Journal of Next Generation Sequencing & Applications
ISSN: 2469-9853

Agrawal et al., Next Generat Sequenc & Applic
2017, 4:2
DOI: 10.4172/2469-9853.1000150

Research Article Open Access

Identification of human genes and its genomics functions via miRNAs of *C. elegans* on bioinformatics platforms

Shashi Prabha Agrawal[1*], Mrityunjay Kumar Sinha[2] and Rajnish Kumar[2]

[1*]*Assistant Professor, Department of Zoology, D.A.V College, Civil Lines Kanpur, India*

[2]*Associate Professor, Department of Zoology, D.A.V College, Civil Lines Kanpur, India*

[*]**Corresponding author:** Shashi Prabha Agrawal, Assistant Professor, Department of Zoology, D.A.V College, Civil Lines Kanpur, India; Tel: +919335014995; E-mail: shashiprabhabioinfo@gmail.com

Rec date: August 28, 2017; **Acc date:** September 20, 2017; **Pub date:** September 22, 2017

Abstract

Very recently microRNAs (miRNAs) have emerged as an important elements of gene regulatory networks. In present study, we use *Caenorhabditis elegans* as a model for identification of targeted genes. In this study, we select 113 precursors and 26 miRNAs retrieves from miRBase and these 26 miRNAs target 610 genes in human. In this prediction, we found exact name of potential genes with their exact name and functions. These potential genes regulate various genomic function with oncogenic transformation. The disease regulated by these potential genes are myelogenous leukemia, Alzheimer's disease, spinal muscular atrophy, poly cystic kidney disease, muscular dystrophyLoeys-Dietz aortic aneurysm syndrome (LDAS) adermatoglyphia, Parkinson disease, lymphoblastic leukemia, hypervalinemia, hyperleucine-isoleucinemia, giant axonal neuropathy (GAN), Turner's syndrome, Klinefelter's syndrome, Triple X syndrome, Color blindness, hemophilia, and Duchenne muscular dystrophy etc.

Keywords: miRNAs; Genes; *Caenorhabditis elegans*; Precursors sequences; Muscular dystrophy

Introduction

MicroRNA is a family of small non-coding RNAs that regulate gene expression in a sequence-specific manner [1] miRNAs are a class of post-transcriptional regulators. miRNAs are a family of 19 to 25 small nucleotide RNAs [2] Analysis of miRNA is leading to new paradigms for control of gene expression during plants and animals. Most noncoding RNAs are characterized by a specific secondary structure that determine their function [3].

Micro RNA is the noncoding class of RNA which plays major role post transcriptional regulation of genes [4] Analysis of miRNAs is leading to new paradigms for control of gene expression during development in plants and animals. MiRNAs arise from larger precursor molecules that can fold into a stable stem-loop structure [5-9]. Those structures are processed by ribonuclease III-like nuclease Dicer in animals and Dicer like in plants and all have a typical stem-loop shape [5-12].

About- *Caenorhabditis elegans*

C. elegans is unsegmented, vermiform, and bilaterally symmetrical. It has a cuticle (a tough outer covering, as an exoskeleton), four main epidermal cords, and a fluid-filled pseudocoelom (body cavity). It also has some of the same organ systems as larger animals. About one in a thousand individuals is male and the rest are hermaphrodites [13]. The basic anatomy of C. elegans includes a mouth, pharynx, intestine, gonad, and collagenous cuticle. Like all nematodes, they have neither a circulatory nor a respiratory system. The four bands of muscles that run the length of the body are connected to a neural system that allows the muscles to move the animal's body only as dorsal bending or ventral bending, but not left or right, except for the head, where the four muscle quadrants are wired independently from one another.

When a wave of dorsal/ventral muscle contractions proceeds from the back to the front of the animal, the animal is propelled backwards. When a wave of contractions is initiated at the front and proceeds posteriorly along the body, the animal is propelled forwards. Because of this dorsal/ventral bias in body bends, any normal living, moving individual tends to lie on either its left side or its right side when observed crossing a horizontal surface. A set of ridges on the lateral sides of the body cuticle, the alae, are believed to give the animal added traction during these bending motions.

The pharynx is a muscular food pump in the head of C. elegans, which is triangular in cross-section. This grinds food and transports it directly to the intestine. A set of "valve cells" connects the pharynx to the intestine, but how this valve operates is not understood. After digestion, the contents of the intestine are released via the rectum, as is the case with all other nematodes [14].

Methodology

First, we do our work on literature survey and also working on computational software's and databases. After that precursor sequences of *C. elegance* are retrieves from miRBase (miRBase: Targets provides an automated pipeline for the prediction of targets for all published animal miRNAs and precursor sequences). Then we focused on thermodynamic study with minimum free energy of fetched precursor sequences for potential miRNA analysis of *C. elegance*. After this analysis we select potential miRNA on behalf of precursor thermodynamics ratings and miRBase high-confidence sequences rating of *C. elegance*s then selected miRNAs of *C. elegance* submitted in Diana Tv3.0 (Tool) for targeted gene identification in human genome via selected miRNAs of *C. elegance* [15-18].

Results and Discussion

Result analysis of this project divide on two categories

Identification of genes via miRNAs of *C. elegance*: MicroRNAs (miRNAs) are a class of short endogenously expressed RNA molecules that regulate gene expression by binding directly to the messenger RNA of protein coding genes. They have been found to confer a novel layer of genetic regulation in a wide range of biological processes. Computational miRNA target prediction remains one of the key means used to decipher the role of miRNAs in development and disease. Here we introduce the basic idea behind the experimental identification of miRNA targets and present some of the most widely used computational miRNA target identification programs. The review includes an assessment of the prediction quality of these programs and their combinations [19].

This REST service can access the DIANA-microT-ANN (v4) web server and identify microRNAs (miRNAs) predicted to target selected genes OR gene targets of selected miRNAs. The user can add 1 or more miRNAs AND/OR 1 or more genes. The base case is the selection of only one miRNA OR only one gene. In this case, the server identifies all targeted genes by the selected miRNA or all the miRNAs predicted to target the selected gene. This case is extended by adding more miRNAs or genes. If the user specifies miRNAs and genes of interest, the server will identify which of the selected genes are targeted by the selected miRNAs Parameterization [20].

Identification of gene function in human via miRNAs of *C. elegance*: DIANA LAB, Fleming: Computational predictive models are a key element of current systems biology. The focus of the DIANA lab is on the development of algorithms, databases and tools for interpreting and archiving genomic data in the framework of a systemic analysis. Current emphasis is on the analysis of microRNA (miRNA) and protein coding genes.

MiRNAs are recently identified to be very abundant in mammalian organisms and play a key role in regulating development. Comprehensive models work and integrate data at various levels of biological detail. Therefore, the activities of the DIANA lab range from the analysis of expression regulation from deep sequencing data, the annotation of miRNA regulatory elements and targets to the interpretation of the role of miRNAs in various diseases (Tables 1-3).

S. No	miRNA id (C. elegance) from mirbase	Accession No.	Sequence of miRNA
1 (A)	>cel-miR-35-3p	MIMAT0000006	UCACCGGGUGGAAACUAGCAGU
2 (B)	>cel-miR-795-5p	MIMAT0004230	UGAGGUAGAUUGAUCAGCGAGCUU
3 (C)	>cel-miR-37-5p	MIMAT0015094	UGUGGGUGUCCGUUGCGGUGCUA
4 (D)	>cel-let-7-5p	MIMAT0000001	UGAGGUAGUAGGUUGUAUAGUU
5 (E)	>cel-miR-2208b-5p	MIMAT0011429	AAGUGUACCCGGAUCUGAUAUCC
6 (F)	>cel-miR-80-3p	MIMAT0000053	UGAGAUCAUUAGUUGAAAGCCGA
7 (G)	>cel-miR-42-5p	MIMAT0015095	GUGGGUGUUUGCUUUUUCGGUGAAG
8 (H)	>cel-lin-4-5p	MIMAT0000002	UCCCUGAGACCUCAAGUGUGA
9 (I)	>cel-miR-79-5p	MIMAT0020322	CUUUGGUGAUUCAGCUUCAAUGA
10 (J)	>cel-miR-239a-3p	MIMAT0020334	AGUGUCUAGUCUAGUGCAAACA
11 (K)	>cel-miR-74-5p	MIMAT0020320	CGGGCUUCCAUCUCUUUCCCAGC
12 (L)	>cel-miR-355-5p	MIMAT0000697	UUUGUUUUAGCCUGAGCUAUG
13 (M)	>cel-miR-791-5p	MIMAT0020350	ACCUUAUCCGUUGUAGCCAAAGU
14 (N)	>cel-miR-50-5p	MIMAT0000021	UGAUAUGUCUGGUAUUCUUGGGUU
15 (O)	>cel-miR-5545-3p	MIMAT0022182	ACUUUUGUAGAUCAAACCGACAU
16 (P)	>cel-miR-4805-5p	MIMAT0019986	UGCGGCAAAUUUGCCGAAUUUGC
17 (Q)	>cel-miR-61-5p	MIMAT0015103	UGGGUUACGGGGCUUAGUCCUU
18 (R)	>cel-miR-786-3p	MIMAT0004221	UAAUGCCCUGAAUGAUGUUCAAU
19 (S)	>cel-miR-231-3p	MIMAT0000286	UAAGCUCGUGAUCAACAGGCAGAA
20 (T)	>cel-miR-45-5p	MIMAT0020772	CUGGAUGUGCUCGUUAGUCAUA
21 (U)	>cel-miR-53-5p	MIMAT0000024	CACCCGUACAUUUGUUUCCGUGCU

22 (V)	>cel-miR-794-3p	MIMAT0020352	UGAAAACGUUGUCUAUCUCGAA
23 (W)	>cel-miR-2208b-3p	MIMAT0011430	AUGCAGAUUUUGGUACACUUCA
24 (X)	>cel-miR-54-3p	MIMAT0000025	UACCCGUAAUCUUCAUAAUCCGAG
25 (Y)	>cel-miR-55-5p	MIMAT0020313	CGGCAGAAACCUAUCGGUUAUA
26 (Z)	>cel-miR-236-3p	MIMAT0000291	UAAUACUGUCAGGUAAUGACGCU

Table 1: Browsed sequences from miRBase for *C. elegance* miRNA. This table explains that accession n number of selected miRNA with sequences of nucleotide where A=Adenine, G=Guanine, C=Cytosine, U=Uracil.

