LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS

MFE	minimum folding energy
si RNA	Silence inducing RNA
miRNA	microRNA
PTGS	post translational gene silencing
RISC	RNA-induced silencing complex
BCL2	B cell lymphoma 2 (BCL2)
UTR	untranslated region
DGCR8	DiGeorge syndrome Critical Region gene 8
RNA	induced silencing complex
PTEN	phosphatase and tensin homology deleted on chromosome ten
Ago2	Argonaute 2
CLL	chronic lymphocitic leukemia
OPN	osteopontin OPN
Skp2	S-phase kinase-associated protein 2 (skp2)
SVM	support vector machine

Mathematical modelling of miRNAs involved in cancer

Prashant kumar vaishla

Delhi Technological University, Delhi, India

Abstract

miRNAs are small ~21 nucleotide, non-coding, endogenous RNA molecule that has regulatory role in the gene expression of plants and animals .miRNAs regulates gene expression by translational repression, mRNA cleavage, and mRNA decay at post transcriptional level. Recent studies have identified the role of miRNAs in different cancer formations. In addition, miRNAs have been found to function as cancer suppressors. Difference in the expression level of certain miRNAs have been related to the promotion of cancer by negative regulation of tumor suppressor genes. This suggests that miRNAs has potential role in cancer therapy as well as in diagnosis. Therapeutics evaluation of these miRNA requires *in vitro* and *in vivo* RNAi based studies. In the current study we have suggested a computational approach to identify these cancer related miRNAs in humans. This approach can be used to skew the dataset of available miRNA for laboratory validation. Further it can be used for prediction of new miRNAs and their role in different cancers.

In this approach we have selected all the miRNAs that are known to be involved in cancer from miR2disease database. Then, different features of the miRNA are identified at different developmental stages (pri-miRNA, pre-miRNA and miRNA) and these features (descriptors) were utilized for the specific classification of cancer and/or non-cancer miRNA. Mathematical modelling of miRNAs involved in cancer includes the finding of cancer and non-cancer miRNAs dataset, for positive and negative control. Then, we identified the sequence, structure and the energy based descriptors. Extraction of values of different descriptors on the basic of structural and sequential features. Computer program was developed to extract these values. Mfold were used to predict secondary structure and libraries of RNAfold were implemented in program to calculate the minimum folding energy. Classification algorithms of WEKA software were used to classify the two datasets. Random Forest, MultilayerPerceptron and J48 classification algorithms has shown higher accuracy.

Introduction:

RNAi has been an evolutionary conserved process for post translational gene silencing (PTGS) (Fumitaka Takeshita *et al*., 2006). Silence inducing RNA (si RNA) and microRNA (miRNA) plays a key role in the process. These RNAs binds to the specific targets in the miRNAs and either decrease or increase their expression. Binding of miRNAs generally shows incomplete homology with its target (Fumitaka Takeshita *et al*., 2006). The process of formation of miRNA initiates with the transcription of long RNA molecule inside the nucleus ,called primary miRNA(pri-miRNA).pri-miRNA is processed into double stranded stem loop RNA by a complex consisting of Drosha (RnaseIII enzyme) and DGCR8(double stranded RNA bonding protein). This stem loop RNA molecule is 60-80 nucleotide long and known as pre-miRNA. pre-miRNA is transported to cytoplasm and further processed by Dicer into mature miRNA (19-25 nucleotides). RNA-induced silencing complex (RISC), a multisub unit ribonucleoprotein forms around the miRNA and takes up one strand of miRNA, which acts as a guide strand for its complimentary sequence in mRNA

microRNAs binds to 3` UTRs regions of their respective mRNAs and block its translation. Individual miRNA can have multi-targets in different mRNAs. They are responsible to regulate multi gene functions. Modulating the expression of a single mRNA can influence an entire network and thereby modify complex disease phenotypes. Recent studies have established that mRNAs play critical role in post-transcriptional gene expression in higher eukaryotes. Evidences for conservation of plant miRNAs has come from genomes and EST sequences data from diverse plants showing sequences containing miRNA hairpins (Kurihara Y, Watanabe,Jones-Rhoades MJ,Bartal Dp,2004).miRNA has regulatory role in neural development ,cell division ,cell death and other cell differentiation processes (Ankur Omer et. al. ,2012). miR-15 and miR-16 targets antiapoptotic gene B cell lymphoma 2 and induces apoptosis(Cimmino *et al*.,2005).Recent evidences suggest that miRNAs also functions as tumor suppressors and oncogenes.

microRNAs as oncogenes

Elevation in the level of miRNA in tumor that results in the inhibition of tumor suppressor gene. Furthermore miRNA promotes tumor development and called oncomir. These miRNAs often promotes tumor by controlling cell differentiation or apoptosis. **Sanchez-Beato *et al***., 2003 identified that miR-15 and miR-16 inhibits the expression of B cell lymphoma 2 (BCL2) gene which promotes human cancers, involving lymphomas and carcinomas. Expression of miR-15a and miR-16a were found down-regulated in B-cell chronic lymphocytic leukemia (B-CLL) (**Calin *et al*.**, 2002).In Breast cancer, the expression profile of miR-145, miR-155, and miR-125a has significant difference from the normal tissue (**lorio *et al*.**, 2005).

microRNAs as tumor suppressor

Tumor suppressor miRNAs negatively inhibiting oncogenes that control cell differentiation or apoptosis. Currently, several miRNAs are considered as tumor suppressor genes, for example, miRNA let-7. Inappropriate expression of let-7 results in oncogenic loss of differentiation. Studies has found that the expression level of let-7 decrease in lung cancer (**Takamizawa *et al***., 2004). Further experiments has concluded that induced overexpression of miRNA gene let-7 in lung adenocarcinoma cell lines prevent cancer growth(**Takamizawa *et al***. (2004)).Interestingly, it was found that let-7 targets RAS oncogene by binding to its 3' UTR site(**Johnson *et al***.,2005). Decrease in the expression of RAS oncogene by let-7 suppresses promotion of lung cancer.

REVIEW OF LITERATURE

Biogenesis of miRNAs involve transcription of mRNAs by RNA polymerase II into a primary transcript called pri-miRNAs. Primary miRNA is a stem like structure with two complimentary strands and a single stranded connecting loop (figure 1). pri-miRNA also contains 7-methyl guasonise cap at 5'end and a polyA tail on 3` end(Cai X *et al*., 2004;Lee Y *et al*., 2004).

