

Phylogenetic Lineage History Analysis of MIR-34A Gene in Type 2 Diabetes Mellitus Patients from Khyber Pakhtunkhwa, Pakistan

Mudusra Mustafa¹, Muhammad Khan¹, Inam Ullah¹, Sajjad Sajjad¹, Mohamed A. Al Abboud², A. El-Shabasy² and Mohamed Mostafa Soliman^{2,3}

¹Biotechnology and Genetic Engineering Department Hazara University 21300, Mansehra, Pakistan

²Department of Biology, College of Science, Jazan University, P.O. Box. 114, Jazan 45142, Saudi Arabia

³Department of Zoology, Faculty of Science, Ain Shams University, Cairo, PO Box 11566, Egypt

*Corresponding author:

A. El-Shabasy,
Department of Biology, College of Science,
Jazan University, P.O. Box. 114, Jazan 45142,
Saudi Arabia

Received: 18 July 2026

Accepted: 28 July 2026

Published: 03 Aug 2026

J Short Name: JCMI

Copyright:

©2026 A. El-Shabasy. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and build upon your work non-commercially

Keywords:

Diabetes Mellitus; MicroRNA; Red Blood Cells; Jukes-Cantor Model; Confidence Value

Citation:

A. El-Shabasy, Phylogenetic Lineage History Analysis of MIR-34A Gene in Type 2 Diabetes Mellitus Patients from Khyber Pakhtunkhwa, Pakistan. *Jour of Clin and Medi Images®* 2026; V9(1): 1-16

1. Abstract

MicroRNA-34a (MiR-34a) is a conserved tumour suppressor gene involved in pancreatic development and function. The genetic variations in MIR34-a may alter the expression and functions leading to disease susceptibility including Type 2 Diabetes (T2DM). The selected population of Khyber Pakhtunkhwa shows a unique genetic landscape, however the phylogenetic structure T2DM related alleles in the region are uncharacterized. This study aimed to investigate the phylogenetic relationships and sequence variation of MIR34A gene among T2DM Patients from different districts of KP, Pakistan, to trace potential lineage identify conserved variable genomic regions. Blood samples were collected from 500 clinically diagnosed T2DM patients across KP province. Genomic DNA was extracted and MIR34A locus was amplified by PCR. A subset of patients (n=9), from diverse districts were selected for Sanger sequencing. The Phylogenetic analysis was performed using Maximum Parsimony Method in MEGA11 software with 1000 bootstrap replicates sequence alignment identified conserved region, variable sites, singletons and parsimony-informative (Pi) sites. Sequencing analysis confirmed significant phylogenetic structured among KP T2DM patients. The parsimony tree revealed distinct genotypic clustering. Conserved region was identified which is critical for miR-34a function and calculated variable sites. These include singletons and Pi-informative sites, which are population specific. Genotype classification based on sequence morphology showed one lineage as predominant while others (Genotype M7) were more divergent. The study demonstrates that for the first time a phylogenetic structure in MIR34A among T2DM patients in KP population. The finding suggests that potential

evolutionary lineage and highlight specific variable regions that may affect genetic susceptibility to T2DM in ethnically distinct group, providing the foundation for future biomarker for functional study.

2. Introduction

Diabetes mellitus is a chronic metabolic disorder characterized by persistent hyperglycemia resulting from defects in insulin secretion and actions [1]. It is a chronic elevation of glucose and fatty acid, characterized by Glucogenesis. Diabetes mellitus causes hyperglycaemia due to insulin deficiency or defect in its receptors (Irvine, 1977). When pancreatic tissue destroys is called "pancreatogenic" diabetes (Cudworth 1976; National Diabetes Data Group 1979). During phosphorylation insulin deficiency synthesizes Glucose 1 Phosphate G-1-P instead of glucose-6-phosphate G-6-P [2]. In Diabetes GLUT4 is the insulin regulatable transporter [3]. Insulin-receptor (InsR) allow glucose molecules to migrate into cell surface, inversely when this Insulin acts on hepatic insulin receptors it suppress glucose production [4].

MicroRNA-34A higher expression level in serum is biomarker for T2DM [5]. MicroRNA-34a is conserved in humans, it three homologs are MiR-34a,b,c on chromosome 1 [6] within inter-genic regions in introns and exons of noncoding genes. The mature MicroRNA-34a is generated by removal of 30 kb intron [7]. RNA polymerase II/III transcribe Mir-34 gene into Primary Mir-34 then polyadenylated and capped by 7-methylguanosine, nuclear cleavage form Precursor Mir-34 by Drosha and DGCR8. The cleavage in the cytoplasm by RNase Dicer make Mir-34 duplex, which is mature Mir-34, found in MirNA-induced silencing

complex (MirISC) [8]. MicroRNAs transcribed from introns are mitrons [9], processed by microprocessor (Drosha and DGCR 8) to form pre-miRNAs. The precursor miRNAs exported into the cytoplasm by exportin-5, processed into miRNA duplex (TRBP, TAR, RNA Binding Protein). The spliceosome splices mirtrons to form looped, then debranched stem loop pre-miRNA structure [10-12].

A lineage is a diagram showing evolutionary relationships, like a family tree for life, where branches represent different evolutionary paths (lineages) from common ancestors, revealing how species diverge over time, built using genetic/physical traits to map shared ancestry from root to tips (Scott and Baum, 2016). Molecular Phylogenetics is aspect of bioinformatics that uses molecular data in phylogenetic trees [13]. These trees are two-dimensional branches related to different groups and endpoints represent present-day Genotypes. The Node connects two branches represents common ancestor [14]. Distance-based tree construction methods involve Parsimony tree method as an example. Parsimony is character-based method have least number of evolutionary changes. The multiple sequence alignment identifies potential positions in the sequences.

MEGA 11 (Molecular Evolutionary Genetic analysis) is a window-based program was used to examine the Genetic Distance and nucleotide variation on the bases of the Kimura 2 parameter (K2P) model and Parsimony tree was constructed in MEGA 11 with 1000 replicates the bootstrap value and with rates and patterns of the gamma distribution method. The pairwise distances represents the evolutionary divergence the larger distance (0.990) between species and while the less distance (0.290) between the studied species and Query to match. The tree forms branches and sub-branches called clades. The studied sample found one clade and rest species were found in disperse form, Query sometimes kept in out group with sister clade. In Mir-34A Sample M1-M9 tree 100 Taxa were selected from NCBI page. The resolution rate was based on the bootstrap value at nodes and each clade. The clade with bootstrap value 50% was all collapsed and Bootstrap greater than 50% implicated to infer phylogenetic relationship [15].

In Phylogeny tree analysis crustal Alignment is made first this gives results like two types of regions shown in Mega 11. The first region is Conserved Region and Second region is Variable region. Variable regions are of two types Pi Ψ and Singleton. The Pi Ψ is PI-Informative are variables are those codes have physical existence. Where Singleton variables have SNPs and have no physical existence as disease or variation. Therefore, the study was designed to perform the phylogenetic analysis of MIR34A gene in T2DM patients from KP. We aimed to sequence the gene, construct phylogenetic tree and analyse sequence conservation and variation to discover protentional evolutionary relationship and identify genotype clusters that correlate with population specific susceptibility. This work will contribute to understand the genetic epidemiology of T2DM in Pakistan and highlight the regions of 34a gene which is critical for future

functional and diagnostic applications.

3. Materials and Methods

The materials and methods apply in this research include intensive studies of the Diabetes Mellitus, MicroRNA, DNA extraction chemicals and procedure and Primer designing of gene. This study started 1st Oct. 2023 and finished at 1st Oct. 2024. The Question forms are written and printed having disease existence and per human organ pain or defect presence. T2DM patients provided their 2cc blood for research study, saved in EDTA tube labeled and stored at -10°C. The samples were taken from various districts Kohat, Bhattagram, Balakot, Mansehra of KPK province of Pakistan. In this study 10 samples of T2DM were amplified and sequenced. Blood samples were collected from 500 T2DM patients. Genomic DNA was extracted and PCR amplified product of a subset of patients (n=9), were sequenced by Sanger sequencing. The trees identify the origins of pathogens and to track the transmission of hereditary diseases. Phylogenetic trees can also be used in forensics and crime scenes. The evolutionary relationships between different species, diversity and distribution of species in ecosystems are found out. Ethical approval statement is mentioned at the end of the manuscript. The consent was informed to patient. Written information was kept from Kath Mansehra Hospital Laboratory Blood Collection Staff was made witnessed for documented consent.

3.1 DNA Extraction and Quantification

The extraction of DNA from blood starts with lysis of blood includes 400μl lysis solution (BME+Proteinase K) and 200 μl Dilution buffer (50XTAE +NaCl). The incubation of blood takes 3 hours at 65 °C in shaker for proper mixing followed by addition of 400 μl PCI (Phenol, Chloroform and Iso-amyl alcohol) in each sample. Then keep it on room temperature for 30 mins and centrifuge at 10000rpm for 20 mins. The liquid divides into 2 layers upper aqueous clear liquid contains DNA, middle organic layer and lower red blood cells. Pick the upper DNA liquid layer and add 200μl ice cold iso-propanol which precipitates the DNA out of solution and Freeze this again at -4 °C for overnight. Next day centrifuge it again for 15 mins at 8000rpm. DNA pellet to the wall of 2nd Eppendorf is either seen in translucent forms. Add 200μl Ethanol and keep it at room temperature for 10 mins, then centrifuge again, and discard upper liquid with light hand. Then invert tube for overnight drying and add 10μl sterile water and stored at lowest temperature of -10°C in Freezer.