S. No	miRNA (miRBase) ID	Name of gene Identified
1	MIMAT0000006	*WDR90, C17orf62*
2	MIMAT0004230	*FIGN, LIN28B, IGF2BP1, ONECUT2, ARID3B, NR6A1, CLCN5, FOXP2, HIC2, IGF1R, BACH1, TRIM71, FBN1, YOD1, C14orf28, CPEB2, CCND2, PKN2, BZW1, ZNF512B, PAPPA, FNDC3A, TGFBR1, PTPRD, DUSP16, GALNT1, PBX2, HOXA1, SMARCAD1, TTLL4, PTCH1, CPEB1, NAP1L1, ABCC5, IGF2BP2, CPEB3, ACVR2A, PBX3, PPP1R15B, PPARGC1B, POU2F1, CPEB4, MACF1, ACVR1C, GDF6, INSR, UTRN, DLC1, NAT12, FNDC3B, GATM, COL1A2, MIB1, PGM2L1, ABL2, COL27A1, C1oorf64, WDFY4, ADAMTS8, DMD, ZFYVE26,*
3	MIMAT0015094	*FOXI2, RSPO4, FGFR1, MEF2D, C8orf13, FOXP4, UCK1, SYNGR1, TCF21, HIVEP3, TSPAN18, PCNT, SDC3, PVRL1, ZMIZ1, KLF13, SCN3B, SPNS2, KIAA1324, TTYH3, SPN, ATXN1, SGIP1,*
4	MIMAT0000001	*FIGN, IGF2BP1, LIN28B, NR6A1, ONECUT2, ARID3B, FBN1, BACH1, CLCN5, YOD1, HIC2, IGF1R, FOXP2, PBX3, TRIM71, CCND2, PKN2, MAP4K3, CDV3, COL1A2, IGF2BP3, PAPPA, CPEB2, NAP1L1, C14orf28, PPP1R15B, PTPRD, RAPGEF6, FNIP1, GALNT1, SOCS4, DUSP16, TGFBR1, HOXA1, ARHGAP28, MAP3K1, LRIG2, UHRF2, LRIG3, BPTF, GAS7, CCNJ, B3GNT7, ABL2, CPEB3, ACVR2A, ABCC5, BCAT1, RANBP2, PCDH19, GAN, MAPK8, COL3A1, PPARGC1B,* *BZW1, POU2F1, CPEB4, FNDC3A, HDX, DLC1, NAT12, IDE*
5	MIMAT0011429	*TNPO1, GLIS3, WDR35, ANKRD57, SLC12A2, ST6GALNAC3, DLG2, BICD2, IVNS1ABP, RNF38, NFIB, TP53INP1, C6orf89, IL6ST, EIF5, FIGN*
6	MIMAT0000053	*LUZP4, ZC3H12C, KIAA1853, CLIC2, GFOD1, NLRC3, CARD8, C4orf15, CRX, PSMB2, UBE2R2, S100A7A, OPHN1, RNF144B, EXOSC6, C3orf64, KIAA0408, METTL7A, C20orf117*
7	MIMAT0015095	*FOXI2, RSPO4, FGFR1, SYNGR1, PVRL1, ZMIZ1, FOXP4, C1orf21, PPARD, ATP11A, KLF13, SCN3B, TRDMT1, NPLOC4, MAML1, HAPLN4, PCNT, PRKCA, ADCY1, MECP2, TCF21, TSPAN18, C10orf105,*
8	MIMAT0000002	*SMEK1, FAM120A, TRIM71, MXD4, ATXN7, RYBP, KIAA0174, OSBPL9, PAFAH1B1, KIAA1522, DAAM1, RAPGEF5, ZSWIM6, AEBP2, KLF13, PSCD1, BAK1, BCL2L7P1, MYT1, C11orf57, ST8SIA4, SEMA4D, ABHD6, ZSWIM5, MFHAS1, ARID3B, ADAM9, LRRC8B, SLC39A9, TAF9B, SH3TC2, LIN28, ANPEP, CDC37L1, IRF4, ZSCAN29, MACF1, ENPP1, GGA2, TMEM77, MLF2, CDR2L, TBC1D1, ZNF385, ANKRD50, NIPA1, SUV39H1, TGOLN2, STARD13, NUP210, KPNA6, SGPL1, ESRRA, FUT4, SCN2B, ACHE, ALPK3, SLC35A4, TSEN54, NCAN, GALNT14, TRIAP1, MSI1, BMF, UBE2R2, FAM78A, CDH5, IER3IP1, LFNG, PCTP, NPL, MCL1, DIRAS1, HCN3, C14orf43, SUV420H2, C10orf105, KLHL6*
9	MIMAT0020322	*ONECUT2, POU2F1, MLL, FOXP2, CPEB2, LIN28B, POU2F2, CNOT6L, VPS13B, NR5A2, DST, FBN1, KIF13A, FOXP1, LDLRAP1, EPB41L3, PRDM6, SRGAP3, PRDM1, PCNP, STC1, KCNA1*
10	MIMAT0020334	*ST8SIA3, MACF1, MECP2, NTRK3, ZNF395, IGF1R, MYO1D, MDN1, ULBP1*
11	MIMAT0020320	*KIAA1045, MACF1, HPS1, FAM86C, ADM2, AFF2, DST, ATP2B3, KIAA1529, CECR6, ZNF783, PLEKHF1, ARC, RAB43, ISY1, LARP1, H6PD, BMP1, SLC12A5, LYNX1, IKZF1, ZNF135*
12L	MIMAT0000697	*TNRC6B, ONECUT2, GK5, ELOVL6, ELAVL4, DST, ZNF462, MACF1, LAMP2, GABRA4, AMMECR1, VGLL3, XKR6,*
13	MIMAT0020350	*POU2F1, CRB1, DST, ADAM22, ANKRD15, ST13, FAM10A6, ST13P7*
14	MIMAT0000021	*GPR123, BACH2, KPNA5, DDEF2, MYO5A, PSD3, KIAA2018, NEUROD1, EXT1, RPS6KA3, STK38L, IKZF2, ERG, TNRC6A, BCL11A, SMAD2, KLHDC5, C12orf5, OSBPL6, PCDH17, MOBKL1A, AMMECR1, ANKRD34, FOXP2, RNF144B, PHF20L1,*
15	MIMAT0022182	*BNC2, MYEF2, POU2F1, CRB1, MACF1, NTRK3, NFIB, PROS1, HS2ST1, CCDC144B, SASH1, DIAPH2, PDE10A, SMAD2, ZBTB41, NOVA1, KIAA0430, PAG1, PHF16, ELOVL6, ACADSB, IPO9, CLIC4, GAD2, TTN, MLL, USP6, CNTNAP2, ANKRD30B, SETD7, CCDC93, ARGFX, GDA, ZNF587,*

16	MIMAT0019986	*No data could be retrieved from the database*
17	MIMAT0015103	*KIAA1853, ADH6, CNOT6L, USP47, PTPRT, ARCN1,*
18	MIMAT0004221	*KIAA1609, NFIB, DST, NUFIP2, GPR92, SGTB, KIAA1239, WDR37, ZNF644, KIAA1147, VGLL3, MLL5, EPC1, KCNJ2, ANKRD28, CLCN6, DCP2, DOCK4, USP48, KIR3DX1, ANKS6, SH3PXD2A, EPAS1, IL1RAP, TMEM164, PIK3R3, HDAC9, EHF, ZNRF1, NR3C2, ARRDC3, ACVR1, RIT1, KCNA1, UBR4, DTNA, APOL6, ARF3, TRHDE, LAMP2, NSL1, RGS9BP, ENAH, TBCEL, PCDH15, ADCY1*
19	MIMAT0000286	*DLGAP2, MDN1*
20	MIMAT0020772	*GSG1L, CLDN19, PRKCA, C18orf1, KIAA0226, C1orf141, SMAD2, KLHL21, C1orf116, SS18L1, CRAMP1L, HN1L, RERE, C1orf21, XRRA1, UBR2, ACACB, VGLL3, CHL1, TMEM125, KLF12, ARC, AFAP1, BNC2, DNMBP, UBR4, TRIOBP, LPAL2, MEGF6, URB1, CDH23, TEAD1, SARM1, NCOA6*
21	MIMAT0000024	*BAZ2A, CDH23, NMT1, TRIB2, SLC25A16, ADCY1, MBNL1, KBTBD8, FZD8, AGO2, FZD5, GIYD1, GIYD2, ITGB1, HS3ST2, FGFR3, MTMR3, HS3ST3B1, LRRC8B*
22	MIMAT0020352	*GABRA4, NTRK3, KITLG, MYO5A, COL19A1, ATP9A, RAB3IP, NHS, POU2F1, PRPF4B, QSER1, BCL2, PSME3, YTHDF3, SFMBT2, DCX, KIAA1600, FGF2, SNTB2, C5orf21, SUDS3, KLHDC5, SOCS4, NFAT5, ADH1B, ADH1C, DTWD2, AKAP11, MAP3K7IP3, DLGAP2, XYLT1, TEAD1, PTEN, PTENP1, BTBD7, PCGF3*
23	MIMAT0011430	*TNRC6B, RAB11FIP2, MBNL2, ZNF264, HSPA12A, RAB3IP, MLL, ERBB4, ZFAND5, NFIB, LUZP1, DST, NAV1, FBXW2, AFF2, OSBPL3, BCL7B, GRIN2A, NCOA1, MAFB, CPSF6, F11R, GAS7, CCDC6, C6orf103, CCDC140, PSD3, MAMDC2, CAMK2G, C18orf1, RC3H1, PARK2, C10orf63, TMEM47, SPATA13, IYD, ITGB1, FAM91A1, KIAA2026, SRSF1, PTGFRN, TTN, MACF1*
24	MIMAT0000025	*TRIB2, MBNL1, KBTBD8, HOXA1, ZZEF1, SLC25A16, NTRK3, C4orf16, CTDSPL, PPP3CA, SMARCA5, LRRC8B, THAP2, CLDN11, EPDR1, ICMT, LIF, BAZ2A, ADCY1, FZD8, EIF2C2, FZD5, CAMTA1, MTMR3, ST6GALNAC4, FRAP2, PRDM1, PPP1CB, C1orf121, INSM1, AMMECR1, BMPR2, CDC25A, SETD1B, TMEM30A, RAC1, CYP26B1, ZNRF2, FIGN, TRIB1, PPFIA3, SMARCD1, ZBTB7A, CCDC21, ST5, RAVER2, PI15, JMJD3*
25	MIMAT0020313	*SV2B, NOTCH2, KIAA1147, SSR2, JARID1A, GRIN2A, C9orf7, ASB1, KRT84, SMU1, GCC1, TNRC6B, MMD, CCND2, FRAS1, RUNX1, HEMK1, C22orf29, WDTC1, SMAD2, G6PC, TOLLIP, UST, ATP9A, PVR, E2F2, FOSL2, FOXK1, XYLT1, GRM4, TBC1D13, ZDHHC8, CYP8B1, CHST3, RNF216,*
26	MIMAT0000291	*ZEB2, FIGN, GPM6A, FBXW7, KIAA1432, RPS6KB1, TSGA14, RAP1B, LRP1B, JHDM1D, NFIB, MAP2, C16orf72, YWHAG, RAP2C, PAPD5, NEGR1, RANBP10, TRIM33, PCTK2, MSN, NOVA1, CRKL, EPS8, CDYL, WIPF1, SYNJ1, SESN1, WAPL, LCORL, PKN2, ELAVL2, MEX3B, CTDSPL2, KIAA0355, RANBP9, BNC2, ESRRG, ERRFI1, AKAP2, PALM2, SCAMP1, CCNJ, MED13*

Table 2: Selected miRNAs (miRBase- ID) for identification of genes in humans.