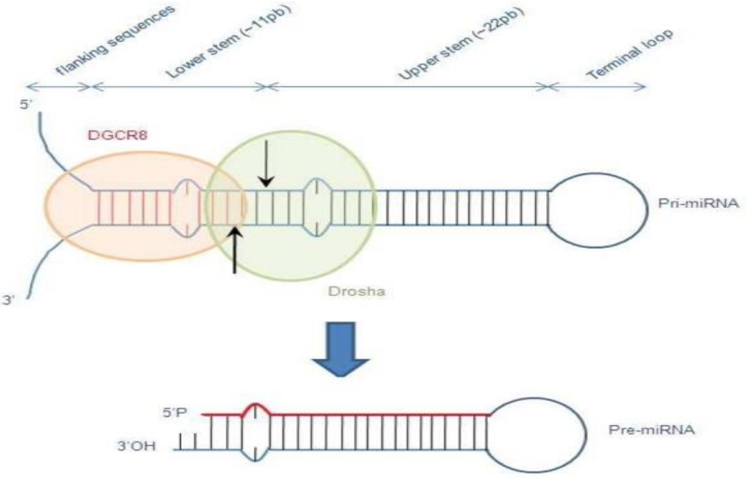

Fig. 1 processing of pri-miRNA into pre-miRNA by Drosha and DGCR8 (**Lee Y *et al***., 2004)

In humans primary miRNA is processed inside the nucleus into pre-miRNA by a protein complex called microprocessor complex. This microprocessor complex is composed of RNAse III Drosha and its cofactor DGCR8 (DiGeorge syndrome Critical Region gene 8) (Gregory RI *et al*., 2004). DGCR8 which binds to double stranded RNA molecule, promotes the binding of Drosha/DGCR8 complex at the junction between flanking region and stem loop (figure 1). After binding of the complex to pri-miRNA Drosha specifically cleaves the pri-miRNA at a distance ~11 bp away from the junction. This cleavage is asymmetric, leaves 3` overhanging (2-unpaired bases) (figure 1) and resulting structure is called pre-miRNA(Lee Y *et al*., 2003). Precursor molecule is then transported from nucleus to cytoplasm with exportin-5 receptor, which is the member of karyopherin beta family. Receptor binds to RNA-GTP protein in the nucleus and then binds to nucleoporine to promote the transfer of pre-miRNA from nucleus to the cytoplasm. Exportin-5 specifically interacts with dsRNAs that has sequence length longer than 14 and contains 3` protruding, hence the binding of pre-miRNA with exp-5/RAN-GTP complex is specific (Zeng Y *et al*., 2004). In the cytoplasm a ribonuclease III enzyme Dicer process pre-miRNA. Dicer is composed of one double-stranded RNA binding domain, two RNAse III domains, one PAZ signature, and a helicase domain. PAZ domain facilitates the recognition of single stranded 3` overhangs of pre-miRNA (Song JJ *et al*., 2003; Patel DJ at al., 2004). Furthermore, dsRNA binding domain promotes the recognition of stem by Dicer and cleavage by RNAse III domain. This cleavage again leads to a 3` overhanging and a miRNA - miRNA* duplex (figure 2). This duplex contains 2nt-3` overhangs on both the ends. RNA induced silencing complex (RISC) which is composed of Dicer/TRBP (TAR (HIV-1) RNA binding protein)/Ago2 (Argonaute 2) complex acts on miRNA - miRNA* duplex (Chendrimada *et al*., 2005, MacRae IJ *et al*., 2008). Unwinding of duplex and selection of one strand occurs in RISC the selected strand (miRNA) is called guide strand and unselected strand is called passenger strand.

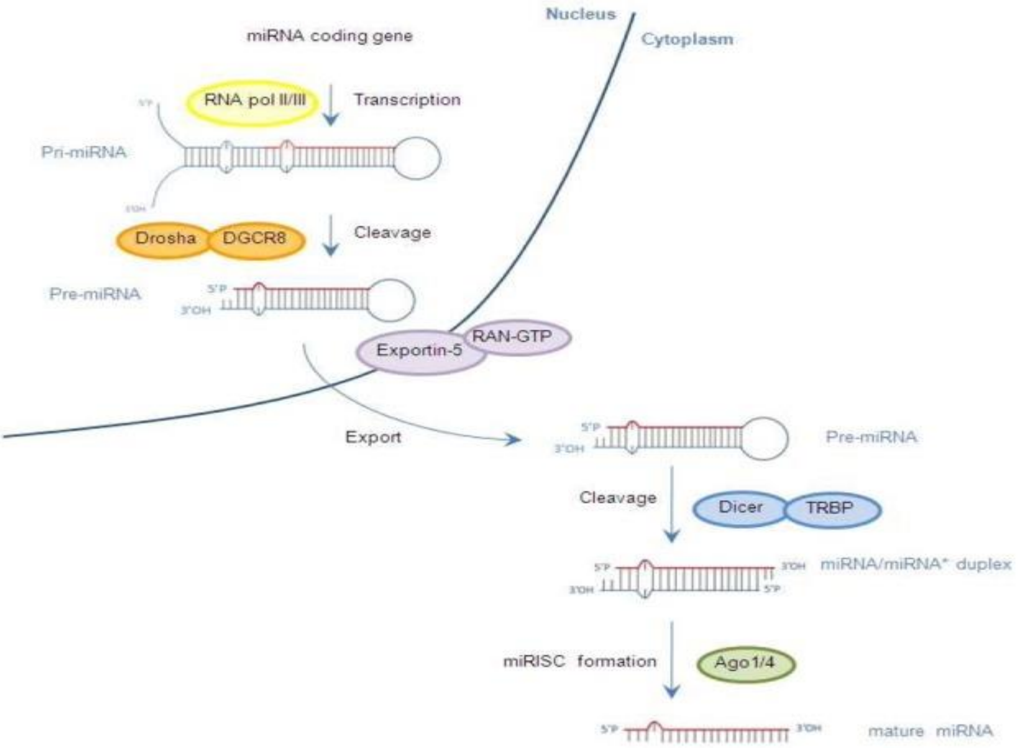

Fig 2: Biogenesis mechanism of micro-RNA (Lee Y *et al*., 2004)

miRNA binds at 3` UTR of its target mRNA. The bind on miRNA is perfectly complimentary in plants. But on the other hand in animals it is not always the case. Studies have reported that the hybridization between miRNA and its target location in mRNA has to be perfectly complimentary between 2 - 7 nucleotide. This region in miRNA called seed region (Lewis *et al, 2003,* Doench *et al*., 2004). Any mismatch in the region leads to the decrease in the efficiency of miRNA.

