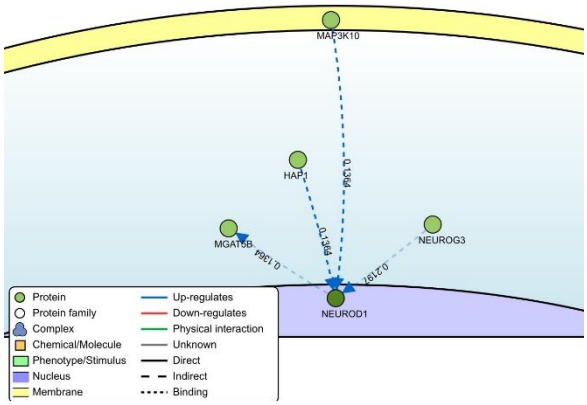


Figure 2: Signaling network of NEUROD1 protein visualized using SIGNOR (https://signor.uniroma2.it/)

Table 1: Relations of NEUROD1 protein in various pathways

Regulator	Mechanism	Target
HAP1	up-regulates activity by binding	NEUROD1
Two proteins that interact with ND, huntingtin-associated protein 1 (HAP1) and mixed-lineage kinase 2 (MLK2). Stimulation of NeuroD activity by huntingtin and huntingtin-associated proteins HAP1 and MLK2		
NEUROD1	up-regulates quantity by expression transcriptional regulation	MGAT5B
NeuroD1 and CTCF that bind to and activate the GnT-IX promoter, NeuroD1- and CTCF-dependent epigenetic mechanism governs brain-specific GnT-IX expression		
NEUROG3	up-regulates quantity by expression transcriptional regulation	NEUROD1
Ngn3 overexpression altered the expression of a number of regulatory genes, including NEUROD1		
MAP3K10	up-regulates activity by binding	NEUROD1

*Direct binding= Indirect Binding=

In 2004, Yan and Wang were the first to use oligonucleotides and small interfering RNA to study the expression and function of neuroD in retinas of chicken embryos. For studying neuroD, Yan and Wang used retrovirus to infect chick embryos, the infected embryos expressed an active repression construct which showed severe photoreceptor deficits. Anti-neuroD antibody which was specifically labelled to the nuclei of the outer nuclear layer was used as marker. The retinal cells were laminated and the results indicated that the outer nuclear retinal layer had become fragmented with areas that had low or no photoreceptor cells. Thus this indicated that repression of neuroD in chick embryos had affected gene related of photoreceptor.

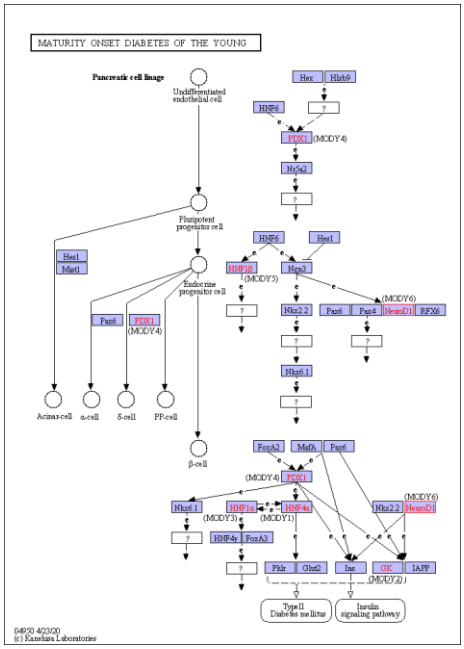


Figure 3: Pathway of Maturity Onset Diabetes of the young and Type II Diabetes mellitus studied using Kegg Pathway (https://www.genome.jp/kegg/pathway.html)

The data suggested a specific and essential role for neuroD in photoreceptor formation in the chick retina. In 2011 Pang et al experimentally discovered that foetal and postnatal human fibroblasts can be converted into induced neuronal cells by combining NEUROD1 protein with POU3F2, ASCL1, and MYT1L genes. Thus, Pang concluded that lineage-determining transcription factors can help in converting non-neuronal human somatic cells, and pluripotent stem cells into neurons.

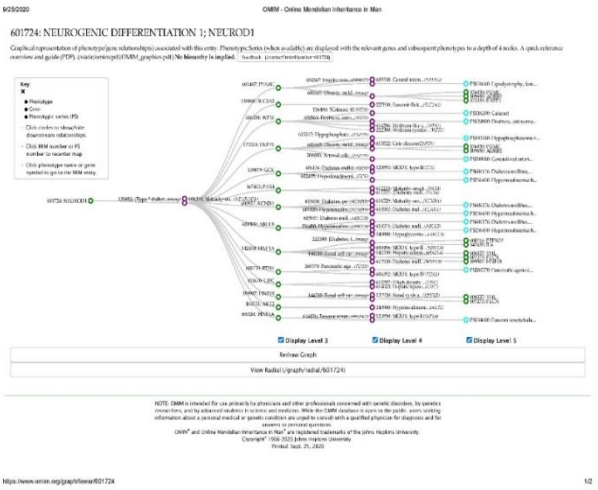


Figure 4: Graphical representation for relation of NEUROD1 protein to various disease associated with it, visualized using OMIM database.

In 1999 Malecki identified two 2 heterozygous mutations (Fig 3, 4) in NEUROD1 that lead to development of type II diabetes. First mutation was missense mutation (arg111-to-leu), which disrupted the DNA-binding domain and inhibited the NEUROD1 activity of binding to E-box. The second mutation, (206+C), gives rise to a shortened protein which lacked the C-terminal transactivation domain. Malecki discovered that patients with the second type of mutation (shortened NEUROD1 protein) had more severe and suggestive case of Maturity-onset diabetes of the young (MODY), whereas that patients with first type of mutation (arg111-to-leu mutation) had more typical of type II diabetes mellitus. In 2001 Fajans et al termed the diabetes due to mutation due to truncated NEUROD1 protein as maturity-onset diabetes of the young type 6. MODY (onset before 25 years of age), is an autosomal dominant inheritant defect, due to absence of β-cell autoimmunity, and sustained pancreatic β-cell function.