S. No.	Potential gene	Functions involves in humans as genomics
1.	*LIN28B*	This gene is highly expressed in testis, fetal liver, placenta, and in primary human tumors and cancer cell lines. It is negatively regulated by microRNAs that target sites in the 3' UTR, and overexpression of this gene in primary tumors is linked to the repression of let-7 family of microRNAs and derepression of let-7 targets, which facilitates cellular transformation.
2.	*CLCN5*	The encoded protein is primarily localized to endosomal membranes and may function to facilitate albumin uptake by the renal proximal tubule.
3.	*FOXP2*	The protein product contains a FOX DNA-binding domain and a large polyglutamine tract and is an evolutionarily conserved transcription factor, which may bind directly to approximately 300 to 400 gene promoters in the human genome to regulate the expression of a variety of genes.
4.	*FBN1*	These microfibrils provide force bearing structural support in elastic and nonelastic connective tissue throughout the body. Mutations in this gene are associated with Marfan syndrome, isolated ectopia lentis, autosomal dominant Weill-Marchesani syndrome, MASS syndrome, and Shprintzen-Goldberg craniosynostosis syndrome.
5.	*YOD1*	Protein ubiquitination controls many intracellular processes, including cell cycle progression, transcriptional activation, and signal transduction. This dynamic process, involving ubiquitin conjugating enzymes and deubiquitinating enzymes, adds and removes ubiquitin.
6.	*C14orf28*	This disorder is characterized by severe lung complications and liver dysfunction. Notably, the immunoglobulin heavy chain locus is found on chromosome 14 and has been identified as a fusion with the chromosome 19 encoded protein BCL3 in the (14;19) translocations found in a variety of B cell malignancies.
7.	*CPEB2*	The protein encoded by this gene is highly similar to cytoplasmic polyadenylation element binding protein (CPEB), an mRNA-binding protein that regulates cytoplasmic polyadenylation of mRNA as a trans factor in oogenesis and spermatogenesis.

8.	*CCND2*	High level expression of this gene was observed in ovarian and testicular tumors. Mutations in this gene are associated with megalencephaly-polymicrogyria-polydactyly-hydrocephalus syndrome 3 (MPPH3).
9.	*PKN2*	PKC-related serine/threonine-protein kinase and Rho/Rac effector protein that participates in specific signal transduction responses in the cell. Plays a role in the regulation of cell cycle progression, actin cytoskeleton assembly, cell migration, cell adhesion, tumor cell invasion and transcription activation signaling processes.
10.	*BZW1*	BZW1 Enhances histone H4 gene transcription but does not seem to bind DNA directly. Belongs to the BZW family. 4 isoforms of the human protein are produced by alternative splicing. BZW1 Enhances histone H4 gene transcription but does not seem to bind DNA directly.
11.	*ZNF512B*	ZNF512B is central to a gene sensor circuitry involving cell-cycle regulators, TGFβ effectors, Drosha and microRNAs with opposite oncogenic potentials
12.	*TGFBR1*	The encoded protein is a serine/threonine protein kinase. Mutations in this gene have been associated with Loeys-Dietz aortic aneurysm syndrome (LDAS). Multiple transcript variants encoding different isoforms have been found for this gene.
13.	*PTPRD*	This PTP contains an extracellular region, a single transmembrane segment and two tandem intracytoplasmic catalytic domains, and thus represents a receptor-type PTP. The extracellular region of this protein is composed of three Ig-like and eight fibronectin type III-like domains.
14.	*PBX2*	This protein is a transcriptional activator which binds to the TLX1 promoter. The gene is located within the major histocompatibility complex (MHC) on chromosome 6.
15.	*SMARCAD1*	This gene encodes a member of the SNF subfamily of helicase proteins. The encoded protein plays a critical role in the restoration of heterochromatin organization and propagation of epigenetic patterns following DNA replication by mediating histone H3/H4 deacetylation.
16.	*PTCH1*	This gene encodes a member of the patched gene family. The encoded protein is the receptor for sonic hedgehog, a secreted molecule implicated in the formation of embryonic structures and in tumorigenesis, as well as the desert hedgehog and indian hedgehog proteins.
17.	*CPEB1*	This gene encodes a member of the cytoplasmic polyadenylation element binding protein family. This highly conserved protein binds to a specific RNA sequence, called the cytoplasmic polyadenylation element, found in the 3' untranslated region of some mRNAs.
18.	*ABCC5*	This protein may be involved in resistance to thiopurines in acute lymphoblastic leukemia and antiretroviral nucleoside analogs in HIV-infected patients. Alternative splicing of this gene has been detected; however, the complete sequence and translation initiation site is unclear.
19.	*IGF2BP2*	The protein encoded by this gene contains four KH domains and two RRM domains. It functions by binding to the 5' UTR of the insulin-like growth factor 2 (IGF2) mRNA and regulating IGF2 translation
20.	*CPEB3*	The mammalian CPEB3 ribozyme is a self-cleaving non-coding RNA located in the second intron of the CPEB3 gene which belongs to a family of genes regulating messenger RNA polyadenylation.
21.	*PBX3*	PBX3 (Pre-B-Cell Leukemia Homeobox 3) is a Protein Coding gene. Among its related pathways are Transcriptional misregulation in cancer.
22.	*CPEB4*	CPEB4 is a candidate biomarker for defining metastatic cancers and directing personalized therapies.
23.	*GDF6*	This gene encodes a member of the bone morphogenetic protein (BMP) family and the TGF-beta superfamily of secreted signaling molecules. It is required for normal formation of some bones and joints in the limbs, skull, and axial skeleton.
24.	*DLC1*	This gene functions as a tumor suppressor gene in a number of common cancers, including prostate, lung, colorectal, and breast cancers. Multiple transcript variants due to alternative promoters and alternative splicing have been found for this gene.
25.	*NAT12*	NAA30 (N (Alpha)-Acetyltransferase 30, NatC Catalytic Subunit) is a Protein Coding gene. Diseases associated with NAA30 include eastern equine encephalitis.
26.	*GATM*	This gene encodes a mitochondrial enzyme that belongs to the amidinotransferase family. This enzyme is involved in creatine biosynthesis, whereby it catalyzes the transfer of a guanido group from L-arginine to glycine, resulting in guanidinoacetic acid, the immediate precursor of creatine.
27.	*COL1A2*	Symptoms associated with mutations in this gene, however, tend to be less severe than mutations in the gene for the alpha1 chain of type I collagen (COL1A1) reflecting the different role of alpha2 chains in matrix integrity. Three transcripts, resulting from the use of alternate polyadenylation signals, have been identified for this gene.
28.	*ABL2*	This gene is expressed in both normal and tumor cells, and is involved in translocation with the ets variant 6 gene in leukemia. Multiple alternatively spliced transcript variants encoding different protein isoforms have been found for this gene
29.	*COL27A1*	This gene encodes a member of the fibrillar collagen family, and plays a role during the calcification of cartilage and the transition of cartilage to bone. The encoded protein product is a preproprotein. It includes an N-terminal signal peptide, which is followed by an N-terminal propetide, mature peptide and a C-terminal propeptide.

30.	*WDFY4*	WDFY4shows polymorphisms in Chinese Patients with rheumatoid arthritis.
31.	*ZFYVE26*	The presence of this domain is thought to target these proteins to membrane lipids through interaction with phospholipids in the membrane. Mutations in this gene are associated with autosomal recessive spastic paraplegia-15.
32.	*RSPO4*	The encoded protein may be involved in activation of Wnt/beta-catenin signaling pathways. Mutations in this gene are associated with anonychia congenital. Alternate splicing results in multiple transcript variants.
33.	*MEF2D*	This gene is a member of the myocyte-specific enhancer factor 2 (MEF2) family of transcription factors. Members of this family are involved in control of muscle and neuronal cell differentiation and development, and are regulated by class II histone deacetylases.
34.	*FOXP4*	Many members of the forkhead box gene family, including members of subfamily P, have roles in mammalian oncogenesis. This gene may play a role in the development of tumors of the kidney and larynx. Alternative splicing of this gene produces multiple transcript variants, some encoding different isoforms.
35.	*TSPAN18*	Tetraspanins are a superfamily of four-transmembrane proteins. They associate laterally withone another and with other transmembrane proteins to form membrane microdomains, inwhich associated proteins are regulated.
36.	*PCNT*	The protein encoded by this gene binds to calmodulin and is expressed in the centrosome. It is an integral component of the pericentriolar material (PCM). The protein contains a series of coiled-coil domains and a highly conserved PCM targeting motif called the PACT domain near its C-terminus.
37.	*ZMIZ1*	This gene encodes a member of the PIAS (protein inhibitor of activated STAT) family of proteins. The encoded protein regulates the activity of various transcription factors, including the androgen receptor, Smad3/4, and p53. The encoded protein may also play a role in sumoylation.
38.	*SPNS2*	Transports sphingosine 1-phosphate (S1P), a secreted lipid mediator that plays critical roles in cardiovascular, immunological, and neural development and function. Mediates the export of S1P from cells in the extraembryonic yolk syncytial layer (YSL), thereby regulating myocardial precursor migration.
39.	*KIAA1324*	Expression of this gene is induced by estrogen and the encoded protein has been characterized as a transmembrane protein. The encoded protein has been found in to correlate with survival in certain carcinomas and may be important for cellular response to stress. Alternative splicing results in multiple transcript variants.
40.	*SPN*	It may be part of a physiologic ligand-receptor complex involved in T-cell activation. During T-cell activation, this protein is actively removed from the T-cell-APC (antigen-presenting cell) contact site, suggesting a negative regulatory role in adaptive immune response.
41.	*SGIP1*	SGIP1 functions as an endocytic protein that affects signaling by receptors in neuronal systems involved in energy homeostasis via its interaction with endophilins
42.	*LRIG2*	This gene encodes a transmembrane protein containing leucine-rich repeats and immunoglobulin-like domains. The encoded protein promotes epidermal growth factor signalling, resulting in increased proliferation. Its expression in the cytoplasm of glioma cells is correlated with poor survival. Mutations in this gene can cause urofacial syndrome. Alternative splicing results in multiple transcript variants.
43.	*LRIG3*	The results on expression of other tumor markers suggest that LRIG3 is influenced by or influences a pattern of tumor markers in cancer
44.	*BPTF*	This gene was identified by the reactivity of its encoded protein to a monoclonal antibody prepared against brain homogenates from patients with Alzheimer's disease.
45.	*CCNJ*	MiR-125b acts as a tumor suppressor in breast tumorigenesis via its novel direct targets ENPEP, CK2-alpha, CCNJ, and MEGF9.
46.	*B3GNT7*	Suppression of B3GNT7 gene expression in colon adenocarcinoma and its potential effect in the metastasis of colon cancer cells.
47.	*BCAT1*	This gene encodes the cytosolic form of the enzyme branched-chain amino acid transaminase. This enzyme catalyzes the reversible transamination of branched-chain alpha-keto acids to branched-chain L-amino acids essential for cell growth.
48.	*PCDH19*	The protein encoded by this gene is a member of the delta-2 protocadherin subclass of the cadherin superfamily. The encoded protein is thought to be a calcium-dependent cell-adhesion protein that is primarily expressed in the brain.
49.	*GAN*	This gene encodes a member of the cytoskeletal BTB/kelch (Broad-Complex, Tramtrack and Bric a brac) repeat family. The encoded protein plays a role in neurofilament architecture and is involved in mediating the ubiquitination and degradation of some proteins. Defects in this gene are a cause of giant axonal neuropathy (GAN).
50.	*MAPK8*	The protein encoded by this gene is a member of the MAP kinase family. MAP kinases act as an integration point for multiple biochemical signals, and are involved in a wide variety of cellular processes such as proliferation, differentiation, transcription regulation and development.
51.	*COL3A1*	This gene encodes the pro-alpha1 chains of type III collagen, a fibrillar collagen that is found in extensible connective tissues such as skin, lung, uterus, intestine and the vascular system, frequently in association with type I collagen.