3.2. PCR Methods and Conditions

The PCR is machine temperature is adjusted to 94°C for 1 minute at Denaturation and 60°C for Annealing then 72°C extension. PCR 2X Master Mix of wiz BioSolutions is mixed with 1μl forward and Reverse Primer and 1μl ddH₂O and 0.8 μl Template DNA in each PCR tube is placed in PCR Machine. The selected genes for mutation study Mir-34a which is 120bp nucleotides in length were amplified 35 cycles. The selected samples were sent to Microgen Korea for Sanger sequencing.

3.3. Phylogenetic Tree construction

The steps involve are the Selection of molecular marker depends on purpose of the study (Mount, 2001). Multiple sequence alignment programs such as T-Coffee can be used. Selection of a model of evolution for nucleotides are the Jukes-Cantor (JC) model and Kimura's two-parameter model. Construction of the phylogenetic tree involve Distance-based method rely on dissimilarity between sequences, while character-based method rely on molecular sequences tracing the character states of the common ancestor. Assessment of the reliability of the tree topology is tested by statistical method called bootstrapping. It involves repeatedly resampling the initial sequence data to generate multiple subsets of derived sequences, referred to as bootstrap samples. Interior branches that are accurately predicted by the new tree are assigned a value of 1 calculated as the bootstrap

value or confidence value. A bootstrap value of > 95 indicates accurate topology, written on the branches of the phylogenetic tree. Besides bootstrapping, other resampling strategies like Jack-knifing and Bayesian Simulation can also be used.

4. Results

The results of Genomics DNA Extraction and PCR of MicroRNA-34A are shown in (Figure 1-3) (Tables 1,2).

4.1. Genotype M1 MIR-34A Phylogenetic Tree

Individual M1 sample has diabetes, weight is 55kg, and experiencing stroke belongs to Mansoura with family history of diabetes. The NCBI analysis showed Genotype 1, sample M1 query cover is 91%, percent identity 93.85% and an E value of 0.0 with Homo Sapiens Chromosome 1 clone CH17-31M (Figure 4) (Table 3, 4).



Figure 1: Gel Pic Genomics samples M1-M9

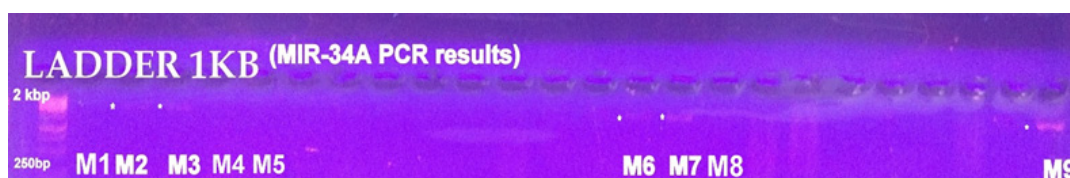


Figure 2: Gel PCR results M1-M9 for Sanger Sequencing

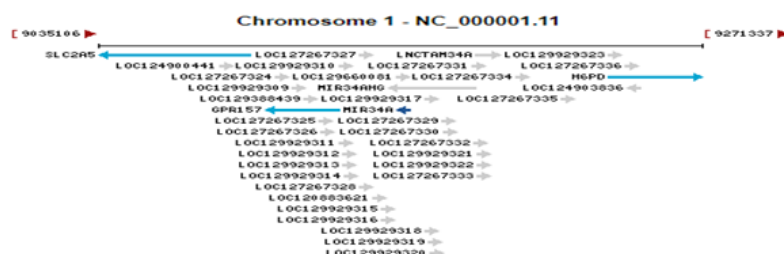


Figure 3: Mir-34a gene Location in Genome Data Viewer

Table 1: Pi-informative detailed Molecular characterization of M1-M9 genotype and conserved regions

Sample	Taxa Selected	Sequence Length	Conserved Regions	Variable Region	Pi. Informative	Singleton
M1	103	1826	958	862	615	243
M2	94	1956	1229	603	440	160
M3	103	1825	949	839	578	255
M4	183	2207	992	1095	695	344
M5	114	1970	834	1037	674	168
M6	102	2011	1074	934	610	321
M7	102	664	365	299	287	12
M8	101	2085	1251	704	450	252
M9	138	1850	554	1281	821	253

**Table 2:** NCBI Query cover, E-Value and percent Identity with closely similar Genotype M1-M9

Sample/Genotype	Nearest Resemblance	Query Cover%	E-Value	Percent Identity (%)
M1	Homo sapiens chromosome 1 clone CH17-31M	91	0.0	93.85
M2	Homo sapiens chromosome 1 clone CH17-31M	99	0.0	92.87
M3	Homo sapiens chromosome 1 clone CH17-31M	93	0.0	94.45
M4	Homo sapiens chromosome 1 clone CH17-31M	99	0.0	92.92
M5	Homo sapiens chromosome 1 clone CH17-31M	100	0.0	96.00
M6	Homo sapiens chromosome 1 clone CH17-31M	93	0.0	92.70
M7	Homo sapiens chromosome 1 clone CH17-31M	62	2e-152	92.13
M8	Homo sapiens chromosome 1 clone CH17-31M	99	0.0	94.75
M9	Homo sapiens chromosome 1 clone CH17-31M	90	0.0	94.33

Table 3: M1 Pi-informative detailed Molecular characterization of M1 genotype and conserved regions

Sample/Genotype	Taxa Selected	Sequence Length	Conserved Regions	Variable Region	Pi. Informative	Singleton
M1	103	1826	958	862	615	243

Table 4. M1 NCBI Query cover, E-Value and percent Identity with closely similar Genotype

Sample/Genotype	Nearest Resemblance	Query Cover (%)	E-Value	Percent Identity
M1	Homo sapiens chromosome 1 clone CH17-31M	91	0.0	93.85

4.2. Genotype M2 MIR-34A Phylogenetic Tree

Individual M2 sample Female with the disease history of diabetes belongs to KOHAT District. The NCBI analysis showed Genotype 2, sample M2 query cover is 99%, percent identity 92.87% and an E value of 0.0 with Homo Sapiens Chromosome 1 clone CH17-31M (Table 5, 6) (Figure 5).

4.3. Genotype M3 MIR-34A Phylogenetic Tree

Individual M3 sample is female with Hypertension and family history of diabetes. The NCBI analysis showed Genotype 3, sample M3 query cover is 93%, percent identity 94.45% and an E value of 0.0 with Homo Sapiens Chromosome 1 clone CH17-31M (Table 7, 8) (Figure 6).

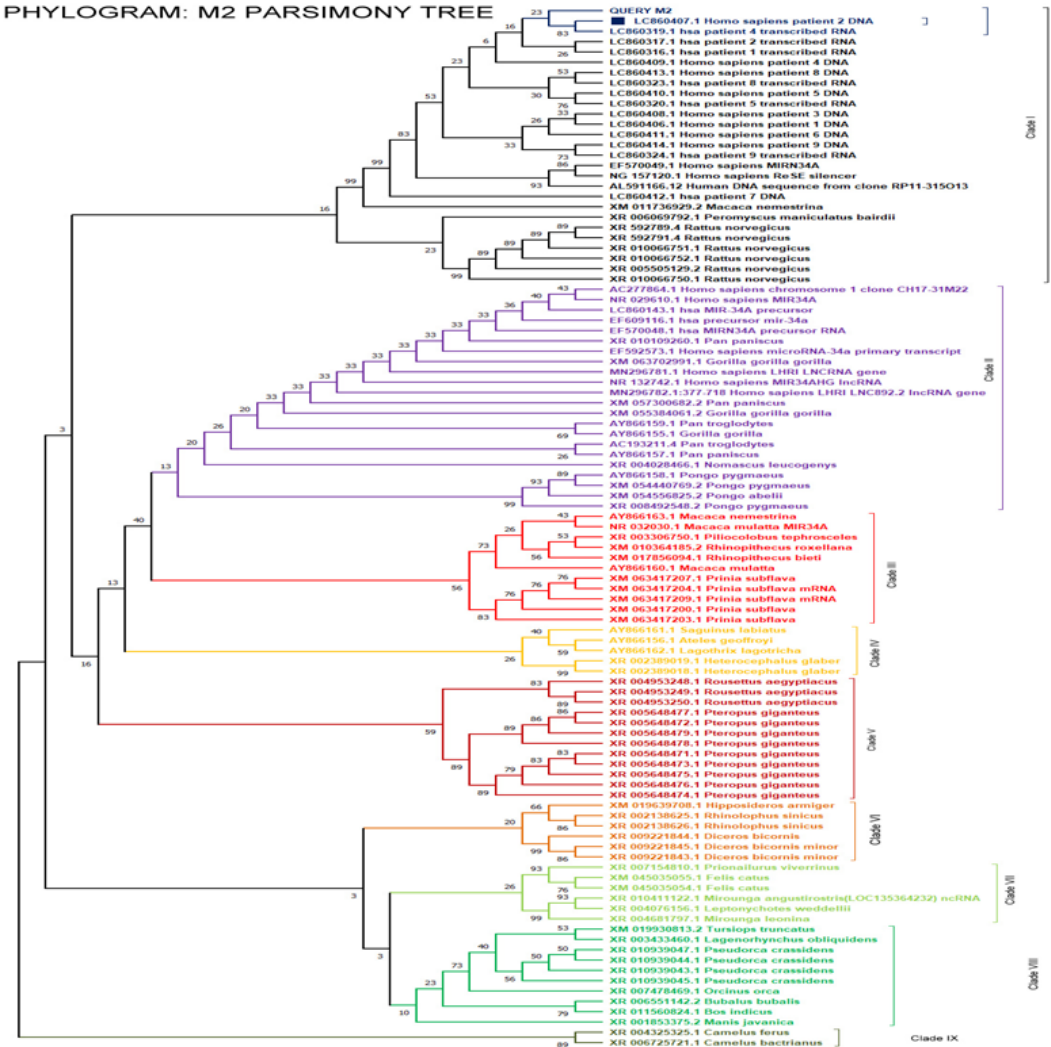


Figure 5: Parsimony tree M2 representative phylogenetic tree of query Genotype 2 showing evolutionary relationship with (Nc_9150668-9152777)