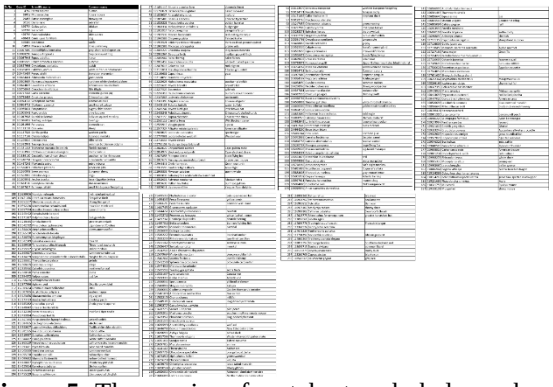


Figure 5: The species of vertebrata subphylum whose protein sequences were studied in this research retrieved from NCBI.

2. Materials and techniques

Free to use internet based bioinformatics tools and softwares were used for examinations of NEUROD1 protein in Vertebrata subphylum

Retrieval of NEUROD1protein sequence
Entirety of protein sequences of NEUROD1 of subphylum Vertebrata (Total Two hundred and ninety-eight protein sequences) were retrieved from protein database of NCBI: (http://www.ncbi.nlm.nih.gov; accessed on September 2020) as recorded in Fig-5.

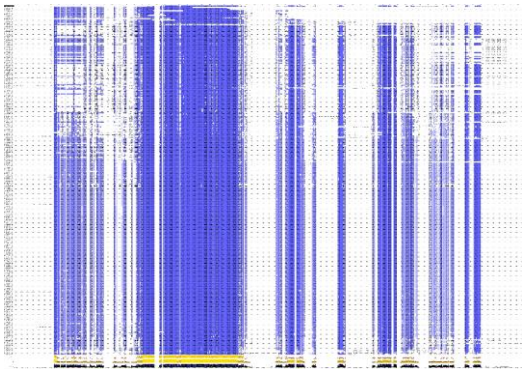


Figure 6: Multiple sequence alignment of the two hundred and ninety eight neuronal differentiation 1 (NEUROD1) protein sequences.(Mentioned in Fig 5)

Bioinformatics examinations

Characteristic features of NEUROD1 gene
SIGnaling Network Open Resource (SIGNOR) (https://signor.uniroma2.it/) tool was used to study the pathway of NEUROD1 at cellular level. Kegg Database was used for studying the role of NEUROD1 in MODY disease in humans (Fig 3).The NEUROD1 gene in human was visualized using 1000 genome browser tool available on NCBI. Online Mendelian Inheritance in Man (OMIM) (https://www.omim.org/) catalog was used to study the diseases caused by NEUROD1 mutation. ProtParam (https://web.expasy.org/protparam/) was used to predict physical and chemical properties of Human NEUROD1 protein. For predicting of presence and location of signal peptide cleavage sites in NEUROD1 protein sequences from different organisms SignalP tool (http://www.cbs.dtu.dk/services/SignalP/) was used.

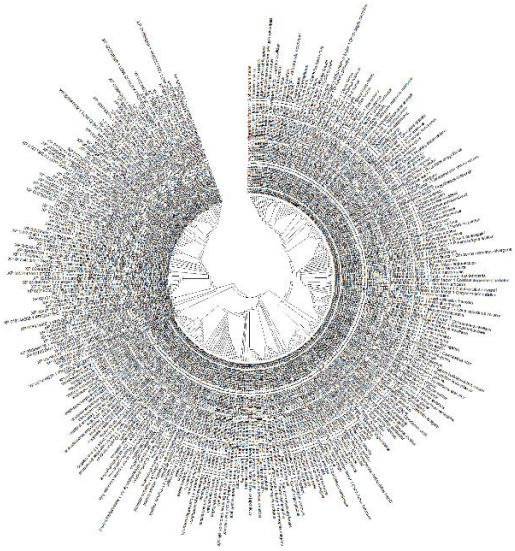


Figure 7: Phylogenetic tree of the two hundred and ninety eight sequences of NEUROD1 protein generated using Neighbour joining method by MEGA-X software.

Multiple sequence alignment
The homology and analogy of nucleotide and protein sequence of Neuronal differentiation 1 (NEUROD1) gene with the already reported sequences was checked through BLASTn and BLASTp(NCBI) and related or similar NEUROD1 gene sequences were aligned by using Clustal Omega utilizing CLUSTAL-W (with character counts) method with default setting.(Fig 6)

Phylogenetic analysis
OMA browser was used to study local synteny of the aligned genes.The desktop version of Jalview was usedfor the generation of phylogenetic tree from aligned sequences of NEUROD1 gene. NeighborJoining (NJ) strategy and blossum62 were used for planning the evolutionary tree (Fig 7). NCBI Tree viewer tool with NeighborJoining (NJ) method and blossum62 parameters were used to generate phylogenetic tree of the species with highest homology with human NEUROD1 (Fig 8).

Visualization of Motifs
The protein sequences of NEUROD1 were analyzedusing Multiple Em for Motif Elicitation tool (MEME; adaptation 5.1.1) for discovering the sequence specific motifs. (http://meme-suite.org/tools/meme). MEME analysis was done using default settings and the motifs to find was increased to 10 motif per protein sequence.(Fig 9, 10)

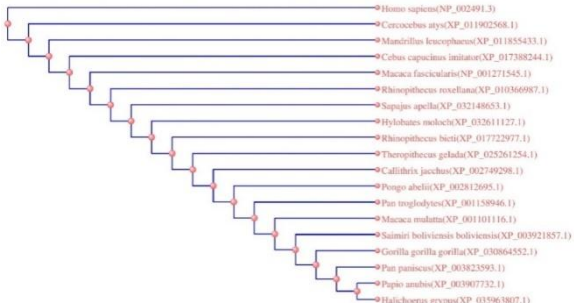


Figure 8: Phylogenetic tree of NEUROD1 protein sequences with highest identity with Human NEUROD1, generated by NCBI Tree viewer tool. (https://www.ncbi.nlm.nih.gov/blast/treeview/treeView.cgi)

Figure 9: Motifs found in the NEUROD1 protein sequence, 10 motif were found in NEUROD1 sequence, discovered using MEME suite tools (http://meme-suite.org/tools/meme).