52.	*HDX*	HDX (highly divergent homeobox), also known as CXorf43, is a 690-amino acid nuclear protein that contains two homeobox DNA-binding domains.
53.	*IDE*	This gene encodes a zinc metallopeptidase that degrades intracellular insulin, and thereby terminates insulins activity, as well as participating in intercellular peptide signalling by degrading diverse peptides such as glucagon, amylin, bradykinin, and kallidin.
54.	*GLIS3*	This gene is a member of the GLI-similar zinc finger protein family and encodes a nuclear protein with five C2H2-type zinc finger domains.
55.	*WDR35*	This gene encodes a member of the WD repeat protein family. WD repeats are minimally conserved regions of approximately 40 amino acids typically bracketed by gly-his and trp-asp (GH-WD), which may facilitate formation of heterotrimeric or multiprotein complexes.
56.	*ANKRD57*	Gastrointestinal tumor; glioma; head and neck tumor; kidney tumor; liver tumor; lung tumor; non-neoplasia; normal; ovarian tumor; prostate cancer; retinoblastoma; soft tissue/muscle tissue tumor; uterine tumor; blastocyst; fetus; juvenile; adult;Protein expression in breast cancer tissues
57.	*IVNS1ABP*	Plays a role in cell division and in the dynamic organization of the actin skeleton as a stabilizer of actin filaments by association with F-actin through Kelch repeats.
58.	*RNF38*	This gene encodes a protein with a coiled-coil motif and a RING-H2 motif (C3H2C2) at its carboxy-terminus. The RING motif is a zinc-binding domain found in a large set of proteins playing roles in diverse cellular processes including oncogenesis, development, signal transduction, and apoptosis.
59.	*NFIB*	NFIB is a potential target for estrogen receptor-negative breast cancers.
60.	*TP53INP1*	Antiproliferative and proapoptotic protein involved in cell stress response which acts as a dual regulator of transcription and autophagy. Acts as a positive regulator of autophagy. In response to cellular stress or activation of autophagy, relocates to autophagosomes where it interacts with autophagosome-associated proteins GABARAP, GABARAPL1/L2, MAP1LC3A/B/C and regulates autophagy.
61.	*EIF5*	Eukaryotic translation initiation factor-5 (EIF5) interacts with the 40S initiation complex to promote hydrolysis of bound GTP with concomitant joining of the 60S ribosomal subunit to the 40S initiation complex. The resulting functional 80S ribosomal initiation complex is then active in peptidyl transfer and chain elongations
62.	*KIAA1853*	SRRM4 promotes alternative splicing and inclusion of neural-specific exons in target mRNAs
63.	*CLIC2*	This gene encodes a chloride intracellular channel protein. Chloride channels are a diverse group of proteins that regulate fundamental cellular processes including stabilization of cell membrane potential, transepithelial transport, maintenance of intracellular pH, and regulation of cell volume.
64.	*S100A7A*	May be involved in epidermal differentiation and inflammation and might therefore be important for the pathogenesis of psoriasis and other diseases.
65.	*OPHN1*	This gene encodes a Rho-GTPase-activating protein that promotes GTP hydrolysis of Rho subfamily members. Rho proteins are important mediators of intracellular signal transduction, which affects cell migration and cell morphogenesis.
66.	*C3orf64*	C3orf64 (chromosome 3 open reading frame 64), also known as AER61, is a 527-amino acid secreted protein that belongs to the glycosyltransferase 61 family and exists as three alternatively spliced isoforms.
67.	*KIAA0408*	KIAA0408 is a 694-amino acid protein that is alternatively spliced into 2 isoforms. The gene encoding KIAA0408 maps to chromosome 6, which makes up nearly 6% of the human genome and contains around 1,200 genes within 170 million base pairs of sequence.
68.	*METTL7A*	METTL7A expression, indicating an editing-independent regulation of ADARs. ADARs regulation of METLL7A was demonstrated to be mediated though microRNA. Moreover, METTL7A was found to be a tumor suppressor in HCC, and primary HCC tumors demonstrated reduced METTL7A expression.
69.	*C20orf117*	Regulates autophagy by playing a role in the reduction of glucose production in an adiponectin- and insulin-dependent manner
70.	*FGFR1*	The protein encoded by this gene is a member of the fibroblast growth factor receptor (FGFR) family, where amino acid sequence is highly conserved between members and throughout evolution
71.	*PVRL1*	This gene encodes an adhesion protein that plays a role in the organization of adherens junctions and tight junctions in epithelial and endothelial cells.
72.	*ZMIZ1*	This gene encodes a member of the PIAS (protein inhibitor of activated STAT) family of proteins. The encoded protein regulates the activity of various transcription factors, including the androgen receptor, Smad3/4, and p53. The encoded protein may also play a role in sumoylation.
73.	*PPARD*	This gene encodes a member of the peroxisome proliferator-activated receptor (PPAR) family. PPARs are nuclear hormone receptors that bind peroxisome proliferators and control the size and number of peroxisomes produced by cells.

74.	*TRDMT1*	This gene encodes a protein responsible for the methylation of aspartic acid transfer RNA, specifically at the cytosine-38 residue in the anticodon loop.
75.	*PCNT*	The protein encoded by this gene binds to calmodulin and is expressed in the centrosome. It is an integral component of the pericentriolar material (PCM). The protein contains a series of coiled-coil domains and a highly conserved PCM targeting motif called the PACT domain near its C-terminus.
76.	*MECP2*	DNA methylation is the major modification of eukaryotic genomes and plays an essential role in mammalian development. Human proteins MECP2, MBD1, MBD2, MBD3, and MBD4 comprise a family of nuclear proteins related by the presence in each of a methyl-CpG binding domain (MBD).
77.	*FAM120A*	FAM120A (Family with Sequence Similarity 120A) is a Protein Coding gene. Diseases associated with FAM120A include myiasis and achondrogenesis.
78.	*KIAA1522*	Chromosome 1 houses a large number of disease-associated genes, including those that are involved in familial adenomatous polyposis, Stickler syndrome, Parkinson's disease, Gaucher disease, schizophrenia and Usher syndrome. Aberrations in chromosome 1 are found in a variety of cancers, including head and neck cancer, malignant melanoma and multiple myeloma.
79.	*MYT1*	The protein encoded by this gene is a member of a family of neural specific, zinc finger-containing DNA-binding proteins. The protein binds to the promoter regions of proteolipid proteins of the central nervous system and plays a role in the developing nervous system.
80.	*C11orf57*	C11orf57 (chromosome 11 open reading frame 57), also known as FLJ10726, is a 292-amino acid protein that exists as 3 alternatively spliced isoforms and is encoded by a gene located on human chromosome 11. With approximately 135 million base pairs and 1,400 genes, chromosome 11 makes up around 4% of human genomic DNA and is considered a gene and disease association dense chromosome.
81.	*ABHD6*	ABHD6 is a serine hydrolyzing enzyme that possesses typical α/β-hydrolase family domains. ABHD6 was first studied because of its over-expression in certain forms of tumours. ABHD6 has been linked to the endocannabinoid system as it controls the accumulation of 2-AG at the cannabinoid receptors. Alongside MAGL and ABHD12, ABHD6 controls 99% of 2-AG signalling in the brain.
82.	*SH3TC2*	This gene encodes a protein with two N-terminal Src homology 3 (SH3) domains and 10 tetratricopeptide repeat (TPR) motifs, and is a member of a small gene family.
83.	*LIN28*	LIN28 encodes a microRNA-binding protein that binds to and enhances the translation of the IGF-2 (insulin-like growth factor 2) mRNA. Lin28binds to the let-7 pre-microRNA.
84.	*ANPEP*	Aminopeptidase N is located in the small-intestinal and renal microvillar membrane, and also in other plasma membranes. In the small intestine aminopeptidase N plays a role in the final digestion of peptides generated from hydrolysis of proteins by gastric and pancreatic proteases
85.	*IRF4*	The protein encoded by this gene belongs to the IRF (interferon regulatory factor) family of transcription factors, characterized by a unique tryptophan pentad repeat DNA-binding domain. The IRFs are important in the regulation of interferons in response to infection by virus, and in the regulation of interferon-inducible genes.
86.	*ENPP1*	This gene is a member of the ecto-nucleotide pyrophosphatase/phosphodiesterase (ENPP) family. The encoded protein is a type II transmembrane glycoprotein comprising two identical disulfide-bonded subunits.
87.	*MLF2*	MLF2 have roles in tumor initiation and metastasis in breast cancer that involve nitric oxide synthase signaling
88.	*NIPA1*	This gene encodes a magnesium transporter that associates with early endosomes and the cell surface in a variety of neuronal and epithelial cells. This protein may play a role in nervous system development and maintenance.
89.	*SUV39H1*	This gene encodes an evolutionarily-conserved protein containing an N-terminal chromodomain and a C-terminal SET domain. The encoded protein is a histone methyltransferase that trimethylates lysine 9 of histone H3, which results in transcriptional gene silencing.
90.	*STARD13*	This gene encodes a protein which contains an N-terminal sterile alpha motif (SAM) for protein-protein interactions, followed by an ATP/GTP-binding motif, a GTPase-activating protein (GAP) domain, and a C-terminal STAR-related lipid transfer (START) domain.
91.	*SCN2B*	The protein encoded by this gene is the beta 2 subunit of the type II voltage-gated sodium channel. The encoded protein is involved in cell-cell adhesion and cell migration. Defects in this gene can be a cause of Brugada Syndrome, atrial fibrillation, or sudden infant death syndrome.
92.	*ALPK3*	A member of the mitogen-activated protein kinase superfamily, MAK, has been proposed to have an important role in spermatogenesis,
93.	*SLC35A4*	SLC35A4 (solute carrier family 35, member A4) is a 324-amino acid multi-pass membrane protein that belongs to the nucleotide-sugar transporter (NST) family and the SLC35A subfamily.
94.	*TSEN54*	This gene encodes a subunit of the tRNA splicing endonuclease complex, which catalyzes the removal of introns from precursor tRNAs. The complex is also implicated in pre-mRNA 3-prime end processing.