Recent studies prove that miRNAs play a direct or indirect role in development and/or pathogenesis of cancer in humans. Some miRNAs directly control cell differentiation and apoptosis while others involve in targeting tumor suppressors and/or oncogenes. Increase or decrease in the level of specific miRNA in particular cancer also suggests that miRNAs can be a new biomarker for cancer diagnosis (table 1). In lung cancer the expression of miR-17-92 increases aggressively and studies has found that miR-17-92 cluster targets PTEN (phosphatase and tensin homology deleted on chromosome ten) and RB2 tumor suppressor genes, which leads the progression of lung cancer (Hayashita *et al*., 2005; He *et al*., 2005b). Another example is identified in chronic lymphocitic leukemia (CLL). In CLL deletion in chromosome 13q14 has been associated with the development and progression of tumor. Experiments has revealed that deletion is not associated with a protein coding gene, this critical region contains two miRNA gene, miR-15a and miR-16-1 (Calin *et al*., 2002). Based on their initial observation Calin *et al*., 2004 mapped all the miRNAs and found that many of the miRNAs genes were located on chromosome loci prone to deletion or amplification. When miRNA negatively regulates the tumor suppressor gene, over amplification of chromosome loci results in the increase in expression of miRNA and suppression of target tumor suppressor gene. Conversely, when miRNAs that suppress oncogenes are located on fragile loci, deletion or mutation in the loci results in the decrease in expression of miRNA and over expression of target oncogenes. Therefore, miRNAs expression level governs the development and progression of human cancer.

Cancer	miRNA	Reference
Brain cancer	miR-2(up),miR-221 (up),miR-181(dwn)	Lorio *et al*., 2005 Chan *et al*.,2005 Ciafree at al., 2005
Brest cancer	miR-125b(dwn),miR-145(dwn),miR-21(dwn),miR-155(dwn)	Lorio *et al*., 2005 Chan *et al*.,2005 Ciafree at al., 2005
Chronic lymphocytic leukemia	miR-15(dwn),miR16(dwn)	Calin at al., 2002,2004a;Cimmino *et al*., 2005
Colorectal neoplasia	miR-143(dwn),miR-145(dwn)	Michael *et al*., 2003
Hepatocellular carcinoma	miR-18(up),miR-224(up),miR-199(dwn),miR-195(dwn),miR-200(dwn),miR-125(dwn)	Murakami *et al*.,2006
Lung cancer	Let-7(dwn),miR-19-92(up)	Takamizawa *et al*.,2004; Johnson *et al*., 2005; Hayashita *et al*., 2005;
Lymphomas	miR-155(up),miR-17-92(up)	Eis *et al*., 2005;Metzler *et al*.,2004;He *et al*.,2005b
Papillary thyroid carcinoma	miR-221(up),miR-222(up),miR-146(up),miR-181(up)	He *et al*., 2005b;Pallante *et al*., 2006
glioblastoma	miR-26a (up)	Huse *et al*. 2009
Testicular germ cell	miR-372(up),miR-373(up)	Voorhoeve *et al*.,2006

Table 1: cancer related miRNAs; up denotes that the expression level of miRNA increased; dwn represents decrease in expression level.

Evidences have proven that RNAi has therapeutic role in cancer suppression. It has shown that the expression level of osteopontin (OPN) protein increases in gastric cancer and it is also associated with cancer pathogenesis (Wu CY *et al*., 2007). Successful RNAi was developed to inhibit the expression of OPN in human gastric cancer and tested in *in vitro* and *in vivo* (song He *et al*., 2009). X-linked inhibitor of apoptosis protein (XIAP), a potent inhibitor of caspase can be suppressed by using antisense oligonucleotides (Hu Y *et al*., 2003). Cyclooxygenase-2 (COX-2) is an enzyme which is responsible for genesis and progression of colorectal cancer (L. Crofford, *et al*., 1998). RNAi mediated study models has been used in vitro to study the molecular and phenotypic effects. Treatment with siCOX-2, shCOX-2 has proven effective in silencing COX-2 protein (Antonio Strillacci *et al*., 2010). miR-101, miR-199a and siCOX-2 could give a highly specific and strong alternative therapeutics solution in future (Antonio Strillacci *et al*., 2010). Over-expression of human S-phase kinase-associated protein 2 (skp2), a specific ubiquitin ligase subunit has been associated with various cancer. Studies have shown that skp2 targets $p27^{Kip1}$ of the cell cycle and negatively regulates its expression (Carrano *et al*., 1999). Experiments has validated that inhibition of skp2 gene leads to increase the expression of $p27^{Kip1}$ protein (Chiarle *et al*., 2004). Tca8113 cells infected to express high level of Skp2 were found to show low level of Skp2 after the treatment with small interfering RNA expression plasmid vector. Furthermore, cell had shown increase in the level of $p27^{kip1}$ protein and suppressed tumor proliferation (Fang *et al*., 2008).

Computational prediction of miRNA often involves the identification of pre-miRNA structure. Mfold (Zuker *et al*., 2003) and RNAfold (Hofacker *et al*., 2003) are the tools available for prediction of secondary structure and MFE of pre-miRNA. It has been found that structural RNA has lower free energy than random (Peter clote *et al*., 2005).

Available miRNA prediction tools includes ProMir (Nam *et al*., 2005, 2006), miPred (Jiang *et al*., 2007), microPred (Batuwita *et al*., 2009), TripletSVM (Xue *et al*., 2005), SSCprofiler (Oulas er al., 2009), SplamiR (Thieme *et al*., 2011), MiRenSVM (Griffiths-Jones *et al*., 2006), miRFinder (Bonnet *et al*., 2004), HHMMiR (Kadri *et al*., 2009), miRabela (Sewer *et al*., 2005) and many others. All these tools are trained on validated miRNA sequences of miRBase database.