Secondary structure prediction

A secondary structure prediction method based on the homologue method SOPMA (Self-Optimized Prediction Method with Alignment) (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) was used for human NEUROD1 protein. For prediction of transmembrane helices in proteins of Homo sapiens NEUROD1 protein, TMpred tool (https://embnet.vital-it.ch/software/TMPRED_form.html).

3-D Structure prediction

Sequences retrieval to predict 3D structure

NEUROD1 Protein from *Homo sapiens* (NP_002491.3) was used as a template to construct a 3D model.

Model building by homology modelling.

To predict 3-D structure of NEUROD1, homology modeling was used that is the most suitable method for building protein models. The SWISS-MODEL (Fig 14) Homology Modelling Report (https://swissmodel.expasy.org/) (PDB ID 2QL2) bioinformatics tool was used for generating fully automated protein structure homology-model of NEUROD1 protein. The Human NEUROD1 protein structure was studied for evolutionary relationships of protein domain using CATH Protein Structure Classification database (https://www.cathdb.info/).Rasmol software was used for 3D structure visualization. Relative Spatial arrangement of atoms (stereochemistry) of 3D model was studied using EMBL-EBI powered PROCHECK tool, (https://servicesn.mbi.ucla.edu/PROCHECK/) which helped in making the Ramachandran plot.

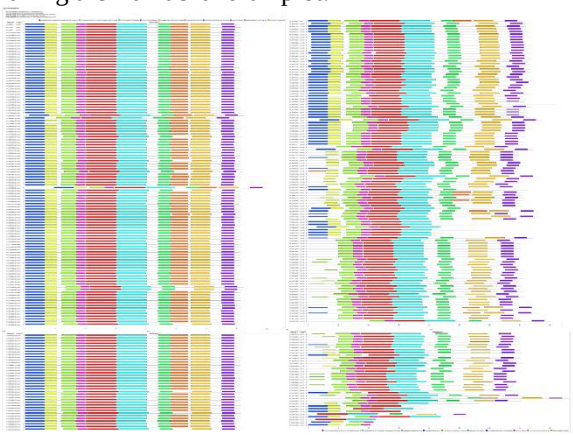


Figure 10: Motifs identified by MEME tool for NEUROD1 proteins. Differentcolored rectangles represent the sequence specific MEME motifs.

Restriction Mapping & Primer Designing

The restriction enzymes for NEUROD1 protein sequence in human were mapped using restriction mapper tool (http://www.restrictionmapper.org/). Primer designing for Human NEUROD1 nucleotide sequence was done using Primer-BLAST tool provided by NCBI (https://www.ncbi.nlm.nih.gov/tools/primer-blast/).

3. Results and Discussion

NEUROD1 are very important transcription factors required in regulation of insulin gene and mutation in these genes can cause MODY and Type II diabetes mellitus thus they play an important role in organisms. Hence to reveal the important factors about NEUROD1, I have analyzed and compared NEUORD1 gene sequence of Homo sapiens.

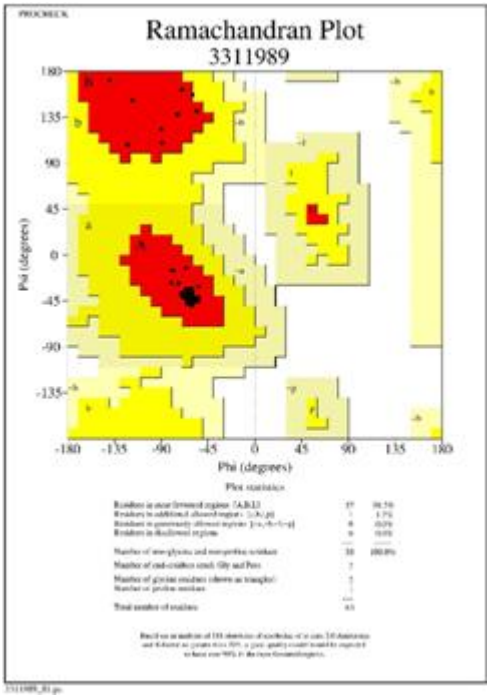


Figure 11: Ramchandran plot of the Human Neuronal differentiation 1 (NEUROD1) protein structure.

Diagnosis of NEUROD1 mutation by RFLP

For detection of NEUROD1 mutations in Homo sapiens by RFLP,online bioinformatics software and tools were used to find appropriate restriction enzymes that cleave the NEUROD1 protein and primers were also designed for amplification of NEUROD1 gene.I mapped restriction enzymes with sites in the protein sequence of NEUROD1 and also designed Primers for RFLP (Fig 14, 15). 63 Restriction enzymes were mapped which have restriction enzymes in the NEUROD1 protein sequence. Primer designing was performed and Primer pair with
5’ GACCCAGAAGCTGAGCAAGA 3’ and
3’ TCGTGGAAGATGGCGTTCAG 5’ and Product length of 669 base pairs was considered to be acceptable.

Alignment of NEUROD1 gene

Basic Local Alignment Search Tool (BLAST) the most accepted search algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to identify local regions of similarity and statistical

significance of nucleotide query sequence of NEUROD1.Multiple sequence alignment (fig 6) of two hundred and ninety-eight NEUROD1 protein sequences belonging to vertebrata subphylum (referenced in table 1), was done by Clustal Omega using CLUSTAL-W (with word counts) method and a similarity/identity matrices graph was generated (Fig 7). The study of multiple sequence alignments and similarity/identitygraph (Fig 7) demonstrated high homology in the protein sequences of vertebrate subphylum.