95.	*TRIAP1*	Determined the TRIAP1 mRNA levels in a panel of human tissues and found its expression to be ubiquitous. Normal breast, as well as non-tumorigenic breast cells, exhibited lower TRIAP1 mRNA levels than breast cancer cells or their drug-resistant derivatives.
96.	*MSI1*	This gene encodes a protein containing two conserved tandem RNA recognition motifs. Similar proteins in other species function as RNA-binding proteins and play central roles in posttranscriptional gene regulation.
97.	*FAM78A*	Regulate characterized by harmful vascular defects, is associated with the chromosome 9 gene encoding endoglin protein, ENG. Familial dysautonomia is also associated with chromosome 9 though through the gene IKBKAP.
98.	*IER3IP1*	This gene encodes a small protein that is localized to the endoplasmic reticulum (ER) and may play a role in the ER stress response by mediating cell differentiation and apoptosis.
99.	*C14orf43*	C14orf43 (chromosome 14 open reading frame 43) is a 1,045-amino acid nuclear protein that contains one ELM2 domain and a SANT domain. The gene encoding C14orf43 maps to human chromosome 14q24.2.
100.	*KLHL6*	The Kelch-like (KLHL) gene family encodes a group of proteins that generally possess a BTB/POZ domain, a BACK domain, and five to six Kelch motifs. BTB domains facilitate protein binding and dimerization.
101.	*MLL*	This gene encodes a transcriptional coactivator that plays an essential role in regulating gene expression during early development and hematopoiesis. The encoded protein contains multiple conserved functional domains.
102.	*CPEB2*	The protein encoded by this gene is highly similar to cytoplasmic polyadenylation element binding protein (CPEB), an mRNA-binding protein that regulates cytoplasmic polyadenylation of mRNA as a trans factor in oogenesis and spermatogenesis
103.	*VPS13B*	This gene encodes a potential transmembrane protein that may function in vesicle-mediated transport and sorting of proteins within the cell.
104.	*NR5A2*	The orphan nuclear receptor (NR5A2), which belongs to the NR5A subfamily of nuclear receptors, is expressed in developing and adult tissues of endodermal origin, and can contribute to the development of several cancers through regulating cell proliferation.
105.	*FOXP1*	This gene belongs to subfamily P of the forkhead box (FOX) transcription factor family. Forkhead box transcription factors play important roles in the regulation of tissue- and cell type-specific gene transcription during both development and adulthood.
106.	*LDLRAP1*	The protein encoded by this gene is a cytosolic protein which contains a phosphotyrosine binding (PTD) domain. The PTD domain has been found to interact with the cytoplasmic tail of the LDL receptor. Mutations in this gene lead to LDL receptor malfunction and cause the disorder autosomal recessive hypercholesterolaemia.
107.	*EPB41L3*	EPB41L3, a potential biomarker in cervical cancer, is often silenced by cancer-specific promoter methylation
108.	*SRGAP3*	The novel Rho-GTPase activating gene MEGAP/ srGAP3 has a putative role in severe mental retardation.
109.	*STC1*	This gene encodes a secreted, homodimeric glycoprotein that is expressed in a wide variety of tissues and may have autocrine or paracrine functions. The gene contains a 5' UTR rich in CAG trinucleotide repeats.
110.	*KCNA1*	This gene encodes a voltage-gated delayed potassium channel that is phylogenetically related to the Drosophila Shaker channel. The encoded protein has six putative transmembrane segments (S1-S6), and the loop between S5 and S6 forms the pore and contains the conserved selectivity filter motif (GYGD).
111.	*NTRK3*	This gene encodes a member of the neurotrophic tyrosine receptor kinase (NTRK) family. This kinase is a membrane-bound receptor that, upon neurotrophin binding, phosphorylates itself and members of the MAPK pathway.
112.	*ZNF395*	Liver cancer is one of leading causes of cancer-related deaths. A deeper mechanistic understanding of liver cancer could lead to the development of more effective therapeutic strategies.
113.	*KIAA1045*	KIAA1045 is a 400-amino acid protein that contains one PHD-type zinc finger. The gene encoding KIAA1045 maps to human chromosome 9p13.3. Chromosome 9 houses over 900 genes and comprises nearly 4% of the human genome. Considered to play a role in gender determination, deletion of the distal portion of chromosome 9p can lead to development of male to female sex reversal, the phenotype of a female with a male XY genotype.
114.	*HPS1*	Mutations in this gene are associated with Hermansky-Pudlak syndrome type 1. Alternative splicing results in multiple transcript variants. A pseudogene related to this gene is located on chromosome 22.
115.	*FAM86C*	\ FAM86A is a 330-amino acid protein that belongs to the FAM86 family and is encoded by a gene located on human chromosome 16. Also, members of the FAM86 family, FAM86B2 and FAM86C are encoded by genes that map to human chromosomes 8 and 11, respectively.
116.	*AFF2*	A repeat polymorphism in the fragile X E locus results in silencing of this gene causing Fragile X E syndrome. Fragile X E syndrome is a form of nonsyndromic X-linked mental retardation. Alternate splicing results in multiple transcript variants.
117.	*KIAA1529*	The KIAA1529 gene product has been provisionally designated KIAA1529 pending further characterization.
118.	*CECR6*	Adenosine deaminase is an enzyme that is present in most tissues and exists predominantly as a monomer, although in some tissues it is associated with adenosine deaminase-binding protein.

119.	*ZNF783*	ZNF783 (Zinc Finger Family Member 783) is a Protein Coding gene. Diseases associated with ZNF783 include acrodysostosis. An important paralog of this gene is ZNF746.
120.	*ARC*	Arc, for activity-regulated cytoskeleton-associated protein (also known as Arg3.1), is a plasticity protein first characterized in 1995. Arc is a member of theimmediate-early gene (IEG) family, a rapidly activated class of genes functionally defined by their ability to betranscribed in the presence of protein synthesis inhibitors.
121.	*RAB43*	The low intrinsic GTPase activity of RAB43 is activated by USP6NL. Involved in retrograde transport from the endocytic pathway to the Golgi apparatus. Involved in the transport of Shiga toxin from early and recycling endosomes to the trans-Golgi network.
122.	*IKZF1*	This gene encodes a transcription factor that belongs to the family of zinc-finger DNA-binding proteins associated with chromatin remodeling.
123.	*ZNF135*	ZNF135 (Zinc Finger Protein 135) is a Protein Coding gene. Diseases associated with ZNF135 include renal pelvis carcinoma and ureteral benign neoplasm. Among its related pathways are Gene Expression and Gene Expression.
124.	*ELAVL4*	ELAVL4 (ELAV Like Neuron-Specific RNA Binding Protein 4) is a Protein Coding gene. Diseases associated with ELAVL4 include paraneoplastic neurologic disorders and lung cancer.
125.	*LAMP2*	The protein encoded by this gene is a member of a family of membrane glycoproteins. This glycoprotein provides selectins with carbohydrate ligands
126.	*AMMECR1*	The exact function of this gene is not known, however, submicroscopic deletion of the X chromosome including this gene, COL4A5, and FACL4 genes, result in a contiguous gene deletion syndrome, the AMME complex (Alport syndrome, mental retardation, midface hypoplasia, and elliptocytosis).
127.	*VGLL3*	High expression of VGLL3 placenta and weak expression in pancreas.
128.	*XKR6*	XKR6 (XK, Kell Blood Group Complex Subunit-Related Family, Member 6) is a Protein Coding gene. Diseases associated with XKR6 include keratolytic winter erythema and hydrocele. An important paralog of this gene is XKR9.
129.	*CRB1*	This gene encodes a protein which is similar to the Drosophila crumbs protein and localizes to the inner segment of mammalian photoreceptors.
130.	*ADAM22*	This gene encodes a member of the ADAM (a disintegrin and metalloprotease domain) family. Members of this family are membrane-anchored proteins structurally related to snake venom disintegrins, and have been implicated in a variety of biological processes involving cell-cell and cell-matrix interactions, including fertilization, muscle development, and neurogenesis.
131.	*ST13*	The expression of this gene is reported to be downregulated in colorectal carcinoma tissue suggesting that it is a candidate tumor suppressor gene. Alternative splicing results in multiple transcript variants encoding different isoforms
132.	*FAM10A6*	FAM10A6 is also known as ST13P6. ST13P6 (Suppression of Tumorigenicity 13 (Colon Carcinoma) (Hsp70 Interacting Protein) Pseudogene 6) is a Pseudogene. Diseases associated with ST13P6 include multiple myeloma.
133.	*ST13P7*	ST13P7 (Suppression of Tumorigenicity 13 (Colon Carcinoma) (Hsp70 Interacting Protein) Pseudogene 7) is a Pseudogene. Diseases associated with ST13P7 include multiple myeloma.
134.	*MYO5A*	This gene is one of three myosin V heavy-chain genes, belonging to the myosin gene superfamily. Myosin V is a class of actin-based motor proteins involved in cytoplasmic vesicle transport and anchorage, spindle-pole alignment and mRNA translocation.
135.	*PSD3*	PSD3 (Pleckstrin And Sec7 Domain Containing 3) is a Protein Coding gene. Diseases associated with PSD3 include hepatocellular carcinoma. Among its related pathways is Endocytosis.
136.	*KIAA2018*	Chromosome 3 is made up of about 214 million bases encoding over 1,100 genes. Notably, there is a chemokine receptor gene cluster and a variety of human cancer related loci on chromosome 3. Particular regions of the chromosome 3 short arm are deleted in many types of cancer cells.
137.	*NEUROD1*	This gene encodes a member of the NeuroD family of basic helix-loop-helix (bHLH) transcription factors. The protein forms heterodimers with other bHLH proteins and activates transcription of genes that contain a specific DNA sequence known as the E-box. It regulates expression of the insulin gene, and mutations in this gene result in type II diabetes mellitus.
138.	*EXT1*	This gene encodes an endoplasmic reticulum-resident type II transmembrane glycosyltransferase involved in the chain elongation step of heparan sulfate biosynthesis. Mutations in this gene cause the type I form of multiple exostoses.
139.	*RPS6KA3*	The activity of this protein has been implicated in controlling cell growth and differentiation. Mutations in this gene have been associated with Coffin-Lowry syndrome (CLS).
140.	*ERG*	This gene encodes a member of the erythroblast transformation-specific (ETS) family of transcriptions factors. All members of this family are key regulators of embryonic development, cell proliferation, differentiation, angiogenesis, inflammation, and apoptosis.
141.	*BCL11A*	This gene encodes a C2H2 type zinc-finger protein by its similarity to the mouse Bcl11a/Evi9 protein. The corresponding mouse gene is a common site of retroviral integration in myeloid leukemia, and may function as a leukemia disease gene, in part, through its interaction with BCL6.