Mathematical modelling

It is the description of a real world system on terms of mathematical equations or language. On the practical basic it involves computer science, statistics and engineering concepts. Mathematical modelling develops interpretation from the input data and use it further for certain predictions related to that data. It is composed of different variables and their relations. Variables defines different features of the data set and their values depend on the occurrence of events. Different operators describes their relationship. Building a model requires the collection of data, extracting values of features (descriptors), processing the input file, training the algorithm, testing the algorithm and implementing the tested algorithm into the computer program.

In life science, simulation of biological systems (e.g. signal transduction, gene expression network, metabolite network, drug pharmacodynamics, protein folding, and molecular modelling) are based on mathematical modelling or machine learning concepts.

There are numerous mathematical models present for the study of different phases for cancer progression (Adam *et al*., 1996; Chaplain *et al*., 2000). Broadly they are classified as discrete cell-based models and continuum models.

METHODOLOGY

Collection of data

The initial step for the process is to collect miRNA that are known to involve in human cancer. MicroRNA ids are downloaded from miR2disease database (http://www.mir2disease.org/). miR2disease is a manually curated online available database. It provides the list of miRNAs and their disease in which they are involved with references. But miR2disease does not provides the sequence details of these miRNA. Sequences are downloaded from another database called miRBase (http://www.mirbase.org/). It is an online available database wich provides sequence datails of miRNA. Database has details of miRNA-ids, primary miRNAs, miRNA, miRNA*, and secondary structure. Primary miRNAs and miRNA sequence dataset are downloaded from mirbase and sequences of pri-miRNAs and miRNAs are extracted using SHELL scripts (Appendix a). Further non cancer miRNAs are sorted from the dataset.

Identification of descriptors

Identification of descriptors was the key step in the development of model. Different miRNA prediction software had implemented structure, sequence and energy based descriptors. In the currents study we have involved most of the known descriptors and few to best of our knowledge new descriptors.

Extraction of descriptors values

Programming is written in C language and the automation part has done by using SHELL scripts. Prediction of secondary structure of pri-miRNA and the calculation of its MFE is done using Mfold and RNAfold softwares. Program is designed in such a way that it has built-in libraries of RNAfold (Appendix), so program autonomously uses these libraries to calculate the MFE of provided sequence. Programme takes three input files pri-miRNA sequence, miRNA, miRNA* and pri-miRNA.out (output file of mfold, having predicted secondary structure). Then, program generates the vales of descriptors based on the given input.

a) Pri-miRNA	**b) has-miR-187-5p**
GGUCGGGCUCACCAUGACACAGUGUGAGACCUCGG GCUACAACACAGGACCCGGGCGCUGCUCUGACCCC UCGUGUCUUGUGUUGCAGCCGGAGGGACGCAGGUC CGCA	GGCUACAACACAGGACCCGGGC
c) has-miR-187-3p	**d) Secondary structure**
UCGUGUCUUGUGUUGCAGCCGG	

```
10        20        30        40        50        60
G|U      CACCAUGACACAG    AGA    G    A           C    CGCUGC
 G CGGGCU             UGUG   CCUC GGCU CAACACAGGAC CGGG      U
 C GCCUGG             ACGC   GGAG CCGA GUUGUGUUCUG GCUC      C
A^-      -------------    AG-    G    C           U    CCCAGU
                     100         90        80        70
```

Table2 Input files for program a) primary miRNA b) miRNA sequence at 5` c) miRNA* sequence at 3` and d) output file of Mfold.

Training set and Classsification:

Further steps involves the division of data into training set and test set. Then, transformation of data file into *.arff file and uploading data into WEKA pre-processor. Weka classifiers has different classification algorithms. Next step is to train the model on the basis of training set. This model is used to test the test datset.

RESULTS

Descriptors that are known to predict candidate miRNAs were identified. These descriptors were categorised into three classes a) sequence based b) structure based and c) energy based.

Table3: List of descriptors

A. Sequence based descriptors

Sequence based information is crucial for the development of model. Study of sequences provide information about the complexity, variation, stability, tandem repeats, contents of nucleic acids and variation in length.

pri-seq	Sequence details of primary miRNA
Explanation: A, U, G, C, GC content present in pri-miRNA divided by its length and G/C ratio of pri-miRNA (Freyhult *et al*. 2005). High G/C ratio indicates high incidence of GU- paring which is energetically more favourable. In hairpin sequences that fold into loop, G/C ratio was much lower as compared to miRNA sequence.	

pre-seq	Sequence details of pre-miRNA
Explanation : A,U,G,C,GC content present in pre-miRNA divided by its length and G/C ratio of pri-miRNA	

mi-RNA seq	Sequence details of pre-miRNA
Explanation : A,U,G,C,GC content present in miRNA divided by its length	

GAsurplusCU	surplus of GA over CU in sequence
Explanation : GAsurplusCU measures the relative abundance of Guanines and Adenosines in a sequence relative to the number of Cytosines and Uracils. With g, c, a and u as the number of Guanines, Cytosines, Adenosines and Uracils in the sequence, (Ate ven der burgt *et al*., 2009) GAsurplusCU is calculated by: $GAsurplusCU = (g + a - c - u) / (g + a + c + u)$	

Poly A hairpin	longest poly-A stretch in hairpin sequence (nt)
Explanation : count of number of adenosines present in the hairpin sequence at a stretch (Ate ven der burgt *et al*., 2009)	

Poly A miRNA	longest poly-A stretch in miRNA sequence (nt)
Explanation : count of number of adenosines present in the miRNA sequence at a stretch	

Poly U hairpin	longest poly-U stretch in hairpin sequence (nt)
Explanation : count of number of uracil present in the hairpin sequence at a stretch (Ate ven der burgt *et al*., 2009)	

Poly U miRNA	longest poly-U stretch in miRNA sequence (nt)
Explanation : count of number of uracil present in the miRNA sequence at a stretch	

Poly G hairpin	longest poly-G stretch in hairpin sequence (nt)
Explanation : count of number of guanine present in the hairpin sequence at a stretch (Ate ven der burgt *et al*., 2009)	

Poly G miRNA	longest poly-U stretch in miRNA sequence (nt)
Explanation : count of number of guanine present in the hairpin sequence at a stretch	

Poly C hairpin	longest poly-C stretch in hairpin sequence (nt)
Explanation : count of number of cytosine present in the hairpin sequence at a stretch (Ate ven der burgt *et al*., 2009)	

Max di base ratio	Highest di based ratio
Explanation : count of two highest occurring bases in the sequence, divided by the length if the sequence	

Min base occurrence	Minimal base occurrence
Explanation : count of least occurring base divided by the length of the sequence	

b. Structure based

Secondary structure of pri-miRNA consists of central loop, upper stem, lower stem and flanking sequences. Bulges within these sequences decides the position of miRNA in pri-miRNA sequence. Wobber matches in secondary structure were also considered.

dP	Adjusted base pair propensity

Explanation: Total number of base pairs present in the structure divided by the length of the sequence in nucleotide (Schultes *et al.*, 1999). It removes the biasing of long sequence tends to have more base pairing (Loong and Mishra 2007).