Table 2: Organisms with highest protein homology with Human NEUROD1 protein

Organism	Percentage similarity with Homo sapiens Neuronal differentiation 1 (NEUROD1) protein
Halichoerus grypus	99.72%
Papio anubis	99.72%
Pan paniscus	99.72%
Gorilla gorilla gorilla	99.72%
Saimiri boliviensis boliviensis	99.72%
Macaca mulatta	99.72%
Pan troglodytes	99.72%
Pongo abelii	99.72%
Callithrix jacchus	99.72%
Theropithecus gelada	99.72%
Rhinopithecus bieti	99.72%
Hylobates moloch	99.72%
Sapajus paella	99.72%
Rhinopithecus roxellana	99.72%
Macaca fascicularis	99.72%
Cebus capucinus imitator	99.72%
Mandrillus leucophaeus	99.72%
Cercocebus atys	99.72%

Phylogenetic analysis of NEUROD1 gene

For studying the evolutionary relationship, protein sequences of NEUROD1 protein in vertebrata subphylum were used and a phylogenetic tree was generated. The evolutionary tree was generated using the NeighbourJoining method. The analysis involved 298 protein sequences (Fig 5).The outcome demonstrated that NEUROD1 protein come from one precursor gene and formed into various branches. As indicated by the phylogenetic tree NEUROD1 protein have relationship with one another (Fig. 7).

Sequence Homology

The pairwise alignment of the samples under study showed that HumanNEUROD1 protein sequence showed 99.72% similarity with 18 protein sequences (Table 2) of vertebrata subphylum. A phylogenetic tree for these sequences was made to study the phylogenetic relationship between the sequences with most homology (Fig 8).

Examination of Sequence Specific Motifs

I further performed motif examinations of these sequences by Multiple EM for Motif Elicitation software for identification of motif specific in NEUROD1 protein of vertebrata subphylum. I found 10 conserved motifs (Fig 9, 10) in the Protein sequences of NEUROD1, which are shown in Table 3.

Table 3: Sequence specific motifs found in NEUROD1 protein sequences

Motif Name	E-value	Sites
RRMKANARERNRMHGLNAALDNLRKVVPCYSKTQKLSKIETLRLAKNYIW	3.5e-16050	295
ALSEILRSGKSPDLVSFVQTLCKGLSQPTTNLVAGCLQLNPRTFLPEQNQ	1.4e-14879	293
EEDEDLEEEEEEEEEDDDQKPKRRG	3.6e-6676	224
CEIPIDNJMSFDSHSHHERVM	1.8e-4615	268
PPLSINGNFSFKHEPSAEFEKNYAFTMHYPAA	2.7e-6565	221
YQSPGLPSPPYGTMDSSHVFH	1.0e-5204	280
MTKSYSESGLMGEPQPQGPPSWTDECLSSQDE	6.2e-6150	209
PKKKKMTKARLERFK	1.5e-3868	289
VKPPPHAYSAALEPFFESPLTDCTSPSFD	6.6e-3732	141
EHEADKKEDDLEAMNAEEDSL	1.6e-3043	210

By study of the results of Table 3, motif 1-RRMKANARERNRMHGLNAALDNLRKVVPCYSKTQKLSKIETLRLAKNYIW was found to be conserved more than other 9 motifs of vertebrata subphylum.By further analysis and characterization of the NEUROD1 protein using pfam tool (https://pfam.xfam.org/), it was found out that the motif – RRMKANARERNRMHGLNAALDNLRKVVPCYSKTQKLSKIETLRLAKNYIW is a domain for (Basic Helix-loop-helix) bHLH. bHLH is among the largest family of dimerizing transcriptional factor, they are very important factors required for cell development.

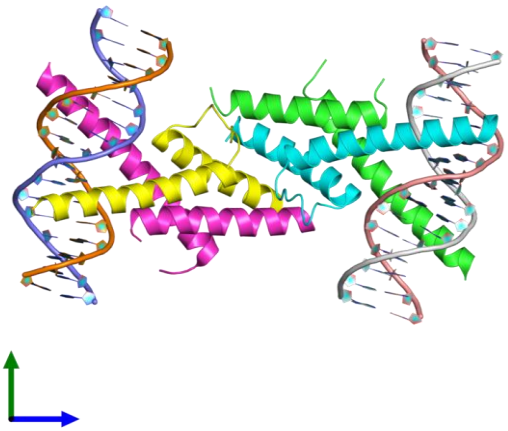


Figure 12: 3D model of the Neuronal differentiation 1 (NEUROD1) protein (colored by chain and viewed from the front.) by PDBe (https://www.ebi.ac.uk/pdbe/entry/pdb/2ql2)the protein contains:

2 copies of Transcription factor E2-alpha,2 copies of Neurogenic differentiation factor 1, 2 copies of DNA (5'-D(*DTP*DAP*DGP*DGP*DCP*DCP*DAP*DTP*DCP*DTP*DGP*DGP*DTP*DCP*DCP*DT)-3'), 2 copies of DNA (5'-D(*DAP*DGP*DGP*DAP*DCP*DCP*DAP*DGP*DAP*DTP*DGP*DGP*DCP*DCP*DTP*DA)-3')

4. Structural and Characteristic Analysis

Table 4: Chemical properties of NEUROD1 protein analyzed using ExPASy-ProtParam tool

Sr.No	Chemical Property	Value
1	Number of amino acids	356
2	Molecular weight	39920.42
3	Theoretical pI	5.19
4	Total number of negatively charged residues (Asp + Glu)	57
5	Total number of positively charged residues (Arg + Lys)	38
6	Formula	$C_{1741}H_{2688}N_{484}O_{561}S_{17}$
7	Estimated half-life((mammalian reticulocytes, in vitro)	30 hours
8	Instability index	74.57
9	Aliphatic index	57.08
10	Grand average of hydropathicity (GRAVY)	-0.839

Chemical properties of Human NEUROD1 are shown in the table 4. On the amino acid level, Human NEUROD1 protein is similar to numerous different NEUROD1 proteins present in vertebrata subphylum under study. For prediction of the presence of signal peptides and the location of their cleavage sites in proteins sequence, NEUROD1 protein was further examined by SignalP tool (http://www.cbs.dtu.dk/administrations/SignalP/) and the results predicted that there is 0.002 probability of presence of Sec signal peptide (Sec/SPI) signal peptide protein sequence.