142.	*SMAD2*	SMAD2 (SMAD Family Member 2) is a Protein Coding gene. Diseases associated withSMAD2 include buschke-ollendorff syndrome and melorheostosis.
143.	*C12orf5*	TIGAR is a recently discovered enzyme that primarily functions as a regulator of glucose breakdown in humancells. In addition to its role in controlling glucose degradation, TIGAR activity can allow a cell to carry outDNA repair, and the degradation of its own organelles.
144.	*PHF20L1*	siRNA-mediated knockdown of PHF20L1 or incubation of a small molecule MBT domain binding inhibitor in cultured cells accelerated the proteasomal degradation of DNMT1. These results demonstrate that the MBT domain of PHF20L1 reads and controls enzyme levels of methylated DNMT1 in cells, thus representing a novel antagonist of DNMT1 degradation.
145.	*MYEF2*	MYEF2 (Myelin Expression Factor 2) is a Protein Coding gene. Among its related pathways are MicroRNAs in cardiomyocyte hypertrophy and Physiological and Pathological Hypertrophy of the Heart.
146.	*PROS1*	This gene encodes a vitamin K-dependent plasma protein that functions as a cofactor for the anticoagulant protease, activated protein C (APC) to inhibit blood coagulation. It is found in plasma in both a free, functionally active form and also in an inactive form complexed with C4b-binding protein.
147.	*SASH1*	The expression levels of SASH1 were strongly reduced in liver cancer tissues compared with adjacent normal tissues. Quantitative methylation analysis by MassArray revealed different CpG sites in SASH1promoter shared similar methylation pattern between liver cancer tissues and adjacent normal tissues
148.	*DIAPH2*	The product of this gene belongs to the diaphanous subfamily of the formin homology family of proteins. This gene may play a role in the development and normal function of the ovaries.
149.	*NOVA1*	NOVA1 (Neuro-Oncological Ventral Antigen 1) is a Protein Coding gene. Diseases associated with NOVA1 include paraneoplastic neurologic disorders and lung cancer.
150.	*CLIC4*	Chloride channels are a diverse group of proteins that regulate fundamental cellular processes including stabilization of cell membrane potential, transepithelial transport, maintenance of intracellular pH, and regulation of cell volume.
151.	*GAD2*	This gene encodes one of several forms of glutamic acid decarboxylase, identified as a major autoantigen in insulin-dependent diabetes. The enzyme encoded is responsible for catalyzing the production of gamma-aminobutyric acid from L-glutamic acid.
152.	*TTN*	This gene encodes a large abundant protein of striated muscle. The product of this gene is divided into two regions, a N-terminal I-band and a C-terminal A-band. The I-band, which is the elastic part of the molecule, contains two regions of tandem immunoglobulin domains on either side of a PEVK region that is rich in proline, glutamate, valine and lysine.
153.	*USP6*	Deubiquitinase with an ATP-independent isopeptidase activity, cleaving at the C-terminus of the ubiquitin moiety. Catalyzes its own deubiquitination. In vitro, isoform 2, but not isoform 3, shows deubiquitinating activity
154.	*CNTNAP2*	This gene encodes a member of the neurexin family which functions in the vertebrate nervous system as cell adhesion molecules and receptors. This protein, like other neurexin proteins, contains epidermal growth factor repeats and laminin G domains.
155.	*ANKRD30B*	ANKRD30B (Ankyrin Repeat Domain 30B) is a Protein Coding gene. Diseases associated with ANKRD30B include breast cancer. An important paralog of this gene is POTEJ.
156.	*SETD7*	Histone methyltransferase that specifically monomethylates 'Lys-4' of histone H3. H3 'Lys-4' methylation represents a specific tag for epigenetic transcriptional activation.
157.	*ARCN1*	This gene maps in a region, which include the mixed lineage leukemia and Friend leukemia virus integration 1 genes, where multiple disease-associated chromosome translocations occur.
158.	*KIAA1609*	KIAA1609 is a 456 amino acid protein that contains one TLD domain and is encoded by a gene that maps to human chromosome 16q24.1. Chromosome 16 encodes over 900 genes in approximately 90 million base pairs, makes up nearly 3% of human cellular DNA and is associated with a variety of genetic disorders.
159.	*KIAA1239*	The sequence of this gene is defined by41 GenBank accessions from 36 cDNA clones, some from brain (seen 10 times), pancreas (4), purified pancreatic islet (4), whole brain (4), total brain (2), germ cell
160.	*ZNF644*	The protein encoded by this gene is a zinc finger transcription factor that may play a role in eye development. Defects in this gene have been associated with high myopia. Three transcript variants encoding two different isoforms have been found for this gene.
161.	*EPC1*	This gene encodes a member of the polycomb group (PcG) family. The encoded protein is a component of the NuA4 histone acetyltransferase complex and can act as both a transcriptional activator and repressor.
162.	*KCNJ2*	Potassium channels are present in most mammalian cells, where they participate in a wide range of physiologic responses. The protein encoded by this gene is an integral membrane protein and inward-rectifier type potassium channel.
163.	*DOCK4*	This gene is a member of the dedicator of cytokinesis (DOCK) family and encodes a protein with a DHR-1 (CZH-1) domain, a DHR-2 (CZH-2) domain and an SH3 domain.

164.	*ANKS6*	This gene encodes a protein containing multiple ankyrin repeats and a SAM domain. It is thought that this protein may localize to the proximal region of the primary cilium, and may play a role in renal and cardiovascular development. Mutations in this gene have been shown to cause a form of nephronophthisis (NPHP16), a chronic tubulo-interstitial nephritis
165.	*EPAS1*	This gene encodes a transcription factor involved in the induction of genes regulated by oxygen, which is induced as oxygen levels fall. The encoded protein contains a basic-helix-loop-helix domain protein dimerization domain as well as a domain found in proteins in signal transduction pathways which respond to oxygen levels. Mutations in this gene are associated with erythrocytosis familial type 4
166.	*TMEM164*	TMEM164 (transmembrane protein 164) is a 297 amino acid protein encoded by a gene mapping to the human X and Y chromosomes. The X and Y chromosomes are the human sex chromosomes. Chromosome X consists of about 153 million base pairs and nearly 1,000 genes.
167.	*PIK3R3*	PI 3-Kinases (phosphoinositide 3-kinases, PI3Ks) are family of lipid kinases capable of phosphorylating the 3'OH of the inositol ring of phosphoinositides. They are responsible for coordinating a diverse range of cell functions including proliferation, cell survival, degranulation, vesicular trafficking and cell migration.
168.	*EHF*	This gene encodes a protein that belongs to an ETS transcription factor subfamily characterized by epithelial-specific expression (ESEs). The encoded protein acts as a transcriptional repressor and may be involved in epithelial differentiation and carcinogenesis.
169.	*ACVR1*	Activins are dimeric growth and differentiation factors which belong to the transforming growth factor-beta (TGF-beta) superfamily of structurally related signaling proteins.
170.	*UBR4*	The protein encoded by this gene is an E3 ubiquitin-protein ligase that interacts with the retinoblastoma-associated protein in the nucleus and with calcium-bound calmodulin in the cytoplasm.
171.	*DTNA*	The protein encoded by this gene belongs to the dystrobrevin subfamily of the dystrophin family. This protein is a component of the dystrophin-associated protein complex (DPC), which consists of dystrophin and several integral and peripheral membrane proteins, including dystroglycans, sarcoglycans, syntrophins and alpha- and beta-dystrobrevin.
172.	*RGS9BP*	The protein encoded by this gene functions as a regulator of G protein-coupled receptor signaling in phototransduction. Studies in bovine and mouse show that this gene is expressed only in the retina, and is localized in the rod outer segment membranes.
173.	*ENAH*	ENAH induces the formation of F-actin rich outgrowths in fibroblasts. Acts synergistically with BAIAP2-alpha and downstream of NTN1 to promote filipodia formation
174.	*DLGAP2*	This gene is biallelically expressed in the brain, however, only the paternal allele is expressed in the testis Alternatively spliced transcript variants encoding different isoforms have been identified
175.	*CLDN19*	The product of this gene belongs to the claudin family. It plays a major role in tight junction-specific obliteration of the intercellular space, through calcium-independent cell-adhesion activity.
176.	*KIAA0226*	The protein encoded by this gene is a negative regulator of autophagy and endocytic trafficking and controls endosome maturation. This protein contains two conserved domains, an N-terminal RUN domain and a C-terminal DUF4206 domain.
177.	*C1orf141,*	Chromosome 1 is the largest human chromosome spanning about 260 million base pairs and making up 8% of the human genome. There are about 3,000 genes on chromosome 1, and considering the great number of genes there are also a large number of diseases associated with chromosome
178.	*C1orf116*	Stickler syndrome, Parkinsons, Gaucher disease and Usher syndrome are also associated with chromosome 1. A breakpoint has been identified in 1q which disrupts the DISC1 gene and is linked to schizophrenia.
179.	*CRAMP1L*	Most normal tissue showed weak to moderate nuclear/nucleolar and occasional cytoplasmic staining. Leydig cells exhibited strong cytoplasmic staining.
180.	*UBR4*	The protein encoded by this gene is an E3 ubiquitin-protein ligase that interacts with the retinoblastoma-associated protein in the nucleus and with calcium-bound calmodulin in the cytoplasm.
181.	*CDH23*	This gene is a member of the cadherin superfamily, whose genes encode calcium dependent cell-cell adhesion glycoproteins. The encoded protein is thought to be involved in stereocilia organization and hair bundle formation.
182.	*TEAD1*	This gene encodes a ubiquitous transcriptional enhancer factor that is a member of the TEA/ATTS domain family. This protein directs the transactivation of a wide variety of genes and, in placental cells, also acts as a transcriptional repressor.
183.	*TRIB2*	This gene encodes one of three members of the Tribbles family. The Tribbles members share a Trb domain, which is homologous to protein serine-threonine kinases, but lacks the active site lysine and probably lacks a catalytic function.
184.	*SLC25A16*	The encoded protein is localized in the inner membrane and facilitates the rapid transport and exchange of molecules between the cytosol and the mitochondrial matrix space. This gene has a possible role in Graves' disease
185.	*FZD8*	This gene is highly expressed in two human cancer cell lines, indicating that it may play a role in several types of cancer. The crystal structure of the extracellular cysteine-rich domain of a similar mouse protein has been determined.