Table 5: M2 Pi-informative detailed Molecular characterization of M2 genotype and conserved regions

Sample	Taxa Selected	Sequence Length	Conserved Regions	Variable Region	Pi. Informative	Singleton
M2	94	1956	1229	603	440	160

Table 6: M2 NCBI Query cover, E-Value and percent Identity with closely similar Genotype

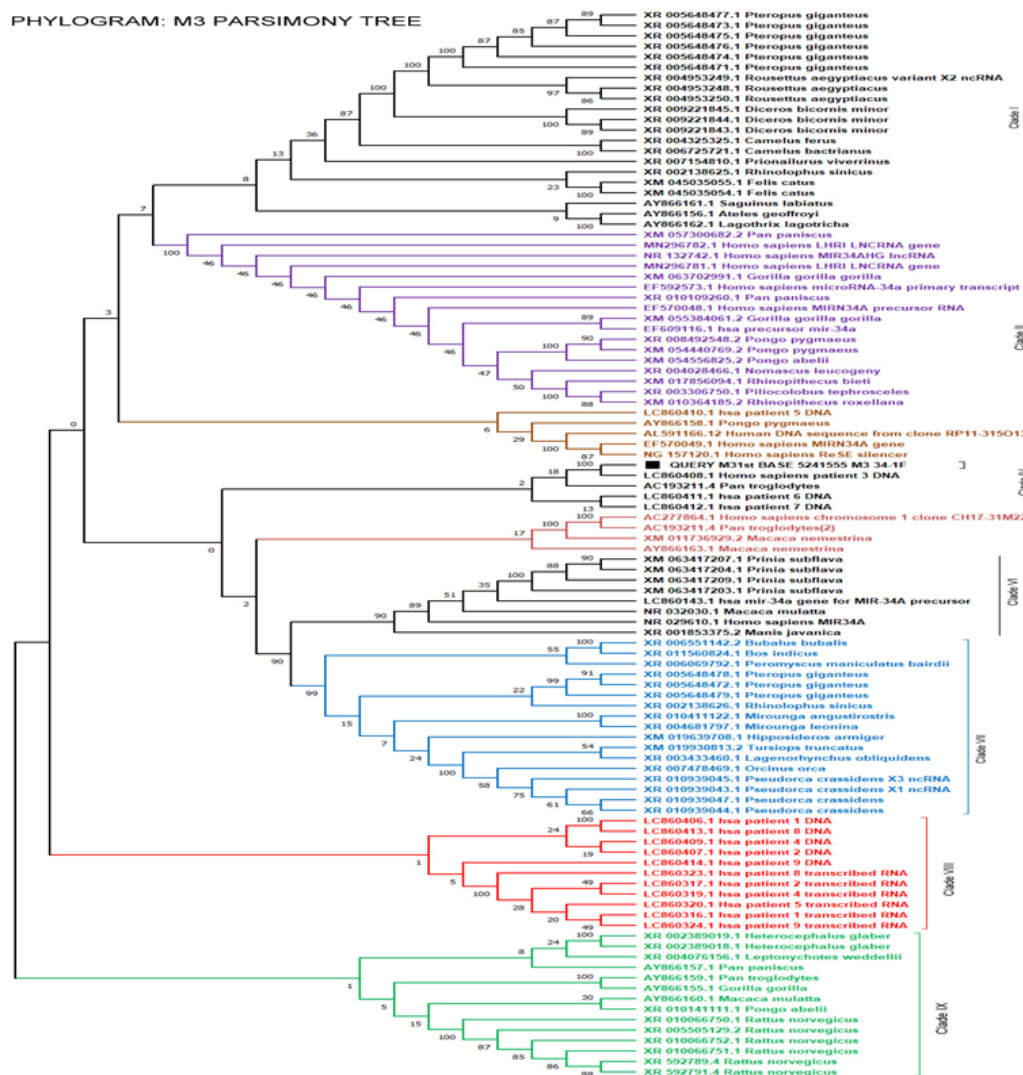
Sample/Genotype	Nearest Resemblance	Query Cover (%)	E-Value	Percent Identity (%)
M2	Homo sapiens chromosome 1 clone CH17-31M	99	0.0	92.87

Table 7: M3 Pi-informative detailed Molecular characterization of M3 genotype and conserved regions

Sample	Taxa Selected	Sequence Length	Conserved Regions	Variable Region	Pi. Informative	Singleton
M3	103	1825	949	839	578	255

Table 8: M3 NCBI Query cover, E-Value and percent Identity with closely similar Genotype

Sample/Genotype	Nearest Resemblance	Query Cover (%)	E-Value	Percent Identity (%)
M3	Homo sapiens chromosome 1 clone CH17-31M	93	0.0	94.45

**Figure 6:** Parsimony tree M3 representative phylogenetic tree of query Genotype 3 showing evolutionary relationship with (Nc_9150668-9152777)

4.4. Genotype M4 MIR-34A Phylogenetic Tree

T2DM Individual M4 sample is male diagnosed with cardiovascular diseases, Periodontal and Hypoglycaemia from KOHAT District. The NCBI analysis showed Genotype 4, sample M4 query cover is 99%, percent identity 92.92% and an E value of 0.0 with Homo Sapiens Chromosome 1 clone CH17-31M (Table 9, 10) (Fig. 7).

4.5. Genotype M5 MIR-34A Phylogenetic Tree

Individual M5 sample Retinopathy, Deafness, Hypertense and Hyperglycaemia with family history of Diabetes belongs to Kohat. The NCBI analysis showed Genotype 5 Query cover is 100%, percent identity is 96% and E-value of 0.0 with Homo Sapiens Chromosome 1 clone CH17-31M (Table 11, 12) (Figure 8).

4.6. Genotype M6 MIR-34A Phylogenetic Tree

Individual M6 sample is male with the disease history of diabetes belongs to Datagram district. The NCBI analysis showed

Genotype 6, sample M6 query cover is 93%, percent identity 92.70% and an E value of 0.0 with Homo Sapiens Chromosome 1 clone CH17-31M (Table 13, 14) (Figure 9).

4.7. Genotype M7 MIR-34A Phylogenetic Tree

Individual M7 sample is male with Hepatobiliary and Hyperglycaemia the disease history of diabetes belongs to Battagram District. The NCBI analysis showed Genotype 7, sample M7 query cover is 62%, percent identity 92.13% and an E value of 2e-152 with Homo Sapiens Chromosome 1 clone CH17-31M (Table 15, 16) (Figure 10).

4.8. Genotype M8 MIR-34A Phylogenetic Tree

Individual M8 sample is female diagnosed with diabetes and Hypertension. The 101 taxa selected for phylogenetic characterizations. The NCBI analysis showed Genotype 8, sample M8 query cover is 99%, percent identity 94.75% and an E value of 0.0 with Homo Sapiens Chromosome 1 clone CH17-31M (Table 17, 18) (Figure 11).