Cloop	Central loop length
Explanation : number of nucleotides present in the central loop of secondary structure of hairpin	

wobber vs. all matches (stem)	Ratio of wobber to all the matches
Explanation : number of wobber base paring allowed in the stable secondary structure divided by the total number of matches	

Bulge ratio stem	Ratio of bulge to length of stem
Explanation : ratio of asymmetric bulge to the length of stem	

Bulge ratio miR	Ratio of bulge to length of stem
Explanation : ratio of asymmetric bulge to the length of stem	

B. Energy

MFE	Minimum folding energy (kcal/mol)
Explanation: it is the thermodynamical measure for the stability of RNA. Libraries of RNAfold (Hofacker 2003) and Mfold (Zuker 2003), were used to calculate the mfe of structure. Lower negative value of mfe results in stable structure	

MFEpri index	MFE of pri-miRNA corrected to gc content
Explanation : mfe of the primary miRNA divided by L*GC	

MFEpril(dG)	MFE corrected to GC
Explanation : mfe divided by GC count	

MFEhai index	MFE of hairpin corrected to gc content
Explanation : mfe of the hairpin divided by L*GC	

MFEhail(dG)	MFE corrected to GC
Explanation : mfe divided by GC count	

Classification results

Validated cancer related miRNA were identified from mir2disease database. Sequences of these miRNAs were extracted from miRBase database. Values of 43 descriptors were extracted for 183 cancer related miRNAs sequences. Same number of sequences that are not related to cancer were processed to extract the values of descriptors. These sequences were validated from miRBase database.

Table4: result of different classifiers on the basic of training and test set

Classifier	Accuracy(training)	Accuracy(test)
LIBsvm	77.4545 %	66 %
MultilayerPerceptron	99.2727 %	67 %
SimpleLogistic	70.9091 %	63 %
SMO	70.9091 %	66 %
RandomForest	99.6364 %	71 %
DecisionTable	73.4545 %	69 %
oneR	74.1818 %	63 %
J48	96.3636 %	70 %
RandomCommittee	100 %	66 %

Details of the results of different classification algorithms

TP rate = true positive rate (rate is the proportion of examples which were classified as class x, among all examples which truly have class x)
FP rate=false positive rate (rate is the proportion of examples which were classified as class x, but belong to a different class, among all examples which are not of class x.)
The Precision is the proportion of the examples which truly have class x, among all those which were classified as class x.

F-Measure is simply 2*Precision*Recall/(Precision+Recall), a combined measure for precision and recall.

LIBsvm

```
                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
Class
                 0.812    0.261    0.745      0.812    0.777      0.552    0.776     0.696
cancer
                 0.739    0.188    0.808      0.739    0.772      0.552    0.776     0.732
noncancer

Weighted Avg.    0.775    0.223    0.777      0.775    0.774      0.552    0.776     0.714
```

```
Confusion Matrix ===

   a   b   <-- classified as
 108  25 |   a = cancer
  37 105 |   b = noncancer
```

multilayerPerceptron

```
                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
class
                 0.992    0.007    0.992      0.992    0.992      0.985    0.987     0.963
cancer
                 0.993    0.008    0.993      0.993    0.993      0.985    0.987     0.984
noncancer
Weighted Avg.    0.993    0.007    0.993      0.993    0.993      0.985    0.987     0.974
```

```
=== Confusion Matrix ===

   a   b   <-- classified as
 132   1 |   a = cancer
```

```
   1 141 |   b = noncancer
```

SimpleLogistic

```
     class           TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
                      0.759    0.338    0.678      0.759    0.716      0.423    0.785     0.755
    cancer
                      0.662    0.241    0.746      0.662    0.701      0.423    0.785     0.793
    noncancer
    Weighted Avg.     0.709    0.288    0.713      0.709    0.709      0.423    0.785     0.774
```

```
=== Confusion Matrix ===

   a   b   <-- classified as
 101  32 |   a = cancer
  48  94 |   b = noncancer
```

SMO

```
     class            TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
    cancer            0.812    0.387    0.663      0.812    0.730      0.432    0.712     0.629
    noncancer         0.613    0.188    0.777      0.613    0.685      0.432    0.712     0.676
    Weighted Avg.    0.709    0.284    0.722      0.709    0.707      0.432    0.712     0.653
```

```
=== Confusion Matrix ===

   a   b   <-- classified as
 108  25 |   a = cancer
  55  87 |   b = noncancer
```

Random forest

```
                     TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
    class             1.000    0.007    0.993      1.000    0.996      0.993    1.000     1.000
    cancer
                      0.993    0.000    1.000      0.993    0.996      0.993    1.000     1.000
    noncancer
    Weighted Avg.     0.996    0.003    0.996      0.996    0.996      0.993    1.000     1.000
```

```
=== Confusion Matrix ===

   a   b   <-- classified as
 133   0 |   a = cancer
   1 141 |   b = noncancer
```

Decision table

```
                          TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC
   Area
   class                   0.737    0.268    0.721      0.737    0.729      0.469    0.779     0.701
   cancer                  0.732    0.263    0.748      0.732    0.740      0.469    0.779     0.774
   noncancer
   Weighted Avg.           0.735    0.265    0.735      0.735    0.735      0.469    0.779     0.739
```

```
=== Confusion Matrix ===

   a   b   <-- classified as
  98  35 |   a = cancer
  38 104 |   b = noncancer
```

oneR

```
                          TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area
   class                   0.782    0.296    0.712      0.782    0.746      0.487    0.743     0.662
   cancer                  0.704    0.218    0.775      0.704    0.738      0.487    0.743     0.699
   noncancer
   Weighted Avg.           0.742    0.256    0.745      0.742    0.742      0.487    0.743     0.681
```

```
=== Confusion Matrix ===

   a   b   <-- classified as
 104  29 |   a = cancer
  42 100 |   b = noncancer
```