186.	*GIYD1*	Two identical copies of this gene are located on the p arm of chromosome 16 due to a segmental duplication; this record represents the more centromeric copy.
187.	*GIYD2*	This gene encodes a protein that is an important regulator of genome stability. The protein represents the catalytic subunit of the SLX1-SLX4 structure-specific endonuclease, which can resolve DNA secondary structures that are formed during repair and recombination processes.
188.	*HS3ST2*	Heparan sulfate biosynthetic enzymes are key components in generating a myriad of distinct heparan sulfate fine structures that carry out multiple biologic activities. The enzyme encoded by this gene is a member of the heparan sulfate biosynthetic enzyme family.
189.	*FGFR3*	This gene encodes a member of the fibroblast growth factor receptor (FGFR) family, with its amino acid sequence being highly conserved between members and among divergent species. FGFR family members differ from one another in their ligand affinities and tissue distribution.
190.	*NTRK3*	This gene encodes a member of the neurotrophic tyrosine receptor kinase (NTRK) family. This kinase is a membrane-bound receptor that, upon neurotrophin binding, phosphorylates itself and members of the MAPK pathway.
191.	*NHS*	This gene encodes a protein containing four conserved nuclear localization signals. The encoded protein functions in eye, tooth, craniofacial and brain development, and it can regulate actin remodeling and cell morphology.
192.	*QSER1*	. As far as is known, this is the smallest deletion as-yet described encompassing the WT1 gene and was detected only once in a total of 32 Portuguese patients with isolated uni- or bilateral cryptorchidism.
193.	*BCL2*	This gene encodes an integral outer mitochondrial membrane protein that blocks the apoptotic death of some cells such as lymphocytes. Constitutive expression of BCL2, such as in the case of translocation of BCL2 to Ig heavy chain locus, is thought to be the cause of follicular lymphoma.
194.	*SFMBT2*	By sequencing clones obtained from a size-fractionated fetal brain cDNA library, cloned SFMBT2, which they designated KIAA1617.
195.	*DCX*	This gene encodes a member of the doublecortin family. The protein encoded by this gene is a cytoplasmic protein and contains two doublecortin domains, which bind microtubules. In the developing cortex, cortical neurons must migrate over long distances to reach the site of their final differentiation.
196.	*C5orf21*	C5orf21 gene contains five splice variants and it is identified for the first time at the level of mRNA. The changes of C5orf21 gene expression are correlated with diabetic macroangiopathy
197.	*XYLT1*	This locus encodes a xylosyltransferase enzyme. The encoded protein catalyzes transfer of UDP-xylose to serine residues of an acceptor protein substrate.
198.	*TEAD1*	This gene encodes a ubiquitous transcriptional enhancer factor that is a member of the TEA/ATTS domain family. This protein directs the transactivation of a wide variety of genes and, in placental cells, also acts as a transcriptional repressor.
199.	*PTENP1*	The phosphatase and tensin homolog (PTEN) gene encodes a tumor suppressor that functions by negatively regulating the AKT/PKB signaling pathway.
200.	*TNRC6B*	TNRC6B has been shown to interact with EIF2C2. It is also known to associate with argonaute proteins and has been shown to be required for miRNA-guided gene silencing in HeLa cells.
201.	*HSPA12A*	We found that HSPA12A (but not HSPA12B) is highly expressed in the human brain and shows a neuron- and region-specific transcript distribution, with strongest expression in the frontal and occipital cortical regions.
202.	*ERBB4*	This gene is a member of the Tyr protein kinase family and the epidermal growth factor receptor subfamily. It encodes a single-pass type I membrane protein with multiple cysteine rich domains, a transmembrane domain, a tyrosine kinase domain, a phosphotidylinositol-3 kinase binding site and a PDZ domain binding motif.
203.	*ZFAND5*	Involved in protein degradation via the ubiquitin-proteasome system. May act by anchoring ubiquitinated proteins to the proteasome. Plays a role in ubiquitin-mediated protein degradation during muscle atrophy.
204.	*NFIB*	NFIB (Nuclear Factor I/B) is a Protein Coding gene. Diseases associated with NFIB include lipoma of colon and polymorphous low-grade adenocarcinoma. NFIB is a potential target for estrogen receptor-negative breast cancers.
205.	*BCL7B*	This gene encodes a member of the BCL7 family including BCL7A, BCL7B and BCL7C proteins. This member is BCL7B, which contains a region that is highly similar to the N-terminal segment of BCL7A or BCL7C proteins.
206.	*MAFB*	The protein encoded by this gene is a basic leucine zipper (bZIP) transcription factor that plays an important role in the regulation of lineage-specific hematopoiesis. The encoded nuclear protein represses ETS1-mediated transcription of erythroid-specific genes in myeloid cells. This gene contains no introns.
207.	*F11R*	Tight junctions represent one mode of cell-to-cell adhesion in epithelial or endothelial cell sheets, forming continuous seals around cells and serving as a physical barrier to prevent solutes and water from passing freely through the paracellular space..

208.	*CCDC6*	This gene encodes a coiled-coil domain-containing protein. The encoded protein is ubiquitously expressed and may function as a tumor suppressor. A chromosomal rearrangement resulting in the expression of a fusion gene containing a portion of this gene and the intracellular kinase-encoding domain of the ret proto-oncogene is the cause of thyroid papillary carcinoma.
209.	*C6orf103*	C6orf103, also known as CAPN7L (calpain-7-like protein), is a 1,667 amino acid protein that is most likely catalytically inactive. C6orf103 belongs to the peptidase C2 family and contains one calpain catalytic domain and one IQ domain.
210.	*PSD3*	PSD3 (Pleckstrin And Sec7 Domain Containing 3) is a Protein Coding gene. Diseases associated with PSD3 include hepatocellular carcinoma.
211.	*PARK2*	The precise function of this gene is unknown; however, the encoded protein is a component of a multiprotein E3 ubiquitin ligase complex that mediates the targeting of substrate proteins for proteasomal degradation.
212.	*IYD*	Mutations in this gene cause congenital hypothyroidism due to dyshormonogenesis type 4, which is also referred to as deiodinase deficiency, or iodotyrosine dehalogenase deficiency, or thyroid hormonogenesis type 4. Alternative splicing results in multiple transcript variants.
213.	*FAM91A1*	FAM91A1 (Family with Sequence Similarity 91, Member A1) is a Protein Coding gene. Among its related pathways are Gastric cancer network 2.
214.	*KIAA2026*	KIAA2026 is a 2,103-amino acid protein that exists as two alternatively spliced isoforms and is encoded by a gene that maps to human chromosome 9p24.1. Consisting of about 145 million bases, chromosome 9 makes up approximately 4% of the human genome and encodes nearly 900 genes.
215.	*TTN*	This gene encodes a large abundant protein of striated muscle. The product of this gene is divided into two regions, a N-terminal I-band and a C-terminal A-band. The I-band, which is the elastic part of the molecule, contains two regions of tandem immunoglobulin domains on either side of a PEVK region that is rich in proline, glutamate, valine and lysine.
216.	*TRIB2*	This gene encodes one of three members of the Tribbles family. The Tribbles members share a Trb domain, which is homologous to protein serine-threonine kinases, but lacks the active site lysine and probably lacks a catalytic function.
217.	*HOXA1*	This gene is part of the A cluster on chromosome 7 and encodes a DNA-binding transcription factor which may regulate gene expression, morphogenesis, and differentiation. The encoded protein may be involved in the placement of hindbrain segments in the proper location along the anterior-posterior axis during development.
218.	*C4orf16*	Tissue expression: ascites; blood; bone; bone marrow; brain; embryonic tissue; eye; heart; intestine; kidney; liver; lung; lymph node; mammary gland; muscle
219.	*CTDSPL*	Expression of CTDSPL in cancer tissue. Methylation of WNT9A/CD558500 and CTDSPL/BC040563 promoters occurs frequently in primary colon cancers and acute lymphoid leukemias
220.	*LIF*	It is involved in the induction of hematopoietic differentiation in normal and myeloid leukemia cells, induction of neuronal cell differentiation, regulator of mesenchymal to epithelial conversion during kidney development, and may also have a role in immune tolerance at the maternal-fetal interface.
221.	*BAZ2A*	BAZ2 is involved in epigenetic alterations in prostate cancer and its overexpression predicts disease recurrence.
222.	*EIF2C2*	This gene encodes a member of the Argonaute family of proteins which play a role in RNA interference. The encoded protein is highly basic, and contains a PAZ domain and a PIWI domain.
223.	*CAMTA1*	CAMTA1 (Calmodulin Binding Transcription Activator 1) is a Protein Coding gene. Diseases associated with CAMTA1 include cerebellar ataxia.
224.	*C1orf121*	C1orf121 is also known as PNS4. PNAS4 induced apoptosis and arreste cell cycle in S phase in A549 human lung adenocarcinoma cells.
225.	*INSM1*	Insulinoma-associated 1 (INSM1) gene is intronless and encodes a protein containing both a zinc finger DNA-binding domain and a putative prohormone domain. This gene is a sensitive marker for neuroendocrine differentiation of human lung tumors.
226.	*CDC25A*	CDC25A is a member of the CDC25 family of phosphatases. CDC25A is required for progression from G1 to the S phase of the cell cycle. It activates the cyclin-dependent kinase CDC2 by removing two phosphate groups.
227.	*CYP26B1*	This gene encodes a member of the cytochrome P450 superfamily. The cytochrome P450 proteins are monooxygenases which catalyze many reactions involved in drug metabolism and synthesis of cholesterol, steroids and other lipids.
228.	*TRIB1*	TRIB1, a member of the Trib family of serine/threonine kinase-like proteins, is an essential factor for prostate cancer cell growth and survival.
229.	*ZBTB7A*	Significantly, reduced ZBTB7A expression correlates with up-regulation of the glycolyticgenes and poor survival in colon cancer patients.
230.	*ST5*	This gene was identified by its ability to suppress the tumorigenicity of Hela cells in nude mice. The protein encoded by this gene contains a C-terminal region that shares similarity with the Rab 3 family of small GTP binding proteins.