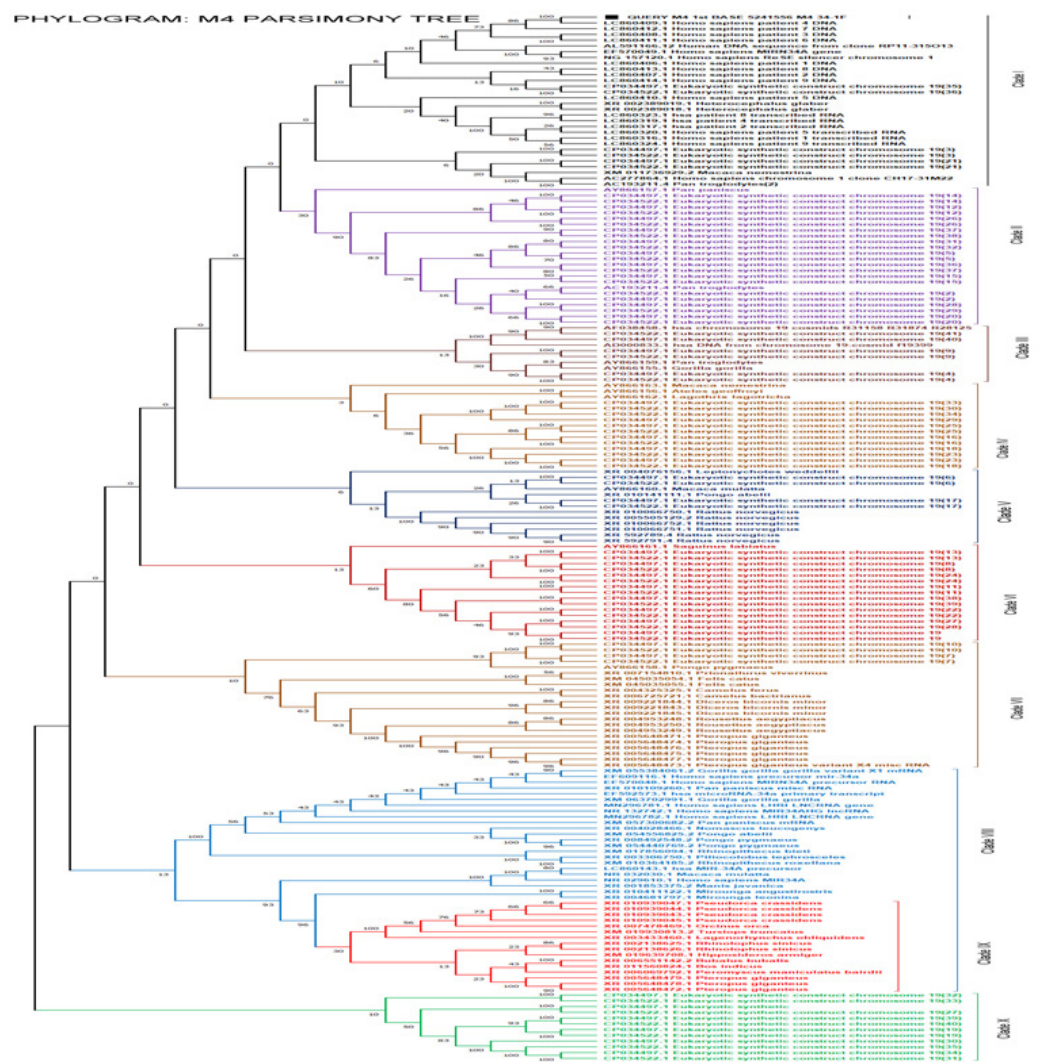


Figure 7: Parsimony tree M4 representative phylogenetic tree of query Genotype 4 showing evolutionary relationship with (Nc_9150668-9152777)

Table 9: M4 Pi-informative detailed Molecular characterization of M4 genotype and conserved regions

Sample	Taxa Selected	Sequence Length	Conserved Regions	Variable Region	Pi. Informative	Singleton
M4	183	2207	992	1095	695	344

Table 10: M4 NCBI Query cover, E-Value and percent Identity with closely similar Genotype

Sample/Genotype	Nearest Resemblance	Query Cover (%)	E-Value	Percent Identity (%)
M4	Homo sapiens chromosome 1 clone CH17-31M	99	0.0	92.92

Table 11: M5 Pi-informative detailed Molecular characterization of M5 genotype and conserved regions

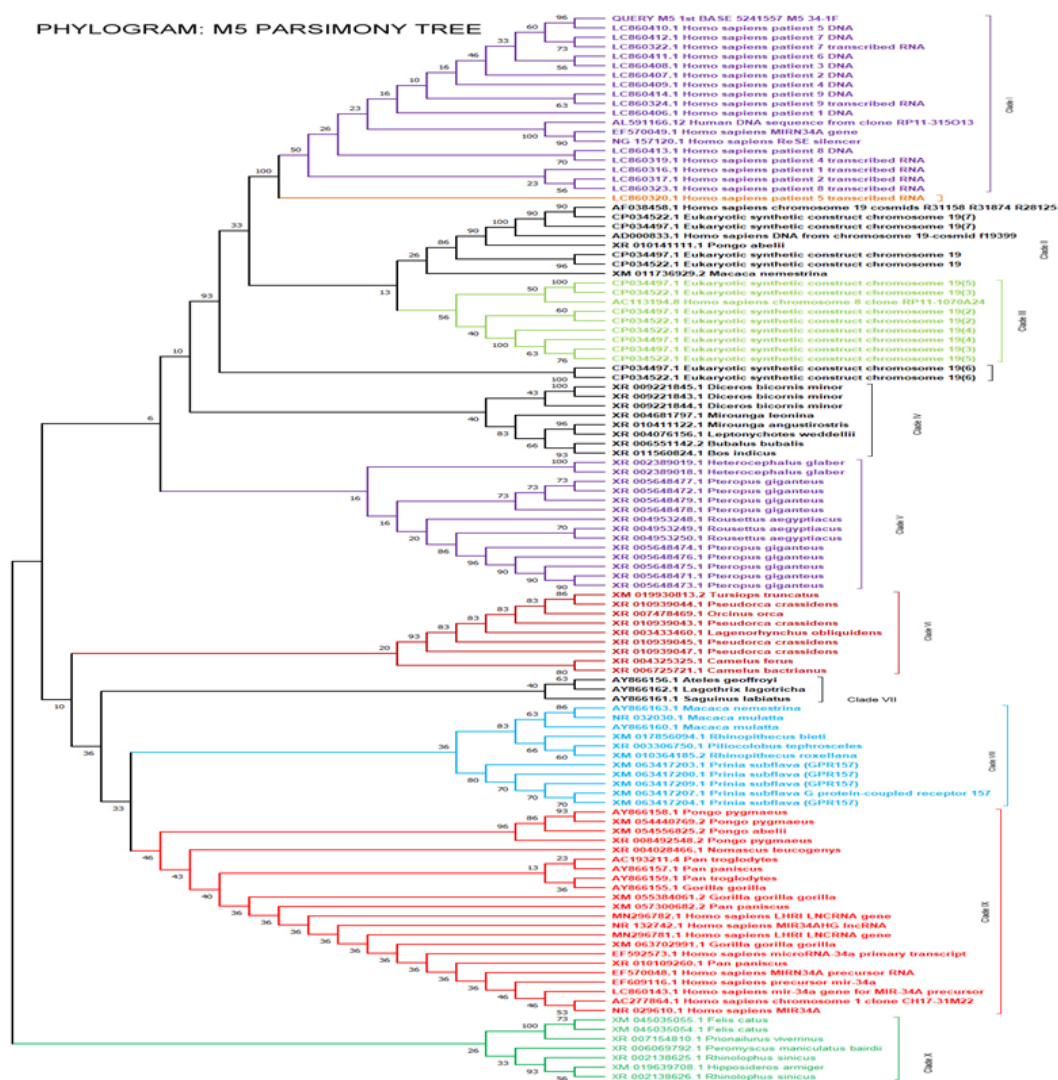
Sample	Taxa Selected	Sequence Length	Conserved Regions	Variable Region	Pi. Informative	Singleton
M5	114	1970	834	1037	674	168

Table 12: M5 NCBI Query cover, E-Value and percent Identity with closely similar Genotype

Sample/Genotype	Nearest Resemblance	Query Cover (%)	E-Value	Percent Identity (%)
M5	Homo sapiens chromosome 1 clone CH17-31M	100	0.0	96.00

Table 13: M6 Pi-informative detailed Molecular characterization of M6 genotype and conserved regions

Sample	Taxa Selected	Sequence Length	Conserved Regions	Variable Region	Pi. Informative	Singleton
M6	102	2011	1074	934	610	321



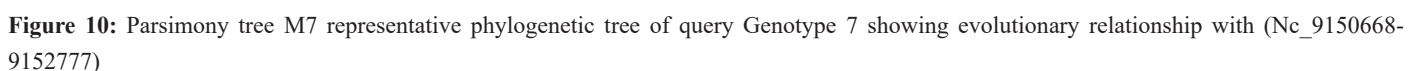




Figure 11: Parsimony tree M8 representative phylogenetic tree of query Genotype-8 showing evolutionary relationship with (Nc_9150668-9152777)

4.9. Genotype M9 MIR-34A Phylogenetic Tree

Individual M9 sample is female with diabetes and Hypertension. The NCBI analysis showed Genotype 9, sample M9 query cov-

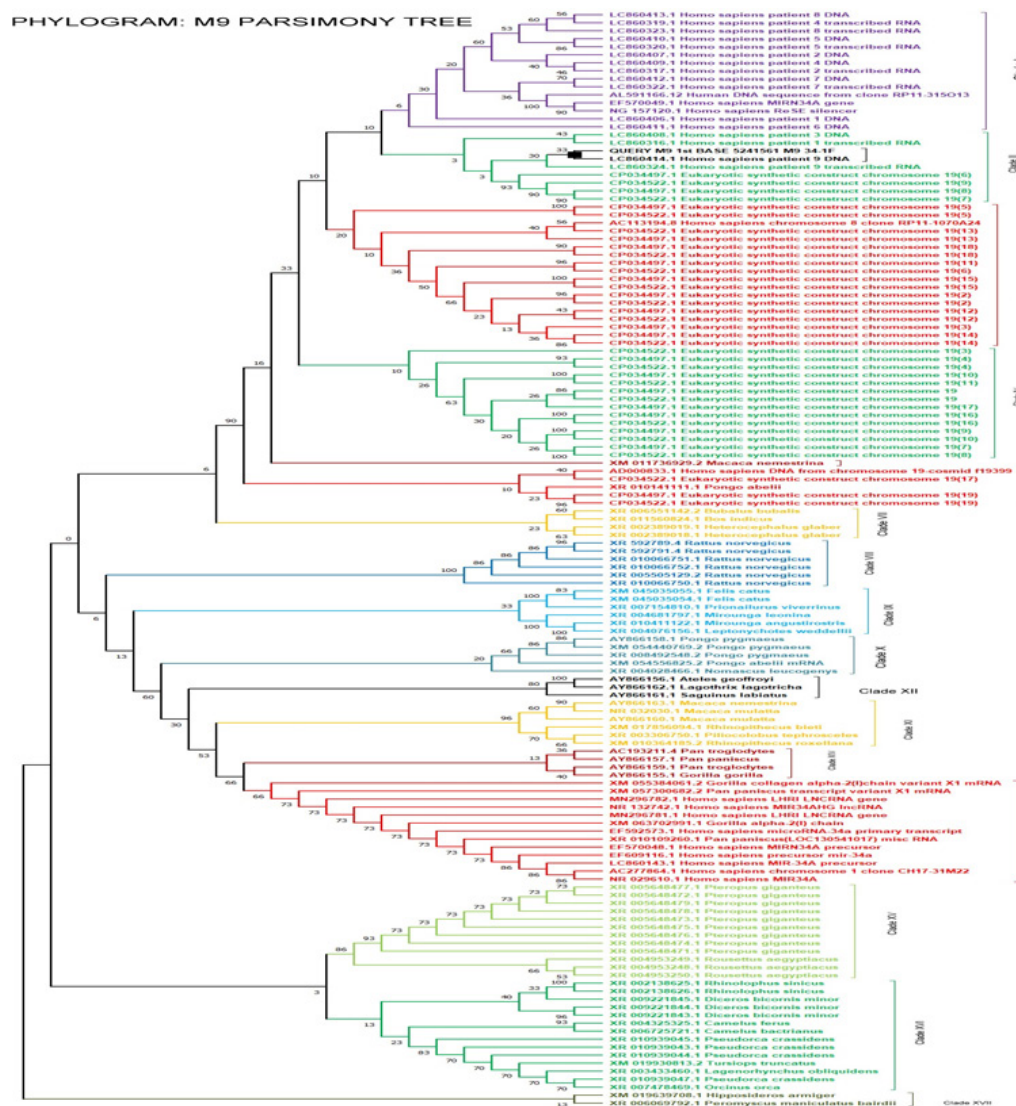
er is 90%, percent identity 94.33% and an E value of 0.0 with Homo Sapiens Chromosome 1 clone CH17-31M (Table 19, 20) (Figure 12).