J48

```
                   TP Rate  FP Rate  Precision  Recall   F-Measure    MCC    ROC Area  PRC Area
   class            0.962    0.035    0.962      0.962    0.962      0.927    0.990     0.986
  cancer
                    0.965    0.038    0.965      0.965    0.965      0.927    0.990     0.988
  noncancer
  Weighted Avg.     0.964    0.036    0.964      0.964    0.964      0.927    0.990     0.987
```

```
=== Confusion Matrix ===

   a   b   <-- classified as
 128   5 |   a = cancer
   5 137 |   b = noncancer
```

RandomCommittee

```
                        TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC
   Area
   class                  1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000
   cancer                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000
   noncancer
   Weighted Avg.           1.000    0.000    1.000      1.000    1.000      1.000    1.000
   1.000
```

```
=== Confusion Matrix ===

   a   b   <-- classified as
 133   0 |   a = cancer
   0 142 |   b = noncancer
```

Accuracy and kappa statistic value of algorithm showing higher accuracy (> 95% on training set)

Classifier	Accuracy(training)	Accuracy (test)	Kappa statistic (training)	Kappa statistic (test)
RandomForest	99.6364%	71%	.9927	0.42
MultilayerPerceptron	99.2727%	67%	0.9854	0.34
J48	96.3636%	70%	0.972	0.4

Supplementary data

Table5: results/ output file of developed program (values of 43 descriptors of 183 cancer sequences)

id	pri-mfe	Hairpin-mfe	hairpin length	MFEpri index	MFEpril(dG)	MFEhai index
hsa-let-7b	-50.599998	-42.599998	73	-0.013855	-0.609639	-0.014372
hsa-let-7c	-33.5	-21.9	65	-0.00997	-0.39881	-0.010788
hsa-let-7d	0	-29.1	2	0	0	00:00
hsa-let-7e	-37.799999	-38.700001	66	-0.010633	-0.478481	-0.012227
hsa-let-7f-1	-43.299999	-34.700001	81	-0.01422	-0.497701	-0.013235
hsa-let-7g	-40.5	-36.299999	74	-0.01148	-0.482143	-0.011707
hsa-let-7i	-39.599998	-18.200001	69	-0.010476	-0.471429	-0.011351
hsa-mir-100	-26.799999	-24.200001	54	-0.009054	-0.335	-0.011742
hsa-mir-106a	-34.700001	-25.799999	57	-0.0126	-0.428395	-0.016971

hsa-mir-106b	-44.900002	-22.49	56	-0.010951	-0.547561	-0.012239
hsa-mir-10a	-32.389999	-24.4	47	-0.006848	-0.294455	-0.014874
hsa-mir-10b	-45	-24.9	58	-0.009298	-0.409091	-0.015136
hsa-mir-122	-46.599998	-46	58	-0.014817	-0.548235	-0.019124
hsa-mir-1228	-46	-26.6	66	-0.011457	-0.630137	-0.011998
hsa-mir-125a	-48.099998	-23	58	-0.010967	-0.559302	-0.012619
hsa-mir-126	-40.200001	-42.599998	58	-0.01051	-0.472941	-0.014839
hsa-mir-1266	-58.299999	-25.200001	57	-0.015423	-0.694048	-0.021472
hsa-mir-127	-41.099998	-30.6	57	-0.008308	-0.423711	-0.014016
hsa-mir-1296	-50.200001	-12.6	65	-0.010295	-0.545652	-0.011815
hsa-mir-130a	-42.599998	-19.700001	58	-0.010637	-0.478652	-0.008129
hsa-mir-130b	-36.400002	-27.700001	58	-0.009445	-0.443902	-0.010591
hsa-mir-132	-48.400002	-27.200001	58	-0.007261	-0.479208	-0.014892
hsa-mir-134	-35.200001	-22	57	-0.010959	-0.482192	-0.012535
hsa-mir-135b	-47.200001	-20.5	56	-0.009541	-0.486598	-0.013142
hsa-mir-136	-45.799999	-31.4	53	-0.015096	-0.558537	-0.015745
hsa-mir-139	-34.900002	-32.200001	55	-0.012831	-0.513235	-0.013688
hsa-mir-140	-55	-26.700001	56	-0.010185	-0.55	-0.017312
hsa-mir-141	-48.900002	-25.799999	63	-0.009712	-0.514737	-0.013622
hsa-mir-143	-51.900002	-20.4	49	-0.008027	-0.489623	-0.015636
hsa-mir-144	-44.400002	-26.92	56	-0.012006	-0.516279	-0.016452
hsa-mir-145	-41.619999	-20.9	49	-0.011261	-0.472955	-0.016081
hsa-mir-146a	-40.299999	-24.5	55	-0.009046	-0.407071	-0.013484
hsa-mir-146b	-37.099998	-17.799999	58	-0.012396	-0.508219	-0.013626
hsa-mir-148a	-28.200001	-16.9	55	-0.013378	-0.414706	-0.011484
hsa-mir-148b	-35	-30.5	57	-0.008222	-0.353535	-0.010484
hsa-mir-149	-55.799999	-28.799999	58	-0.009796	-0.626966	-0.011713
hsa-mir-150	-57.5	-22.4	58	-0.012224	-0.684524	-0.012903
hsa-mir-152	-48.200001	-20.4	58	-0.010073	-0.554023	-0.01129
hsa-mir-154	-37.5	-23.299999	53	-0.01313	-0.446429	-0.01371
hsa-mir-155	-29.700001	-24.9	58	-0.017574	-0.456923	-0.016339
hsa-mir-15a	-30.74	-19.1	57	-0.010582	-0.370361	-0.015447
hsa-mir-15b	-32.700001	-22.700001	55	-0.009269	-0.333673	-0.012836
hsa-mir-17	-34.299999	-23.700001	55	-0.011343	-0.408333	-0.01356
hsa-mir-181c	-44.66	-27.5	51	-0.006767	-0.406	-0.011946
hsa-mir-181d	-69.099998	-21.040001	66	-0.006151	-0.50438	-0.011224
hsa-mir-182	-47.139999	-22.5	56	-0.006802	-0.428545	-0.009696
hsa-mir-183	-41.900002	-28.299999	56	-0.007054	-0.380909	-0.013441
hsa-mir-185	-53.099998	-25.139999	56	-0.012453	-0.647561	-0.013041
hsa-mir-186	-41.240002	-19.299999	48	-0.015469	-0.479535	-0.018431
hsa-mir-18a	-22	-18.5	65	-0.009683	-0.309859	-0.009507
hsa-mir-18b	-19.799999	-26.200001	67	-0.008451	-0.278873	-0.008525
hsa-mir-191	-49	-29.49	66	-0.009684	-0.532609	-0.010694
hsa-mir-192	-39.490002	-29	54	-0.005698	-0.359	-0.012037
hsa-mir-193b	-43.099998	-29.74	53	-0.012665	-0.519277	-0.015591
hsa-mir-195	-46.439999	-24.6	48	-0.010894	-0.533793	-0.015473
hsa-mir-196b	-34.299999	-29.4	52	-0.010208	-0.408333	-0.013682
hsa-mir-197	-42.400002	-22.9	58	-0.012028	-0.565333	-0.012816
hsa-mir-199b	-42.900002	-23	58	-0.006724	-0.39	-0.014206
hsa-mir-19a	-39.099998	-28	55	-0.013624	-0.476829	-0.018548
hsa-mir-200a	-47.5	-25.1	57	-0.010556	-0.527778	-0.015573
hsa-mir-200b	-44.299999	-27.799999	55	-0.008181	-0.466316	-0.014994
hsa-mir-200c	-31.299999	-27.76	54	-0.012113	-0.460294	-0.013587
hsa-mir-204	-42.259998	-25.620001	47	-0.007249	-0.384182	-0.015439
hsa-mir-205	-49.02	-26.6	49	-0.008408	-0.445636	-0.014758