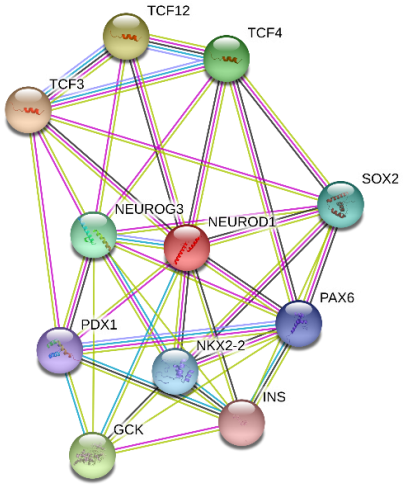


Figure 13: Predicted Protein-protein interaction networks of NEUROD1 with other proteins generated using STRING Database.

NEUROD1 protein was examined using TMpred tool (http://www.cbs.dtu.dk/administrations/TMHMM) and the results indicated presence of 1 transmembrane helice, which is inside to outside helices, the helice is between acid amine 298 to 551.SOPMA examination of Human NEUROD1 protein was performed for studying the secondary structure of NEUROD1 protein sequence and the results stated that the peptide of Human Neuronal differentiation 1 (NEUROD1) protein had 39.89% of alpha helices, 3.37% of beta turns, 3.97% of extended strands, and 52.81 % of random coil. SOPMA results indicate that extended strand had values lower than alpha helices and random coil in Human NEUROD1 protein. PROCHECK tool was used to study the quality of stereochemistry of Human NEUROD1 protein by generating the Ramachandran plot (Fig. 11), which suggested Human NEUROD1protein had 98.3% residues in the most favoredregion, thus stating that it has good and stable stereochemistry.

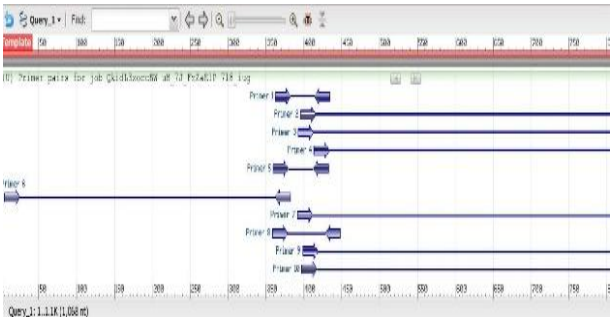


Figure 14: Graphical representation of Primer pairs generated through Primer-BLAST NCBI database.

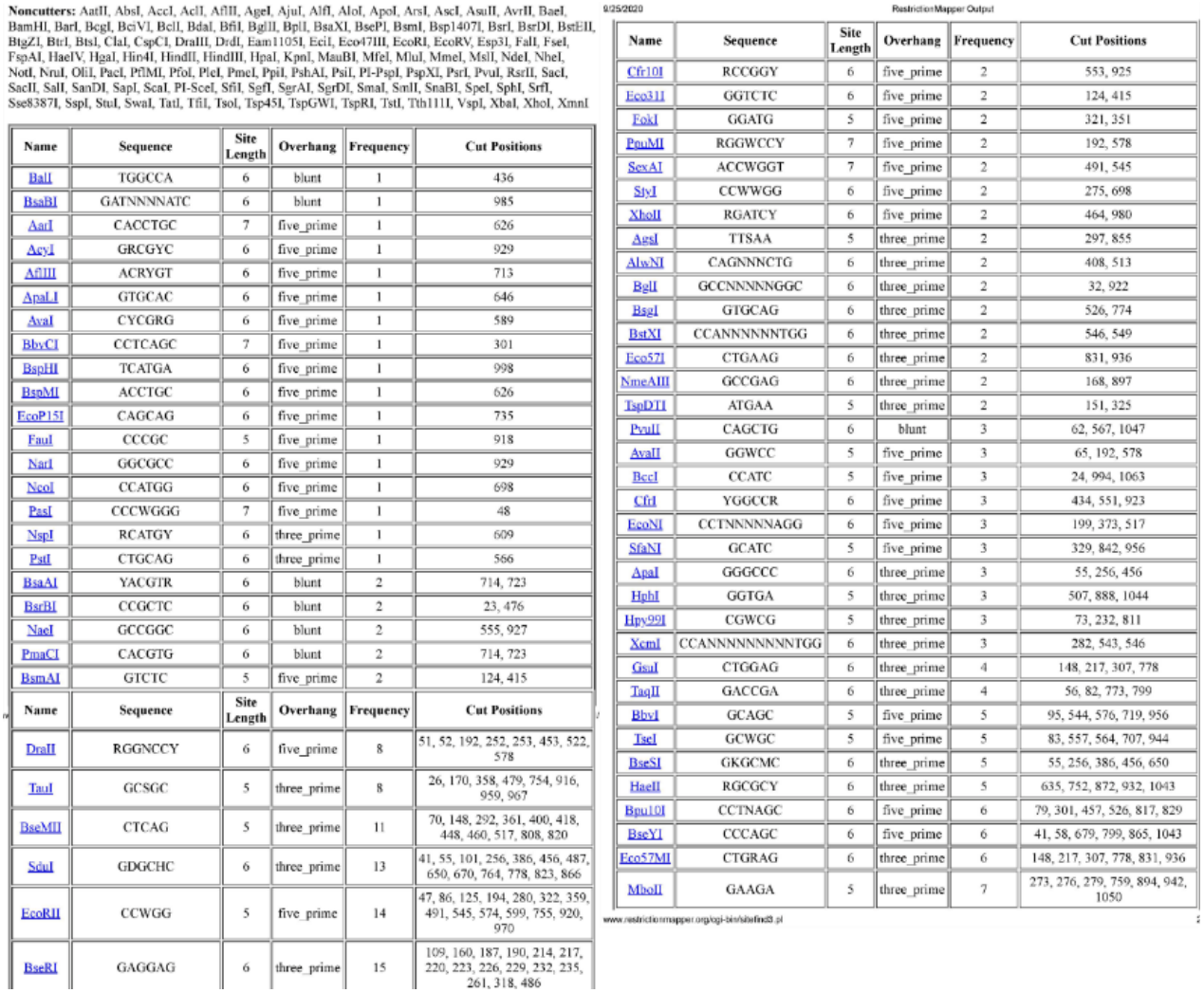
Homology modeling

SWISS-MODEL Homology Modelling Report tool was used (https://swissmodel.expasy.org/) for prediction of 3D structure of Human NEUROD1 protein, and the predicted model was visualized using PROCHECK and PDBe (https://www.ebi.ac.uk/pdbe/entry/pdb/2ql2) (Fig 12).