Table 19: M9 Pi-informative detailed Molecular characterization of M9 genotype and conserved regions

Sample	Taxa Selected	Sequence Length	Conserved Regions	Variable Region	Pi. Informative	Singleton
M9	138	1850	554	1281	821	253

Table 20: M9 NCBI Query cover, E-Value and percent Identity with closely similar Genotype

Sample/Genotype	Nearest Resemblance	Query Cover (%)	E-Value	Percent Identity (%)
M9	Homo sapiens chromosome 1 clone CH17-31M	90	0.0	94.33



5. Discussion

MEGA a window-based program, was used to examine the Genetic Distance and nucleotide variation on the bases of the Kimura 2 parameter (K2P) model and Parsimony tree was constructed in MEGA 11 with 1000 replicates the bootstrap value and with rates and patterns of the gamma distribution method

The M1 Parsimony Tree characterizations for query genotype M1 start with selection of 103 taxa, total sequence length 1826 nucleotide bases, aligned sequence length 824 base pairs with 958 conserved sites, 862 variable sites, 615 parsimony-informative sites, and 243 singletons. This Parsimony tree consists of 11 clades. Phylogram for query 1st BASE 5241553 M1 34A-1F constructed with 100 BLAST hits representing distance scale 96 bootstrap values. The query sample M1 patient hsa-Mir-34a gene is grouped with Hsa-patients 1 DNA, transcribed RNA, Pan troglodytes in Clade X, nearest group is hsa- ReSE silencer, clone RP-11, *Macaca nemestrina*, *Pongo abelii*, *Mirounga leonina*, *Leptonychotes weddellii*, *Mirounga angustirostris*. The Clade I contains *Pan paniscus*, Gorilla variants X3 mRNA, Hsa-LHRI LNCRNA gene. The clade II contains *Macaca nemestrina*, *Macaca mulatta*, *Rhinopithecus bieti* variant X2 mRNA,

Piliocolobus tephrosceles, *Rhinopithecus rexellana*, *Nomascus leucogenys*. The clade III contains *Nomascus leucogenys*. The Clade IV contains *Pongo Pygmaeus X3 misc RNA*, *Pongo abellii*. The clade V contains *Ateles geoffroyi*, *Lagothrix Lagotricha*, *Saguinus labiatus*, *Prinia subflava*. The clade VI is *Rousettus aegyptiacus*, *Pteropus giganteus*. The clade VII contains *Rhinolophus Sinicus*, *Hipposideros armiger*, *Camelus ferus*, *Camelus bactrianus*. The clade VIII consists of *Diceros bicornis minor*, *Peromyscus Maniculatus bairdii*. The clade XI consists of *Bubalus bubalis*, *Bos indicus*, *Heterocephalus glaber*, *Pseudorca Crassidens*, *Tursiops truncates*, *Lagenorhynchus obliquidens*, *Orcinus orca*.

The M2 Parsimony Tree characterizations for query genotype M2 start with selection of 94 taxa, total sequence length 1956 nucleotide bases, aligned sequence length 824 bp, 1229 bp conserved sites, 603 bp variable sites, 440 bp parsimony-informative sites, and 160 bp singletons. This Parsimony tree consists of 9 clades. The Phylogram for query M2 hsa-Mir-34a gene constructed with 100 BLAST hits representing distance scale 83 bootstrap values are grouped with patient 2 DNA along with 1, 2, 3, 4, 5, 6, 8, 9 and ReSE silencer, clone RP11, *Macaca nemestrina*, *Peromyscus maniculatus bairdii* in subgroup of clade I *Rattus norvegicus*. The clade II consists of hsa-chr1 clone CH17, precursor RNA, *Pan paniscus*, primary transcript, *Gorilla*, LHRI LNCRNA, *Pan troglodytes*, *Nomascus leucogenys*, *Pongo pygmaeus*. The clade III consists of *Macaca nemestrina*, *Macaca mulatta*, *Piliocolobus tephrosceles*, *Rhinopithecus roxellana*, *Rhinopithecus bieti*, *Prinia subflava*. The clade IV consists of *Saguinus labiatus*, *Ateles geoffroyi*, *Lagothrix lagotricha*, *Heterocephalus glaber*. The clade V consists of *Rousettus aegyptiacus*, *Pteropus giganteus*. The clade VI consists of *Mirounga Leonina*, *Leptonychotes weddellii*, *Mirounga angustirostris*, *Felis catus*, and *Prionailurus viverrinus*.

The M3 Parsimony Tree characterizations for query genotype M3 start with selection of 103 taxa, total sequence length 1825 nucleotide bases, aligned sequence length 824 base pairs with 949 conserved sites, 839 variable sites, 578 parsimony-informative sites, and 255 singletons. This M3 Parsimony Tree consists of 9 clades. The Phylogram for query M3 1st Base 5241555 M3 34-1F constructed with 100 BLAST hits representing distance scale 100 bootstrap values are grouped in clade IV with hsa-patients 3 DNA, *Pan troglodytes*, hsa-patient 6 and 7 DNA. The clade I contains *Pteropus giganteus*, *Rousettus aegyptiacus* variant X2 ncRNA, *Diceros bicornis minor*, *Camelus ferus*, *Camelus bactrianus*, *Prionailurus viverrinus*, *Rhinolophus Sinicus*, *Felis catus*, *Saguinus labiatus*, *Ateles geoffroyi*, *Lagothrix lagotricha*. The clade II consists of *Pan Paniscus*, LHRI LNCRNA gene, *Gorilla gorilla*, *Pongo Pygmaeus*, *Pongo abellii*, *Nomascus Leucogeny*, *Rhinopithecus bieti*, *Piliocolobus tephrosceles*, *Rhinopithecus roxellana*. The clade III consists of hsa-patient 5 DNA, *Pongo pygmaeus*, and clone RP 11, hsa-ReSE silencer. The clade V consists of hsa-clone CH17, *Pan troglodytes*, *Macaca nemestrina*. The clade VI consists of *Prinia subflava*, *Macaca mulatta*,

Manis Javanica. The clade VII consists of *Bubalus bubalis*, *Bos indicus*, *Peromyscus maniculatus bairdii*, *Pteropus giganteus*, *Rhinolophus Sinicus*, *Mirounga angustirostris*, *Mirounga leonine*, *Hipposideros armiger*, *Tursiops truncates*, *Lagenorhynchus obliquidens*, *Orcinus Orca*, *Pseudorca crassidens X3 ncRNA*. The clade VIII consists of hsa-patient 1, 2, 4, 8, 9 DNA and 1, 2, 4, 5, 8, 9 transcribed RNA. The clade IX consists of *Heterocephalus glaber*, *Leptonychote weddellii*, *Pan paniscus*, *Gorilla*, *Macaca mulatta*, and *Rattus norvegicus*.