hsa-mir-20a	-31	-25.4	56	-0.016171	-0.43662	-0.021452
hsa-mir-20b	-29.799999	-24.9	55	-0.013932	-0.431884	-0.014631
hsa-mir-21	-35.799999	-25	57	-0.014206	-0.497222	-0.014875
hsa-mir-210	-60.700001	-27.5	59	-0.006898	-0.551818	-0.009601
hsa-mir-211	-45.799999	-24.9	57	-0.006406	-0.416364	-0.012673
hsa-mir-212	-50.799999	-34.32	59	-0.005701	-0.461818	-0.010569
hsa-mir-214	-67.620003	-22.1	50	-0.009915	-0.614727	-0.015565
hsa-mir-215	-33.400002	-25.700001	51	-0.00799	-0.303636	-0.016974
hsa-mir-216a	-40.700001	-25.9	58	-0.007115	-0.37	-0.015352
hsa-mir-22	-39.799999	-25.799999	53	-0.010179	-0.468235	-0.014919
hsa-mir-221	-47.200001	-29.200001	58	-0.008252	-0.429091	-0.014626
hsa-mir-222	-54.099998	-25.200001	51	-0.008782	-0.491818	-0.01624
hsa-mir-223	-47.900002	-26.6	61	-0.007381	-0.435455	-0.012414
hsa-mir-224	-36.599998	-26.4	65	-0.012551	-0.451852	-0.013571
hsa-mir-23a	-33.200001	-21.34	52	-0.010828	-0.454795	-0.014194
hsa-mir-23b	-36.040001	-23.700001	48	-0.007583	-0.371546	-0.013768
hsa-mir-25	-37.799999	-23	55	-0.008491	-0.45	-0.010618
hsa-mir-26b	-40.099998	-28.6	56	-0.012111	-0.520779	-0.014268
hsa-mir-27a	-37.5	-25.700001	54	-0.010229	-0.480769	-0.01318
hsa-mir-27b	-50.400002	-21.6	60	-0.011295	-0.519588	-0.014067
hsa-mir-29a	-24.9	-24.6	51	-0.018527	-0.389062	-0.017419
hsa-mir-29c	-35.400002	-20.5	57	-0.010315	-0.402273	-0.015871
hsa-mir-301a	-32.799999	-31.799999	59	-0.010897	-0.381395	-0.013777
hsa-mir-30a	-37.299999	-24.4	59	-0.013825	-0.525352	-0.015774
hsa-mir-30b	-37.599998	-27.299999	50	-0.011244	-0.427273	-0.015136
hsa-mir-30c-1	-35.400002	-26.700001	74	-0.008839	-0.397753	-0.008974
hsa-mir-30d	-28.1	-30.799999	56	-0.012545	-0.401429	-0.01485
hsa-mir-30e	-53	-27.700001	66	-0.012257	-0.576087	-0.015172
hsa-mir-31	-36	-26.200001	54	-0.014085	-0.507042	-0.015956
hsa-mir-32	-31.9	-41.900002	59	-0.016878	-0.455714	-0.018081
hsa-mir-328	-46.299999	-25.219999	59	-0.012105	-0.617333	-0.015467
hsa-mir-330	-46.52	-21.1	49	-0.008682	-0.494894	-0.011622
hsa-mir-335	-41.5	-26	56	-0.012264	-0.441489	-0.017016
hsa-mir-339	-49.200001	-26.9	58	-0.007586	-0.523404	-0.010753
hsa-mir-33a	-37.299999	-19.4	57	-0.016381	-0.54058	-0.014961
hsa-mir-340	-36	-23.9	63	-0.010242	-0.378947	-0.011548
hsa-mir-342	-47.700001	-28	66	-0.009833	-0.481818	-0.011014
hsa-mir-345	-51.299999	-28	58	-0.008872	-0.523469	-0.011885
hsa-mir-34a	-50.700001	-24.799999	66	-0.008231	-0.460909	-0.013793
hsa-mir-34b	-33.799999	-22	57	-0.010317	-0.402381	-0.015385
hsa-mir-34c	-30	-24	51	-0.012568	-0.38961	-0.015385
hsa-mir-363	-25.4	-21.799999	65	-0.010925	-0.338667	-0.012698
hsa-mir-370	-32.599998	-22.6	55	-0.010109	-0.434667	-0.010342
hsa-mir-372	-27.1	-27.940001	51	-0.011896	-0.404478	-0.013018
hsa-mir-373	-35.939999	-25.9	48	-0.014882	-0.52087	-0.015539
hsa-mir-374a	-35.200001	-22.52	50	-0.021256	-0.488889	-0.036895
hsa-mir-376c	-24.02	-21.9	49	-0.019155	-0.363939	-0.019117
hsa-mir-379	-26.6	-29.5	58	-0.014179	-0.397015	-0.014129
hsa-mir-381	-35	-26.5	59	-0.015556	-0.466667	-0.017343
hsa-mir-383	-27.299999	-26.9	54	-0.010388	-0.373973	-0.014247
hsa-mir-424	-42.299999	-24.5	57	-0.008633	-0.431633	-0.018078
hsa-mir-425	-36.299999	-27.6	56	-0.008345	-0.417241	-0.010234
hsa-mir-433	-38.200001	-26.9	72	-0.008215	-0.410753	-0.009563
hsa-mir-452	-40.900002	-47	59	-0.013748	-0.481176	-0.01478
hsa-mir-486	-48.5	-24.5	53	-0.016587	-0.713235	-0.018651