E47 homo- and heterodimers of the peotein structure are stabilized by two bonds of hydrogen and due to favorable interactions during packing. Instability of class II homodimers favors heterodimerization.

DNA specific contacts by E47 are possible due to orientation of E47-NeuroD1 heterodimer on DNA.

Argb10 contiguity to purine base N7 opposite the pyrimidine in position 3 is possible because of binding of CA(C/T) half sites toE47.NeuroD1 is prevented from forming this DNA contact by HisH1.1 blocks the shift in phosphate backbone binding, which inhibits the NEUROD1 to form DNA contact. Without HisH1.1 Argb10 would bind the DNA in a specific conformation. E47-NeuroD1 heterodimer should be oriented, as it may be beneficial for regulating other transcription factors, initiating formation of an active pre-initiation complex. Further studies of bHLH mutations and complexes with interacting transcription factors are required to know how the bHLH domain is once it binds to the E-box.



Noncutters: AatII, AbsI, AccI, AclI, AflII, AgeI, AjuI, AlfI, AloI, ApoI, ArsI, AscI, AsuII, AvrII, BaeI, BamHI, BarI, BcgI, BciVI, BclI, BdaI, BfiI, BglII, BplI, BsaXI, BsePI, BsmI, Bsp1407I, BsrI, BsrDI, BstEII, BtgZI, BtrI, BtsI, ClaI, CspCI, DraIII, DrdI, Eam1105I, EciI, Eco47III, EcoRI, EcoRV, Esp3I, FalI, FseI, FspAI, HaeIV, HgaI, Hin4I, HindII, HindIII, HpaI, KpnI, MauBI, MfeI, MluI, MmeI, MslI, NdeI, NheI, NotI, NruI, OliI, PacI, PflMI, PfoI, PleI, PmeI, PpiI, PshAI, PsiI, PI-PspI, PspXI, PsrI, PvuI, RsrII, SacI, SacII, SalI, SanDI, SapI, ScaI, PI-SceI, SfiI, SgfI, SgrAI, SgrDI, SmaI, SmlI, SnaBI, SpeI, SphI, SrfI, Sse8387I, SspI, StuI, SwaI, TatI, TfiI, TsoI, Tsp45I, TspGWI, TspRI, TstI, Tth111I, VspI, XbaI, XhoI, XmnI

Name	Sequence	Site Length	Overhang	Frequency	Cut Positions
BalI	TGGCCA	6	blunt	1	436
BsaBI	GATNNNNATC	6	blunt	1	985
AarI	CACCTGC	7	five_prime	1	626
AcyI	GRCGYC	6	five_prime	1	929
AflIII	ACRYGT	6	five_prime	1	713
ApaLI	GTGCAC	6	five_prime	1	646
AvaI	CYCGRG	6	five_prime	1	589
BbvCI	CCTCAGC	7	five_prime	1	301
BspHI	TCATGA	6	five_prime	1	998
BspMI	ACCTGC	6	five_prime	1	626
EcoP15I	CAGCAG	6	five_prime	1	735
FauI	CCCGC	5	five_prime	1	918
NarI	GGCGCC	6	five_prime	1	929
NcoI	CCATGG	6	five_prime	1	698
PasI	CCCWGGG	7	five_prime	1	48
NspI	RCATGY	6	three_prime	1	609
PstI	CTGCAG	6	three_prime	1	566
BsaAI	YACGTR	6	blunt	2	714, 723
BsrBI	CCGCTC	6	blunt	2	23, 476
NaeI	GCCGGC	6	blunt	2	555, 927
PmaCI	CACGTG	6	blunt	2	714, 723
BsmAI	GTCTC	5	five_prime	2	124, 415
DraII	RGGNCCY	6	five_prime	8	51, 52, 192, 252, 253, 453, 522, 578
TauI	GCSGC	5	three_prime	8	26, 170, 358, 479, 754, 916, 959, 967
BseMII	CTCAG	5	three_prime	11	70, 148, 292, 361, 400, 418, 448, 460, 517, 808, 820
ScuI	GDGCHC	6	three_prime	13	41, 55, 101, 256, 386, 456, 487, 650, 670, 764, 778, 823, 866
EcoRII	CCWGG	5	five_prime	14	47, 86, 125, 194, 280, 322, 359, 491, 545, 574, 599, 755, 920, 970
BseRI	GAGGAG	6	three_prime	15	109, 160, 187, 190, 214, 217, 220, 223, 226, 229, 232, 235, 261, 318, 486