The M4 Parsimony Tree characterizations for query genotype M4 start with selection of 183 taxa, total sequence length 2207 nucleotide bases, aligned sequence length 824 base pairs with 992 conserved sites, 1095 variable sites, 695 parsimony-informative sites, and 344 singletons. This M4 Parsimony Tree consists of 10 clades. The phylogram for query M4 1st Base 5241556 M4 34-1F constructed with 100 BLAST hits representing distance scale 100 bootstrap values is grouped in clade I with hsa-patients 4 DNA, 1,2, 3, 6,7, 8, 9 DNA, hsa-clone RP 11, chr 1, hsa-MIRN 34A gene, hsa-ReSE silencer, Eukaryotic synthetic construct chr 19, hsa-patient 5 DNA, *Heterocephalus glaber*, hsa-patient 2, 4, 8 transcribed RNA, hsa-Patient 1, 5, 9 transcribed RNA, *Macaca nemestrina*, hsa-clone CH17, *Pan troglodytes*. The clade II consists of *Pan paniscus*, *Pan troglodytes*, and *Gorilla gorilla*. The clade IV consists of *Macaca nemestrina*, *Ateles geoffroyi*, *Lagothrix Lagotricha*. The clade V consists of *Leptonychotes weddellii*, *Pongo abellii*, and *Rattus norvegicus*. Clade VI consists of *Saguinus Labiatus*, clade VII consists of *Pongo pygmaeus*, *Prionailurus viverinus*, *Felis catus*, *Camelus bactrianus*, *Diceros bicornis minor*, *Rousettus aegyptiacus*, *Pteropus giganteus*. The clade VIII consists of *Gorilla*, hsa-precursor Mir34A, *Pan paniscus* misc RNA, Mir-34A primary transcript, *Gorilla*, hsa-LHRI LNCRNA gene, *Nomascus leucogenys*, *Pongo abellii*, *Pongo pygmaeus*, *Rhinopithecus tephrosceles*, *Rhinopithecus roxellana*, *Macaca mulatta*, *Manis javanica*. The clade IX consists of *Mirounga angustirostris*, *Mirounga Leonina*, *Pseudorca crassidens*, *Orcinus orca*, *Tursiops truncates*, *Lagenorhynchus obliquidens*, *Rhinolophus sinicus*, *Hipposideros armiger*, *Bubalus bubalis*, *Bos indicus*, *Peromyscus maniculatus bairdii*, *Pteropus giganteus*.

The M5 Parsimony Tree characterizations for query genotype M5 start with selection of 114 taxa, total sequence length 1970 nucleotide bases, aligned sequence length 824 base pairs with 834 conserved sites, 1037 variable sites, 674 parsimony-informative sites, and 168 singletons. This M5 Parsimony Tree consists of 10 clades. The Phylogram for query M5 1st Base 5241557 M5 34-1F constructed with 100 BLAST hits representing distance scale 96 bootstrap values is grouped in clade I with hsa-patient 5 DNA, 1,2,3,4, 6,7,9 DNA and hsa-1,2,4,5,7,8,9 transcribed RNA, clone RP 11, hsa-ReSE silencer. The clade II consists of hsa-chromosome 19 cosmids RP 31158, Eukaryotic synthetic construct Chr 19 cosmids R31158, *Pongo abellii*, *Macaca nemestrina*. The clade III consists of Eukaryotic synthetic construct Chr 19 (2, 3, 4, 5), hsa-clone RP11. The clade IV consists of Di-

ceros bicornis minor, *Mirounga Leonina*, *Mirounga angustirostris*. The clade V consists of *Heterocephalus glaber*, *Pteropus giganteus*, and *Rousettus aegyptiacus*. Clade VI consists of *Tursiops truncatus*, *Pseudorca crassidens*, *Orcinus orca*, *Lagenorhynchus obliquidens*, *Camelus ferus*, and *Camelus bactrianus*. The clade VII consists of *Ateles geoffroyi*, *Lagothrix lagotricha*, and *Saguinus labiatus*. The clade VIII consists of *Macaca mulatta*, *Rhinopithecus tephrosceles*, *Rhinopithecus roxellana*, *Prinia subflava* (GPR 157). The clade IX consists of *Pongo pygmaeus*, *Pongo abelli*, *Nomascus leucogenys*, *Pan troglodytes*, *Pan paniscus*, *Gorilla gorilla*, hsa-LHRI LNCRNA gene, Mir34AHG, hsa-Chr 1, clone CH17. The clade X consists of *Felis catus*, *Prionailurus viverrinus*, *Peromyscus maniculatus bairdii*, *Rhinolophus*.

The M6 Parsimony Tree characterizations for query genotype M6 start with selection of 102 taxa, total sequence length 2011 nucleotide bases, aligned sequence length 824 base pairs with 1074 conserved sites, 934 variable sites, 610 parsimony-informative sites, and 321 singletons. This M6 Parsimony Tree consists of 14 clades. The phylogram for query M6 1st BASE 5241558 M6 34-1F constructed with 100 BLAST hits representing distance scale 53 bootstrap values is grouped in clade XII with hsa-patient 6 DNA, hsa-patient 1 transcribed RNA. The clade I consists of hsa-LHRI LNCRNA gene, *Pan paniscus* misc RNA, hsa-Mir34a primary transcript, *Gorilla collagen alpha-2 (I) chain*. The clade II consists of *Piliocolobus tephrosceles*, *Macaca mulatta* MIR34A, hsa-MIR 34A precursor RNA, *Gorilla collagen alpha 2(I) chains*. The clade III consists of *Pan troglodytes*, *Gorilla gorilla*, hsa-MIR34a precursor, hsa-chr 1 clone CH17, *Pan troglodytes*. The clade IV consists of *Macaca mulatta*, *Pongo pygmaeus*, *Saguinus Labiatus*, *Ateles geoffroyi*, *Lagothrix Lagotricha*. The clade V consists of *Prinia subflava* (GPR 157) *Rhinolophus Sinicus*. The clade VI consists of *Peromyscus maniculatus bairdii*, *Rattus norvegicus*. The clade VII consists of *Pseudorca crassidens*, *Lagenorhynchus obliquidens*, *Tursiops truncatus*, *Pseudorca crassidens*, *Orcinus orca*, *Diceros bicornis minor*, *Camelus ferus*, *Camelus bactrianus*, *Felis catus*, *Prionailurus viverrinus*, *Mirounga Leonina*, *Mirounga angustirostris*, *Leptonychotes weddellii*. The clade IX consists of *Manis javanica*, *Rousettus aegyptiacus*, *Pteropus giganteus*. The clade X consists of *Hipposideros armiger*, *Heterocephalus glaber*, *Bubalus bubalis*, and *Bos indicus*. The clade XIII consists of hsa-patient 1,2,3,4,5,7,8,9 DNA, hsa-patient 2,4,5,8,9 transcribed RNA, Human DNA sequence from clone RP 11-315013, hsa-ReSE silencer chr 1. The clade XIV consists of *Macaca nemestrina*, *Pongo abelli*.

The M7 Parsimony Tree characterizations for query genotype M7 start with selection of 102 taxa, total sequence length 664 nucleotide bases, aligned sequence length 824 base pairs with 365 conserved sites, 299 variable sites, 287 parsimony-informative sites, and 12 singletons. The M7 Parsimony Tree consists of 8 clades. This phylogram for query M7 1st BASE 5241559 M7 34-1F constructed with 100 BLAST hits representing distance scale 99 bootstrap values is grouped in clade I with hsa-patient

7 transcribed mir-34a, hsa-patient 1,2,3,4,5,6,8,9 DNA, hsa-Patient 1,2,5,8,9 transcribed RNA, hsa-ReSE silencer, hsa-RP11. The clade II consists of *Felis Catus*, *Leptonychotes weddellii*, ncRNA, *Mirounga angustirostris*, *Prionailurus viverrinus*, and *Mirounga leonina*. The clade III consists of *Saguinus labiatus*, *Ateles geoffroyi*, *Lagothrix lagotricha*, *Rousettus aegyptiacus*, *Heterocephalus glaber*, *Rousettus aegyptiacus*, and *Pteropus giganteus*. The clade IV consists of *Nomascus leucogeny*. The clade V consists of *Prinia subflava* GPR 157, *Rhinopithecus roxellana*, *Piliocolobus tephrosceles*, *Macaca Nemestrina*, *Macaca mulatta*, *Pongo abelli* mRNA, *Pongo pygmaeus*. The clade VI consists of *Camelus ferus*, *Camelus bactrianus*, hsa-LHRI LNCRNA gene, *Gorilla gorilla* variant X3 mRNA, hsa-MIR34A primary transcript, *Pan paniscus*, hsa-chr 1 clone CH17. The clade VII consists of *Rhinolophus sinicus*, *Hipposideros armiger*, *Peromyscus maniculatus bairdii*, *Monodelphis domestica*, *Rattus norvegicus*. The clade VIII consists of *Diceros bicornis minor*, *Bubalus bubalis*, *Bos indicus*, *Tursiops truncatus*, *Orcinus orca*, *Lagenorhynchus obliquidens*, and *Pseudorca crassidens*.

The M8 Parsimony Tree characterizations for query genotype M8 start with selection of 101 taxa, total sequence length 2085 nucleotide bases, aligned sequence length 824 base pairs with 1251 conserved sites, 704 variable sites, 450 parsimony-informative sites, and 252 singletons. This M8 Parsimony Tree consists of 12 clades. The phylogram for query M8 1st Base 5241560 M8 34-1F constructed with 100 BLAST hits representing distance scale 83 bootstrap values is grouped in clade I with hsa-patient 8 DNA along with patient 1,2,3,4,5,6,7,9 DNA and hsa-patient 1,2,4,5,8,9 transcribed RNA, hsa-ReSE silencer, clone RP11. The clade II consists of *Macaca nemestrina*, *Prionailurus viverrinus*, *Felis catus*, *Mirounga angustirostris*, *Leptonychotes weddellii*, *Mirounga leonine*. The clade III consists of hsa-chromosome 1 clone CH17, hsa-MIR-34A precursor, *Pan paniscus*, *Gorilla gorilla gorilla*, hsa-LHRI LNCRNA gene, MIR34AHG LncRNA gene, *Bubalus bubalis*, *Bos indicus*, *Heterocephalus glaber*. The clade IV consists of *Pan troglodytes*, *Pan paniscus*, and *Gorilla gorilla*. The clade V consists of *Macaca nemestrina*, *Macaca mulatta*, *Piliocolobus tephrosceles*, *Rhinopithecus roxellana*, *Rhinopithecus bieti*, *Macaca mulatta*, *Prinia subflava* GPR 157, and *Pongo pygmaeus*. The clade VI consists of *Ateles geoffroyi*, *Lagothrix lagotricha*, and *Saguinus labiatus*. The clade VII consists of *Camelus ferus*, *Camelus bactrianus*, and *Rattus norvegicus*. The clade VIII consists of *Pseudorca crassidens*, *Orcinus orca*, *Lagenorhynchus obliquidens*, and *Tursiops truncatus*. The clade IX consists of *Rhinolophus sinicus*, *Hipposideros armiger*, and *Peromyscus maniculatus bairdii*. The clade X consists of *Rousettus aegyptiacus*. The clade XI consists of *Pteropus giganteus*. The clade XII consists of *Diceros bicornis minor*, *Manis javanica*.