231.	*JMJD3*	JMJD3 as an epigenetic regulator in development and disease. role and mechanisms of JMJD3 has been extensively studied for their involvement in development, cell plasticity, immune system, neurodegenerative disease, and cancer
232.	*JARID1A*	Knockdown of JARID1A promoted osteogenic differentiation of human adipose-derived stromal cells in vitro and in vivo and resulted in marked increases of mRNA expression of osteogenesis-associated genes such as alkaline phosphatase (ALP), osteocalcin (OC), and osterix (OSX).
233.	*TNRC6B*	Single nucleotide polymorphisms in the TNRC6B gene is associated with uterine fibroid.
234.	*CCND2*	Cyclins function as regulators of CDK kinases. Different cyclins exhibit distinct expression and degradation patterns which contribute to the temporal coordination of each mitotic event. This cyclin forms a complex with CDK4 or CDK6 and functions as a regulatory subunit of the complex, whose activity is required for cell cycle G1/S transition.
235.	*FRAS1*	This gene encodes an extracellular matrix protein that appears to function in the regulation of epidermal-basement membrane adhesion and organogenesis during development. Mutations in this gene cause Fraser syndrome, a multisystem malformation that can include craniofacial, urogenital and respiratory system abnormalities. Alternative splicing results in multiple transcript variants.
236.	*RUNX1*	Core binding factor (CBF) is a heterodimeric transcription factor that binds to the core element of many enhancers and promoters. The protein encoded by this gene represents the alpha subunit of CBF and is thought to be involved in the development of normal hematopoiesis.
237.	*HEMK1*	HEMK1 (HemK Methyltransferase Family Member 1) is a Protein Coding gene. Diseases associated with HEMK1 include acrocephalosyndactylia.
238.	*G6PC*	Mutations in this gene cause glycogen storage disease type I (GSD1). This disease, also known as von Gierke disease, is a metabolic disorder characterized by severe hypoglycemia associated with the accumulation of glycogen and fat in the liver and kidneys.
239.	*E2F2*	The protein encoded by this gene is a member of the E2F family of transcription factors. The E2F family plays a crucial role in the control of cell cycle and action of tumor suppressor proteins and is also a target of the transforming proteins of small DNA tumor viruses.
240.	*XYLT1*	The encoded protein catalyzes transfer of UDP-xylose to serine residues of an acceptor protein substrate. This transfer reaction is necessary for biosynthesis of glycosaminoglycan chains. Mutations in this gene have been associated with increased severity of pseudoxanthoma elasticum.
241.	*GRM4*	L-glutamate is the major excitatory neurotransmitter in the central nervous system and activates both ionotropic and metabotropic glutamate receptors. Glutamatergic neurotransmission is involved in most aspects of normal brain function and can be perturbed in many neuropathologic conditions.
242.	*CHST3*	This gene encodes an enzyme which catalyzes the sulfation of chondroitin, a proteoglycan found in the extracellular matrix and most cells which is involved in cell migration and differentiation. Mutations in this gene are associated with spondylepiphyseal dysplasia and humerospinal dysostosis.
243.	*ZEB2*	The protein encoded by this gene is a member of the Zfh1 family of 2-handed zinc finger/homeodomain proteins. It is located in the nucleus and functions as a DNA-binding transcriptional repressor that interacts with activated SMADs.
244.	*FBXW7*	This gene encodes a member of the F-box protein family which is characterized by an approximately 40 amino acid motifs, the F-box. The F-box proteins constitute one of the four subunits of ubiquitin protein ligase complex called SCFs (SKP1-cullin-F-box), which function in phosphorylation-dependent ubiquitination.
245.	*KIAA1432*	KIAA1432 is a novel gene activated by amplification in Small Cell Lung Cancers.
246.	*RPS6KB1*	This gene encodes a member of the ribosomal S6 kinase family of serine/threonine kinases. The encoded protein responds to mTOR (mammalian target of rapamycin) signaling to promote protein synthesis, cell growth, and cell proliferation. Activity of this gene has been associated with human cancer.
247.	*TSGA14*	This gene encodes a centrosomal and microtubule-binding protein which is predicted to have two coiled-coil domains and a rhodanese domain. In human retinal pigment epithelial cells, the protein localized to centrioles and cilia.
248.	*JHDM1D*	The central role of the histone demethylase JHDM1D in the regulation of tumor associated genes in bone tumor-related cells.
249.	*PAPD5*	PAPD5 Plays a role in replication-dependent histone mRNA degradation. May be involved in the terminal uridylation of mature histone mRNAs before their degradation is initiated.
250.	*CRKL*	This gene encodes a protein kinase containing SH2 and SH3 (src homology) domains which has been shown to activate the RAS and JUN kinase signaling pathways and transform fibroblasts in a RAS-dependent fashion.
251.	*CDYL*	Chromodomain Y is a primate-specific Y-chromosomal gene family expressed exclusively in the testis and implicated in infertility. Although the Y-linked genes are testis-specific, this autosomal gene is ubiquitously expressed.
252.	*WIPF1*	This gene encodes a protein that plays an important role in the organization of the actin cytoskeleton. The encoded protein binds to a region of Wiskott-Aldrich syndrome protein that is frequently mutated in Wiskott-Aldrich syndrome, an X-linked recessive disorder.

253.	*PKN2*	PKN2 an AGC kinase of the PKN family. A PKC-related serine/threonine-protein kinase and Rho/Rac effector protein. Plays a role in the regulation of cell cycle progression, actin cytoskeleton assembly, cell migration, cell adhesion, tumor cell invasion and transcriptional activation the cytoplasm.
254.	*RANBP9*	This gene encodes a protein that binds RAN, a small GTP binding protein belonging to the RAS superfamily that is essential for the translocation of RNA and proteins through the nuclear pore complex
255.	*ESRRG*	All members of the ESRR family share an almost identical DNA binding domain, which is composed of two C4-type zinc finger motifs. and some of which encode protein isoforms differing in the N-terminal region.
256.	*CCNJ*	Acts as a tumor suppressor in breast tumorigenesis via its novel direct targets ENPEP, CK2-alpha, CCNJ, and MEGF9.

Table 3: Name of potential functional genes with genomic involvement.

Conclusion

Finally, this theory reveals that study of 26 potential high confidence miRNAs from miRBase These potential miRNAs target 610 genes in human. 256 genes perform a potential function to regulate some specific activities. These genes are : *LIN28B, CLCN5, FOXP2, FBN1, YOD1, C14orf28, CPEB2, CCND2, PKN2, BZW1, ZNF512B, TGFBR1, PTPRD, PBX2, SMARCAD1, PTCH1, CPEB1, ABCC5, IGF2BP2, CPEB3, PBX3, CPEB4, GDF6, DLC1, NAT12, GATM, COL1A2, ABL2, COL27A1, WDFY4, ZFYVE26, RSPO4, MEF2D, FOXP4, TSPAN18, PCNT, ZMIZ1, SPNS2, SPNS2, SPN, SGIP1, LRIG2, LRIG3, BPTF, CCNJ, B3GNT7, BCAT1, PCDH19, GAN, MAPK8, COL3A1, HDX, IDE, GLIS3, WDR35, ANKRD57, IVNS1ABP, RNF38, NFIB, TP53INP1, EIF5, KIAA1853, CLIC2, S100A7A, OPHN1, C3orf64, KIAA0408, METTL7A, C20orf117, FGFR1, PVRL1, ZMIZ1, PPARD, TRDMT1, PCNT, MECP2, KIAA1522, MYT1, C11orf57, ABHD6, SH3TC2, LIN28, ANPEP, IRF4, ENPP1, MLF2, NIPA1, SUV39H1, STARD13, SCN2B, ALPK3, SLC35A4, TSEN54, TRIAP1, MSI1, FAM78A, IER3IP1, C14orf43, KLHL6, MLL, CPEB2, VPS13B, NR5A2, FOXP1, LDLRAP1, EPB41L3, SRGAP3, STC1, KCNA1, NTRK3, ZNF395, KIAA1045, HPS1, FAM86C, AFF2, KIAA1529, CECR6, ZNF783, ARC, RAB43, IKZF1, ZNF135, ELAVL4, LAMP2, AMMECR1, VGLL3, XKR6, CRB1, ADAM22, ST13, FAM10A6, ST13P7, MYO5A, PSD3, KIAA2018, NEUROD1, EXT1, RPS6KA3, ERG, BCL11A, SMAD2, C12orf5, PHF20L1, MYEF2, PROS1, SASH1, DIAPH2, NOVA1, CLIC4, GAD2, TTN, USP6, CNTNAP2, ANKRD30B, SETD7, ARCN1, KIAA1609, KIAA1239, ZNF644, EPC1, KCNJ2, DOCK4, ANKS6, EPAS1, TMEM164, PIK3R3, EHF, ACVR1, UBR4, DTNA, RGS9BP, ENAH, DLGAP2, CLDN19, KIAA0226, C1orf141, C1orf116, CRAMP1L, UBR4, CDH23, TEAD1, TRIB2, SLC25A16, FZD8, GIYD1, GIYD2, HS3ST2, FGFR3, NTRK3, NHS, QSER1, BCL2, SFMBT2, DCX, C5orf21, XYLT1, TEAD1, PTENP1, TNRC6B, HSPA12A, ERBB4, ZFAND5, NFIB, BCL7B, MAFB, F11R, CCDC6, C6orf103, PSD3, PARK2, IYD, FAM91A1, KIAA2026, TTN, TRIB2, HOXA1. C4orf16, CTDSPL, LIF, BAZ2A, EIF2C2, CAMTA1, C1orf121, INSM1, CDC25A, CYP26B1, TRIB1, ZBTB7A, ST5, JMJD3, JARID1A, TNRC6B, CCND2, FRAS1, RUNX1, HEMK1, G6PC, E2F2, XYLT1, GRM4, CHST3, ZEB2, FBXW7, KIAA1432, RPS6KB1, TSGA14, JHDM1D, PAPD5, CRKL, CRKL, WIPF1, PKN2, RANBP9, ESRRG, CCNJ.*

These genes play an important role in the regulation of different diseases. Functions of different genes are shown in Table 3. In conclusion, we can say that if we control malfunction of protein at transcriptional level, then we will be able to control the disease and thus save patients suffering from different type of disease.

Acknowledgement

The corresponding authors wish to thank Er. NEELESH YADAV, Scientist and Coordinator, Bioinformatics Centre, Forest Research Institute, Dehradun, INDIA for providing valuable scientific platform and giving us support for scientific favorable rights. I am also thankful to DBT center for providing necessary facilities.

References

1. Saxena VL, Dwivedi A (2013) In silico identification of miRNAs and their target prediction from Japanese encephalitis. J Bioinform Seq Anal 5: 25-33.
2. Dubey A, Shanker U, Kalra SS, Yadav N (2013) Viral micro RNA analysis via the bioinformatics approaches basis on their statistics values. AJBCPS 1: 42-66.
3. Dubey A, Agarwal SP, Yadav N, Kumar R (2016) Analysis of Caenorhabditis elegans via bioinformatics approaches basis on their precursors statistics values. Next Generat Sequenc Applic 3: 126.
4. Dubey A, Kalra SS, Trivedi N (2013) Computational prediction of miRNA in Gmelina arborea and their role in human metabolomics. J Biosci Bioeng 1: 62-74.
5. Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75: 843-854.
6. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001) Identification of novel genes coding for small expressed RNAs. Science 294: 853-858.
7. Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science 294: 858-862.
8. Llave C, Kasschau KD, Rector MA, Carrington JC (2002) Endogenous and silencing-associated small RNAs in plants. Plant Cell 14: 1605-1619.
9. Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP (2002) MicroRNAs in plants. Genes Dev 16: 1616-1626.
10. Hutvagner G, Zamore PD (2002) RNAi: Nature abhors a doublestrand. Curr Opin Genet Dev 12: 225-232.
11. Schauer SE, Jacobsen SE, Meinke DW, Ray A (2002) DICER-LIKE1: Blind men and elephants in Arabidopsis development. Trends Plant Sci 7: 487-491.
12. Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, et al. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature 403: 901-906.
13. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, et al. (2007) Molecular biology of the cell (5th edn) Garland Science, USA. pp. 1321.
14. Mango SE (2017) The C. elegans pharynx: A model for organogenesis.
15. Kozomara A, Griffiths-Jones S (2014) miRBase: Annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 42: D68-D73.

16. Kozomara A, Griffiths-Jones S (2011) miRBase: Integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 39: D152-D157.
17. Griffiths-Jones S, Saini HK, Van Dongen S, Enright AJ (2008) miRBase: Tools for microRNA genomics. Nucleic Acids Res 36: D154-D158.
18. Griffiths-Jones S, Grocock RJ, Van Dongen S, Bateman A, Enright AJ (2006) miRBase: MicroRNA sequences, targets and gene nomenclature. Nucleic Acids Res 34: D140-D144.
19. Alexiou P, Maragkakis M, Papadopoulos GL, Reczko M, Hatzigeorgiou AG (2009) Lost in translation: An assessment and perspective for computational microRNA target identification. Bioinformatics 25: 3049-3055.
20. Paraskevopoulou MD, Georgakilas G, Kostoulas N, Vlachos IS, Vergoulis T, et al. (2013) DIANA-microT web server v5.0: Service integration into miRNA functional analysis workflows. Nucleic Acids Res 4: W169-173.