9/25/2020 RestrictionMapper Output

Name	Sequence	Site Length	Overhang	Frequency	Cut Positions
Cfr10I	RCCGGY	6	five_prime	2	553, 925
Eco31I	GGTCTC	6	five_prime	2	124, 415
FokI	GGATG	5	five_prime	2	321, 351
PpuMI	RGGWCCY	7	five_prime	2	192, 578
SexAI	ACCWGGT	7	five_prime	2	491, 545
StyI	CCWWGG	6	five_prime	2	275, 698
XhoII	RGATCY	6	five_prime	2	464, 980
AgsI	TTSAA	5	three_prime	2	297, 855
AlwNI	CAGNNNCTG	6	three_prime	2	408, 513
BglI	GCCNNNNNGGC	6	three_prime	2	32, 922
BsgI	GTGCAG	6	three_prime	2	526, 774
BstXI	CCANNNNNNTGG	6	three_prime	2	546, 549
Eco57I	CTGAAG	6	three_prime	2	831, 936
NmeAIII	GCCGAG	6	three_prime	2	168, 897
TspDTI	ATGAA	5	three_prime	2	151, 325
PvuII	CAGCTG	6	blunt	3	62, 567, 1047
AvaII	GGWCC	5	five_prime	3	65, 192, 578
BccI	CCATC	5	five_prime	3	24, 994, 1063
CfrI	YGGCCR	6	five_prime	3	434, 551, 923
EcoNI	CCTNNNNNAGG	6	five_prime	3	199, 373, 517
SfaNI	GCATC	5	five_prime	3	329, 842, 956
ApaI	GGGCCC	6	three_prime	3	55, 256, 456
HghI	GGTGA	5	three_prime	3	507, 888, 1044
Hpy99I	CGWCG	5	three_prime	3	73, 232, 811
XcmI	CCANNNNNNNNNTGG	6	three_prime	3	282, 543, 546
GsuI	CTGGAG	6	three_prime	4	148, 217, 307, 778
TaqII	GACCGA	6	three_prime	4	56, 82, 773, 799
BbvI	GCAGC	5	five_prime	5	95, 544, 576, 719, 956
TseI	GCWGC	5	five_prime	5	83, 557, 564, 707, 944
BseSI	GKGCMC	6	three_prime	5	55, 256, 386, 456, 650
HaeII	RGCGCY	6	three_prime	5	635, 752, 872, 932, 1043
Bpu10I	CCTNAGC	6	five_prime	6	79, 301, 457, 526, 817, 829
BseYI	CCCAGC	6	five_prime	6	41, 58, 679, 799, 865, 1043
Eco57MI	CTGRAG	6	three_prime	6	148, 217, 307, 778, 831, 936
MboII	GAAGA	5	three_prime	7	273, 276, 279, 759, 894, 942, 1050

www.restrictionmapper.org/cgi-bin/sitefind3.pl 2

Figure 15: Restriction enzymes that have restriction sites in the NEUROD1 protein sequence. Restriction Mapper tool was used to map the restriction enzymes

Prediction of protein interactions

STRING database (https://string-db.org/) (Fig 13) was used to predict physical and functional protein-protein association networks of human NEUROD1 protein with other proteins of the family. The neural networks (fig 13) showed the interaction of NEUROD1 protein with other proteins of Hominidae family. There are different lines which have the results of curated databases, experimentally determined data, gene neighbourhood, gene fusions, gene co-occurrence, text mining, co-expression, and protein homology. Small nodes presented that they have low interaction with NEUROD1 protein sequence and large nodes showed high interaction with query protein. Protein-protein interactions are very important in various biological processes including cell development and communication, gene expression, metabolism and immune responses. CATH databasewas used to study the evolutionary relationships of protein domains and it was found that NEUROD1 belongs to super family of helix-loop-helix DNA binding domain.

5. Conclusion

Studying of NEUROD1 gene can help in understanding onset of type II diabetes mellitus and Maturity-onset diabetes of the young, study of NEUROD1 can help in preimplantation genetic diagnosis of the NEUROD1 mutations. In this investigation I have effectively annotated the full sequence of NEUROD1 protein and uncovered data about protein-protein association. The examination additionally gave insights about physicochemical nature of NEUROD1, secondary structure of NEUROD1 was predicted through different bioinformatics tools, and examination of the consensus sequence in NEUROD1 was done. Gene ontology and sequence specific motif study was also done and a phylogenetic investigation of NEUROD1 protein was executed.

The knowledge of bioinformatics is beneficial to understand the nature of pathways associated with Neuronal differentiation 1 (NEUROD1), viz. Regulation of gene expression in beta cells and regulation of gene expression in endocrine-committed (NEUROG3+) progenitor cells have been studied and the proteins associated with it are identified. The characterization of proteins have helped in biological understanding of MODY6 and NIDDM.

6. Acknowledgement

I would like to thank my Head of the Department (B.Sc. Biotechnology) Dr. (Mrs.) Aparna Deshmukh for giving me exposure to the field of bioinformatics and for generating interest for research. I would also like to thank Thakur

College of Science and Commerce for always supporting me in such extra-curricular activities.

7. Competing Interests Disclaimer

Authors have declared that no competing interests exist. The products used for this research are commonly and predominantly use products in our area of research and country. There is absolutely no conflict of interest between the authors and producers of the products because we do not intend to use these products as an avenue for any litigation but for the advancement of knowledge. Also, the research was not funded by the producing company rather it was funded by personal efforts of the authors.

References

[1] Fajans, S. S., Bell, G. I., Polonsky, K. S. Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. New Eng. J. Med. 345: 971-980, 2001. [PubMed: 11575290, related citations] [Full Text]

[2] Malecki, M. T., Jhala, U. S., Antonellis, A., Fields, L., Doria, A., Orban, T., Saad, M., Warram, J. H., Montminy, M., Krolewski, A. S. Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus. Nature Genet. 23: 323-328, 1999. [PubMed: 10545951, related citations] [Full Text]

[3] Fajans, S. S., Bell, G. I., Polonsky, K. S. Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. New Eng. J. Med. 345: 971-980, 2001. [PubMed: 11575290, related citations] [Full Text]

[4] Lee, J. E., Hollenberg, S. M., Snider, L., Turner, D. L., Lipnick, N., Weintraub, H. Conversion of Xenopus ectoderm into neurons by NeuroD, a basic helix-loop-helix protein. Science 268: 836-844, 1995. [PubMed: 7754368, related citations] [Full Text]

[5] Liu, M., Pereira, F. A., Price, S. D., Chu, M., Shope, C., Himes, D., Eatock, R. A., Brownell, W. E., Lysakowski, A., Tsai, M.-J. Essential role of BETA2/NeuroD1 in development of the vestibular and auditory systems. Genes Dev. 14: 2839-2854, 2000. [PubMed: 11090132, images, related citations] [Full Text]

[6] Liu, M., Pleasure, S. J., Collins, A. E., Noebels, J. L., Naya, F. J., Tsai, M.-J., Lowenstein, D. H. Loss of BETA2/NeuroD leads to malformation of the dentate gyrus and epilepsy. Proc. Nat. Acad. Sci. 97: 865-870, 2000. Note: Erratum: Proc. Nat. Acad. Sci. 97: 5679 only, 2000. [PubMed: 10639171, images, related citations] [Full Text]