The M9 Parsimony Tree characterizations for query genotype M9 start with selection of 138 taxa, total sequence length 1850 nucleotide bases, aligned sequence length 824 base pairs with 554 conserved sites, 1281 variable sites, 821 parsimony-infor-

mative sites, and 253 singletons. This Parsimony tree consists of 17 clades. The phylogram for query M9 1st Base 5241561 M9 34-1F constructed with 100 BLAST hits representing distance scale 93 bootstrap values is grouped in clade II with has patient 9 DNA, transcribed RNA along with Eukaryotic synthetic construct 19(5,6,7,9). The clade I consist of hsa-patient 1,2,3,4,5,6,7,8 DNA and 1,2,4,5,7,8 transcribed RNA has-clone RP11, has ReSE silencer. The clade III consist of ESC chromosome 19 (2, 5, 6, 12, 13, 15, 18). The clade IV Eukaryotic synthetic construct chromosome 19 (4, 11, 16) is performed. The clade V consists of *Macaca nemestrina*. The clade VI consists of hsa chromosome 19 cosmid, Eukaryotic synthetic construct chromosome 19, *Pongo abelii*. The clade VII consists of *Rattus norvegicus*. The clade VIII consists of *Bubalus bubalis*, *Bos indicus*, *Heterocephalus glaber*, *Prionailurus viverrinus*, *Felis Catus*, *Mirounga leonine*, *Mirounga angustirostris*, *Leptonychotes weddellii*. The clade X consists of *Pongo Pygmaeus*, *Pongo abelii*, and *Nomascus Leucogenys*. The clade XI consists of *Ateles geoffroyi*, *Lagothrix Lagotricha*, *Saguinus labiatus*. The clade XII consists of *Macaca nemestrina*, *Macaca mulatta*, *Rhinopithecus bieti*, *Ptilocolobus tephrosceles*, and *Rhinopithecus roxellana*. The clade XIII consists of *Pan troglodytes*, *Pan paniscus*, and *Gorilla gorilla*. The clade XIV consists of *Gorilla gorilla*, alpha 2 (I) chain variant X1 mRNA, *Pan paniscus* transcript variant X1 mRNA, hsa-LHRI LNCRNA gene, hsa-MIR34AHG LncRNA, *Gorilla alpha 2 (I) chain*, hsa-Mir34a primary transcript, hsa-chr1 clone CH17-31M22. The clade XV consists of *Pteropus giganteus*, *Rousettus aegyptiacus*. The clade XVI consists of *Rhinolophus Sinicus*, *Diceros bicornis minor*, *Camelus ferus*, *Camelus bactrianus*, *Pseudorca crassidens*, *Tursiops truncatus*, *Lagenorhynchus obliquidens*, *Orcinus orca*. The clade XVII consists of *Hipposideros armiger*, *Peromyscus maniculatus bairdii*.

The gene Mir-34a Query M1-M10 analysis determines similarity, monophyly and evolutionary lineage to known sequences in public database. The Tamura-Nei method to compute distances in phylogenetic tree is based on number of base substitutions per site.

Pi-informative sites detailed Molecular characterization of Genotypes M1-M9 is discussed. The highest available selected Taxa are 183 of Genotype M4 and least are 94 of Genotype M2. The lengthiest Sequence is 2207 bp of Genotype M4 and shortest of M7 664bp. The highest conserved region of Genotype M2 is 1229bp and least is 365bp of Genotype M7. The variable region of length 1281 bp of Genotype M9 and least 603 bp nucleotide of Genotype M2. The Pi Ψ informative site is longest of 821 bp for Genotype M9 and shortest of Genotype M7 with 287 bp nucleotide. The Singleton Highest value is 344 bp of Genotype M4 and least of 12 bp nucleotide of Genotype M7.

The NCBI Query cover, E-value and Percent Identity with closely similar Genotype M1-M9 with *Homo sapiens* chromosome 1 clone CH17-31M. The Query cover is 100% and Percent Identity is 96% with Genotype M5 at top.

The T2DM patients had a history of diabetes in parents and they are Patient number Genotype M4, M7 and M5. The M4 is T2DM Individual 45 years old male from KOHAT District with the family history of Diabetes from Bangash tribe. The M5 Individual is 70 years old male of KOHAT prone to diseases of Diabetes with family history of jutt tribe. The M7 Individual belongs to Battagram District 40 year's old male with the maternal disease history of diabetes with a Hepatobiliary disease.

6. Conclusion

The phylogenetic analysis was performed based on Parsimony tree method by using MEGA 11 software. The nearest similar organism to the Query is mostly Human hsa-DNA and Transcribed RNA. The Query M1 is like hsa-Patient 1 DNA, hsa ReSE silencer, *Pan troglodytes*, clone RP11-315013, hsa-Mir-34a primary transcript, and clone CH12. The Query M2 is similar with hsa-patient 2 DNA and Transcribed RNA, clone RP11-315013, ReSE silencer, hsa-patient 8 transcribed RNA, hsa-Patient 4,5 DNA & transcribed RNA, hsa-Patient 2 DNA Patient 4 transcribed RNA, The Query M3 is similar to hsa-Patient 3 DNA, *Pan troglodytes*, hsa-ReSE Silencer, *Rattus norvegicus*, hsa-Patient 4 transcribed RNA, hsa-Mir-34a precursor, The Query M4 is similar to hsa-Patient 4 DNA, hsa- clone RP 11, hsa-patient 6,7 DNA, hsa-patient 4, 8 transcribed RNA. The Query M5 is like hsa-Patient 4 DNA, hsa-patient 5 DNA, hsa-patient 2, 3, 6, hsa-patient 5 DNA and transcribed RNA, hsa-Patient 7 DNA and transcribed RNA. The Query M6 hsa-Patient 6 DNA, hsa-patient 1 transcribed RNA, Mir-34a gene, hsa-clone RP11-315013, Patient 7 DNA. The Query M7 hsa-Patient 7 DNA and Transcribed RNA, *Felis Catus*, hsa-patient 5 DNA and transcribed RNA, Human DNA sequence Clone RP. The Query M8 hsa-Patient 8 DNA and transcribed RNA, hsa-clone RP11, hsa-Patient 4 DNA, hsa-Patient 5 DNA and transcribed RNA. The Query M9 is similar to hsa-Patient 9 DNA and transcribed RNA, Eukaryotic synthetic construct chromosome 19, hsa-Patient 5 transcribed MicroRNA-34A, hsa-LHRI LNCRNA.

The phylogenetic relation of every sample regarding Mir-34a gene is BLAST in NCBI Gene Bank and Ancestor tree is drawn has similarities with more than 100 organisms creating susceptibility of presence of Diabetes and Cancer in other organisms including invertebrates and vertebrates. The Genotype M4 has highest selected Taxa 183 with longest sequenced length of 2207 bp nucleotide and Singleton value of 344 bp. The Genotype M7 has at least Sequence Length 664 bp and Conserved region of 365 bp and Pi Ψ Informative site of 287 bp and singleton value of 12 bp nucleotide.