[7] Malecki, M. T., Jhala, U. S., Antonellis, A., Fields, L., Doria, A., Orban, T., Saad, M., Warram, J. H., Montminy, M., Krolewski, A. S. Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus. Nature Genet. 23: 323-328, 1999. [PubMed: 10545951, related citations] [Full Text]

[8] Naya, F. J., Huang, H.-P., Qiu, Y., Mutoh, H., DeMayo, F. J., Leiter, A. B., Tsai, M.-J. Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/neurod-deficient mice. Genes Dev. 11: 2323-2334, 1997. [PubMed: 9308961, images, related citations] [Full Text]

[9] Naya, F. J., Stellrecht, C. M., Tsai, M. J. Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix transcription factor. Genes Dev. 9: 1009-1019, 1995. [PubMed: 7774807, related citations] [Full Text]

[10] Ochocinska, M. J., Munoz, E. M., Veleri, S., Weller, J. L., Coon, S. L., Pozdeyev, N., Iuvone, P. M., Goebbels, S., Furukawa, T., Klein, D. C. NeuroD1 is required for survival of photoreceptors but not pinealocytes: results from targeted gene deletion studies. J. Neurochem. 123: 44-59, 2012. [PubMed: 22784109, images, related citations] [Full Text]

[11] Pang, Z. P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D. R., Yang, T. Q., Citri, A., Sebastiano, V., Marro, S., Sudhof, T. C., Wernig, M. Induction of human neuronal cells by defined transcription factors. Nature 476: 220-223, 2011. [PubMed: 21617644, images, related citations] [Full Text]

[12] Pennesi, M. E., Cho, J.-H., Yang, Z., Wu, S. H., Zhang, J., Wu, S. M., Tsai, M.-J. BETA2/NeuroD1 null mice: a new model for transcription factor-dependent photoreceptor degeneration. J. Neurosci. 23: 453-461, 2003. [PubMed: 12533605, related citations] [Full Text]

[13] Tamimi, R., Steingrimsson, E., Copeland, N. G., Dyer-Montgomery, K., Lee, J. E., Hernandez, R., Jenkins, N. A., Tapscott, S. J. The NEUROD gene maps to human chromosome 2q32 and mouse chromosome 2. Genomics 34: 418-421, 1996. [PubMed: 8786144, related citations] [Full Text]

[14] Wang, F., Li, H., Xu, M., Li, H., Zhao, L., Yang, L., Zaneveld, J. E., Wang, K., Li, Y., Sui, R., Chen, R. A homozygous missense mutation in NEUROD1 is associated with nonsyndromic autosomal recessive retinitis pigmentosa. Invest. Ophthal. Vis. Sci. 56: 150-155, 2015.

[15] Yan, R.-T., Wang, S.-Z. Requirement of neuroD for photoreceptor formation in the chick retina. Invest. Ophthal. Vis. Sci. 45: 48-58, 2004. [PubMed: 14691153, images, related citations] [Full Text]

[16] Kuwabara, T., Hsieh, J., Muotri, A. et al. Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. Nat Neurosci 12, 1097–1105 (2009). https://doi.org/10.1038/nn.2360

[17] Malecki, M., Jhala, U., Antonellis, A. *et al.* Mutations in *NEUROD1* are associated with the development of type 2 diabetes mellitus. *Nat Genet* 23, 323–328 (1999). https://doi.org/10.1038/15500

[18] Gao, Z., Ure, K., Ables, J. *et al. Neurod1* is essential for the survival and maturation of adult-born neurons. *Nat Neurosci* 12, 1090–1092 (2009). https://doi.org/10.1038/nn.2385

[19] NeuroD1/beta2 contributes to cell-specific transcription of the proopiomelanocortin gene.

[20] G Poulin, B Turgeon, J Drouin Molecular and Cellular Biology Nov 1997, 17 (11) 6673-6682; DOI: 10.1128/MCB.17.11.6673

[21] NeuroD1 induces terminal neuronal differentiation in olfactory neurogenesis

[22] Camille Boutin, Olaf Hardt, Antoine de Chevigny, Nathalie Coré, Sandra Goebbels, Ralph Seidenfaden, A ndreas Bosio, Harold Cremer
[23] Proceedings of the National Academy of Sciences Jan 2010, 107 (3) 1201-1206; DOI: 10.1073/pnas.0909015107
[24] Neurogenin2 Directs Granule Neuroblast Production and Amplification while NeuroD1 Specifies Neuronal Fate during Hippocampal Neurogenesis Roybon L, Hjalt T, Stott S, Guillemot F, Li JY, et al. (2009) Neurogenin2 Directs Granule Neuroblast Production and Amplification while NeuroD1 Specifies Neuronal Fate during Hippocampal Neurogenesis. PLOS ONE 4(3): e4779. https://doi.org/10.1371/journal.pone.0004779
[25] NeuroD1 reprograms chromatin and transcriptionfactor landscapes to induce the neuron al programAbhijeet Pataskar1,†, Johannes Jung1,†, Pawel Smialowski2, Florian Noack3, Federico Calegari3,Tobias Straub2& Vijay K Tiwari https://doi.org/10.15252/embj.201591206
[26] Essential role of BETA2/NeuroD1 in development of the vestibular and auditory systems Min Liu,1 Fred A. Pereira,1 Steven D. Price,2 Mei-jin Chu,1 Cindy Shope,3 Donna Himes,3 Ruth Anne Eatock,3 William E. Brownell,3 Anna Lysakowski,2 and Ming-Jer Tsai1,4,5 doi:10.1101/gad.840500
[27] Sauve, S., Naud, J. F., and Lavigne, P. (2007) The mechanism of discrimination between cognate and non-specific DNA by dimeric b/HLH/LZ transcription factors, J. Mol. Biol. 365, 1163-1175.
[28] Longo, A., Guanga, G. P., & Rose, R. B. (2008). *Crystal Structure of E47−NeuroD1/Beta2 bHLH Domain−DNA Complex: Heterodimer Selectivity and DNA Recognition†,‡. Biochemistry, 47(1), 218–229.* doi:10.1021/bi701527r