References

1. Anderson J R, Bothell D, Byrne M D, Douglass S, Lebiere C, Qin Y. An integrated theory of the mind. *Psychological review*. 2004; 111(4):1036.
2. Baek D, Villen J, Shin C, Camargo F D, Gygi SP, Bartel D P. The impact of microRNAs on protein output. *Nature*. 2008; 455(7209):64-71.
3. Bawono P. and Heringa J. Phylogenetic Analyses. *Comprehensive Biomedical Physics*. 2014; 93-110.
4. Bell G I, Kayano T, Buse J B, Burant C F, Takeda J. Molecular biology of mammalian glucose transporters. *Diabetes care*. 1990; 13(3): 198-208.
5. Brennan M A, Lalonde C E, Bain J L. Body image perceptions: Do gender differences exist. *Psi Chi Journal of Undergraduate Research*. 2010; 15(3): 130-8.
6. Chakraborty C, Priya D and Bandyopadhyay S. MirNAs in insulin resistance and diabetes-associated pancreatic cancer: The ‘minute and Miracle’ molecule moving as a monitor in the ‘genomic galaxy’. *Current drug targets*. 2013; 14(10): 1110-1117.
7. Chang T C, Wentzel E A, Kent O A, Ramachandran K, Mullendore M, Lee K H. Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. *Mol Cell*; 2007. 26: 745–752.
8. Choudhuri S. Phylogenetic Analysis. *Bioinformatics for Beginners*. 2014; 209-18.
9. Concepcion C P, Han Y C, Mu P, Bonetti C, Yao E, D’Andrea A. Intact p53-dependent responses in miR-34-deficient mice. *PLoS Genet* 2012; 8(7): e 1002797.
10. Corney D C, Hwang C I, Matoso A, Vogt M, Flesken-Nikitin A, Godwin A K. Frequent downregulation of miR-34 family in human ovarian cancers. *Clin Cancer Res* 2010; 16(4): 1119-28.
11. Cudworth A G. The aetiology of diabetes mellitus. *British Journal of Hospital Medicine*. 1976; 16: 207.
12. Di Martino M T, Leone E, Amodio N, Foresta U, Lionetti M, Pitari M R. Synthetic miR-34a mimics as a novel therapeutic agent for multiple myeloma: in vitro and in vivo evidence. *Clin Cancer Res*; 2012; 18(22): 6260-6270.
13. Dijkstra M K, Van Lom K, Tieleman D, Elstrodt F, Langerak A W, Van Veer M B. 17p13/TP53 deletion in B-CLL patients is associated with microRNA-34a downregulation. *Leukemia*. 2009; 23: 625-627.
14. Ding Y, Ren G, Zhao Z, Xu Y, Luo Y, Li Q, Zhang J. Detection, causes and projection of climate change over China: an overview of recent progress. *Advances in Atmospheric Sciences*. 2007; 24(6): 954-971.
15. Ding Y, Zhang C, Zhang J, Zhang N, Li T, Fang J. Mir-145 inhibits proliferation and migration of breast cancer cells by directly or indirectly regulating TGF- β 1 expression. *Int J Oncol*. 2017; 50: 1701-1710.
16. Edgerton DS, Lautz M, Scott M, Everett CA, Stettler KM, Neal DW, Chu CA, Cherrington AD. Insulin’s direct effects on the liver dominate the control of hepatic glucose production. *The Journal of clinical investigation*. 2006; 116(2): 521-527.
17. Elmer S, Hänggi J, Meyer M, Jäncke L. Increased cortical surface area of the left planum temporale in musicians facilitates the categorization of phonetic and temporal speech sounds. *Cortex*. 2013; 49(10): 2812-2821.
18. Federal Interagency Forum on Aging-Related Statistics (US). Older Americans update: Key indicators of well-being. Federal Interagency Forum on Aging Related Statistics. 2006.
19. Ha S Y, Lee J, Kang S Y, Do I G, Ahn S, Park J O. MET overexpression assessed by new interpretation method predicts gene amplification and poor survival in advanced gastric carcinomas. *Mod Pathol*. 2013; 26(12): 1632-41.
20. Haberle V, Stark A. “Eukaryotic core promoters and the functional basis of transcription initiation”. *Nature Reviews. Molecular Cell Biology*. 2018; 19(10): 621-637.
21. He L, He X, Lim L P, de Stanchina E, Xuan Z. A microRNA component of the p53 tumour suppressor network. *Nature*. 2007; 447(7148): 1130-114.
22. Irvine W J. Classification of idiopathic diabetes. *The Lancet*. 1977; 638-642.
23. Joost H G, Bell G I, Best J D, Birnbaum M J, Charron M J. Nomenclature of the GLUT/SLC2A family of sugar/polyol transport facilitators. *American Journal of Physiology-Endocrinology and Metabolism*. 2002; 282(4): E974-976.
24. Li H, Diao S, Li J, Ma B and Yuan S. An updated meta-analysis of 23 case-control studies on the association between Mir-34b/c polymorphism and cancer risk. *Oncotarget*. 2017; 8(17): 28888-28896.
25. Lim L P, Lau N C, Garrett-Engle P, Grimson A, Schelter J M. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature*. 2005; 433(7027): 769-773.
26. Lin Y, Liu Z, Luan H, Sun M, Rao S, Liu S. Modeling relation paths for representation learning of knowledge bases. *arXiv preprint arXiv*. 2015; 1506.00379.
27. Liu Y, Cao X. Characteristics and significance of the pre-metastatic niche. *Cancer cell*. 2016; 30(5): 668-681.
28. Liu C, Kelnar K, Liu B, Chen X, Calhoun-Davis T. The microRNA Mir-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. *Nat Med*. 2011; 17(2): 211-215.
29. Luan S, Sun L, Huang F. MicroRNA-34a: a novel tumor suppressor in p53-mutant glioma cell line U251. *Arch Med Res*. 2010; 41: 67-74.
30. Mount D. W. *Bioinformatics: sequence and genome analysis*. Cold Spring Harbor Laboratory Press. 2001.
31. Mraz M, Malinova K, Kotaskova J, Pavlova S, Tichy B, Malcikova J. MiR-34a, miR-29c and miR-17-5p are downregulated in CLL patients with TP53 abnormalities. *Leukemia*; 2009; 23(6): 1159-63.
32. Munjal G, Hanmandlu M, and Srivastava S. *Phylogenetics Algorithms and Applications*. Ambient Communications and Computer Systems, 2018; 904: 187–194.
33. National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *diabetes*, 1979; 28(12): 1039-1057.

34. Okamura K, Hagen J W, Duan H, Tyler D M, Lai E C. The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell*. 2007; 130(1): 89-100.
35. Olivier M, Hollstein M, Hainaut P. TP53 mutations in human cancers: origins, consequences, and clinical use. *Cold Spring Harb Perspect Biol*. 2010; 2(1): a001008.
36. Pan X M, Sun R F, Li Z H, Guo X M, Qin H J and Gao L B. Pri-Mir-34b/c rs4938723 polymorphism is associated with a decreased risk of gastric cancer. *Genet Test Mol Biomarkers*. 2015; 19(4): 198-202.
37. Pu X, Wang T J, Huang X, Melas D, Zanis P, Papanastasiou D K, Poupkou A. Enhanced surface ozone during the heat wave of 2013 in Yangtze River Delta region, China. *Science of the Total Environment*. 2017; 603-604: 807-816.
38. Rodriguez A, Griffiths-Jones S, Ashurst J L and Bradley A. Identification of mammalian microRNA host genes and transcription units. *Genome Res*. 2004; 14(10a): 1902-1910.
39. Raver-Shapira N, Efi M, Eti M, Yael S, Nitzan R, Neta M, Zvi B, and Moshe O. 2007. Molecular Cell Article Transcriptional Activation of Mir-34a Contributes to p53-Mediated Apoptosis. *Science*. 2007; 26(5): 731-743.
40. Park, Rehovot 76706, Israel j. molcel. Israel 2 Rosetta Genomics, 10 Plaut Street.
41. Raver-Shapira N, Marciano E, Meiri E, Spector Y, Rosenfeld N. Transcriptional activation of miR-34a contributes to p53-mediated apoptosis. *Mol Cell* 2007; 26(5): 731-43.
42. Rokavec M, Öner MG, Li H, Jackstadt R, Jiang L, Lodygin D. IL-6R/STAT3/miR-34a feedback loop promotes EMT-mediated colorectal cancer invasion and metastasis. *J Clin Invest*; 2014; 124(4): 1853-1867.
43. Rottiers V, Näär A M. MicroRNAs in metabolism and metabolic disorders. *Nature reviews Molecular cell biology*. 2012; 13(4): 239-250.
44. Scanlon D M, Gelzheiser L M, Vellutino F R, Schatschneider C, Sweeney J M. Reducing the incidence of early reading difficulties: Professional development for classroom teachers versus direct interventions for children. *Learning and individual differences*; 2008; 18(3): 346-359.
45. Scott A D, and Baum D. A. Phylogenetic Tree. *Encyclopedia of Evolutionary Biology*. 2016; 270-276.
46. Sicree R. The global burden of diabetes. *Diabetes Atlas*. 2003.
47. Siemens H, Jackstadt R, Hunten S, Kaller M, Menssen A, Gotz U. MiR-34 and SNAIL form a double-negative feedback loop to regulate epithelial-mesenchymal transitions. *Cell Cycle*. 2011; 10(24): 4256-4271.
48. Sun Y, Peng R, Li A, Zhang L, Liu H, Peng H. Sequence variation in microRNA-34a is associated with diabetes mellitus susceptibility in a southwest Chinese han population. *Int J Exp Pathol*. 2018; 11(3): 1637-1644.
49. Verheul TC, van Hijfte L, Perenthaler E, Barakat T S. "The Why of YY1: Mechanisms of Transcriptional Regulation by Yin Yang 1". *Frontiers in Cell and Developmental Biology*. 2020; 8: 592164.
50. Wang A M, Huang T T, Hsu K W, Huang K H, Fang W L, Yang M H. Yin Yang 1 is a target of microRNA-34 family and contributes to gastric carcinogenesis. *Oncotarget*. 2007; 5: 5002-5016.
51. Wei J S. The MYCN oncogene is a direct target of Mir-34a. *Oncogene*, 2008 Sep 4;27(39): 5204-13.
52. Wong M Y, Yu Y, Walsh W R, Yang J L. MicroRNA-34 family and treatment of cancers with mutant or wild-type p53 (Review). *Int J Oncol*. 2011; 38(5): 1189-1195.
53. Xiong J. *Essential Bioinformatics* Cambridge: Cambridge University Press. 2006.
54. Yamakuchi M, Lowenstein C J. Mir-34, SIRT1 and p53: the feedback loop. *Cell Cycle*. 2009; 8(5): 712-715.
55. Yamakuchi M, Ferlito M, Lowenstein C J. MiR-34a repression of SIRT1 regulates apoptosis. *Proc Natl Acad Sci USA*. 2008; 105(36): 13421-13426