hsa-mir-493	-47.700001	-18.299999	59	-0.011166	-0.535955	-0.01341
hsa-mir-494	-34.099998	-20.4	46	-0.01358	-0.420988	-0.015124
hsa-mir-495	-35.599998	-28.200001	53	-0.014005	-0.434146	-0.014956
hsa-mir-497	-55.5	-24.200001	56	-0.006607	-0.495536	-0.013378
hsa-mir-501	-43.099998	-32.5	53	-0.011402	-0.513095	-0.013011
hsa-mir-503	-42.099998	-23	58	-0.014118	-0.592958	-0.013943
hsa-mir-510	-29.6	-21.200001	58	-0.012121	-0.4	-0.015457
hsa-mir-511	-33	-25.200001	56	-0.011156	-0.37931	-0.017097
hsa-mir-519d	-41.5	-34.900002	54	-0.011502	-0.471591	-0.015054
hsa-mir-520c	-51.299999	-33.900002	57	-0.015119	-0.589655	-0.022516
hsa-mir-616	-51.400002	-41.599998	54	-0.013587	-0.529897	-0.020251
hsa-mir-625	-74.900002	-39.200001	51	-0.021492	-0.881176	-0.026839
hsa-mir-627	-56.5	-29.299999	58	-0.017132	-0.582474	-0.02529
hsa-mir-675	-48.299999	-28.6	45	-0.013233	-0.661644	-0.015221
hsa-mir-874	-37.200001	-32.119999	57	-0.008832	-0.476923	-0.011251
hsa-mir-92b	-66.120003	-23.1	50	-0.009566	-0.68875	-0.013073
hsa-mir-93	-44.799999	-20.6	56	-0.011667	-0.56	-0.011643
hsa-mir-95	-37.5	-29.5	56	-0.014029	-0.462963	-0.018459
hsa-mir-96	-34.400002	-31.4	64	-0.01192	-0.441026	-0.014048
hsa-mir-98	-57	-25.6	75	-0.009392	-0.478992	-0.011778
hsa-mir-99a	-39.599998	-22.1	54	-0.01137	-0.488889	-0.012903
hsa-let-7a-2	-25.200001	-28.700001	65	-0.012069	-0.35	-0.011693
hsa-let-7f-2	-40.700001	-25.799999	73	-0.01401	-0.490361	-0.013141
hsa-miR-125b-1	-43.400002	-25.5	57	-0.010493	-0.493182	-0.013871
hsa-miR-125b-2	-40.599998	-25.200001	58	-0.010368	-0.45618	-0.015233
hsa-miR-128-1	-35.900002	-22.9	55	-0.010678	-0.437805	-0.014728
hsa-miR-128-2	-39.400002	-26.299999	54	-0.00998	-0.469048	-0.011542
hsa-miR-129-1	-29.4	-23.5	66	-0.011667	-0.408333	-0.011741
hsa-miR-129-2	-48.799999	-31.4	60	-0.011296	-0.542222	-0.010829
hsa-miR-138-1	-58.799999	-30.4	57	-0.01024	-0.593939	-0.014896
hsa-miR-138-2	-36	-24.299999	64	-0.008929	-0.428571	-0.011136
hsa-miR-16-1	-37.700001	-22.700001	60	-0.011767	-0.423596	-0.015093
hsa-miR-16-2	-30.799999	-21	62	-0.012675	-0.380247	-0.015442
hsa-miR-181a-2	-54	-23.719999	51	-0.008926	-0.490909	-0.012545
hsa-miR-19b-1	-38.419998	-27.700001	49	-0.011621	-0.441609	-0.014715
hsa-miR-19b-2	-40.599998	-27.1	66	-0.012816	-0.422917	-0.015829
hsa-miR-218-1	-42.900002	-25.4	61	-0.007959	-0.39	-0.012905
hsa-miR-218-2	-46.299999	-19.120001	66	-0.0069	-0.420909	-0.011339
hsa-miR-24-1	-26.32	-20	49	-0.011384	-0.387059	-0.011422
hsa-miR-24-2	-27.299999	-23.9	53	-0.009349	-0.373973	-0.011123
hsa-miR-26a-1	-37.299999	-24.74	55	-0.011265	-0.484416	-0.012849

hsa-miR-26a-2	-42.240002	-23.92	48	-0.012571	-0.502857	-0.016626
hsa-miR-29b-1	-34.02	-24.040001	57	-0.013125	-0.42	-0.015532
hsa-miR-30c-2	-25.24	-21.700001	48	-0.011685	-0.350556	-0.014361
hsa-miR-376a-2	-28.200001	-21.84	57	-0.013056	-0.3525	-0.015909
hsa-miR-29b-2	-31.24	-25	56	-0.012052	-0.385679	-0.013
hsa-miR-7-1	-48	-19.4	66	-0.009486	-0.436364	-0.017007
hsa-miR-7-2	-46.900002	-31.5	50	-0.007614	-0.426364	-0.012516
hsa-miR-92a-1	-36.200001	-25.299999	56	-0.012213	-0.464103	-0.016935
hsa-miR-92a-2	-30.200001		59	-0.011843	-0.402667	-0.015114
hsa-let-7a-3	-34.400002	-30.4	67	-0.013672	-0.464865	-0.014476
hsa-mir-9-3	-41.099998	-28.4	56	-0.011417	-0.456667	-0.017618
hsa-mir-101-2	-31.4	-23.1	58	-0.012421	-0.397468	-0.012222
hsa-mir-105-2	-41.299999	-29	55	-0.012436	-0.509877	-0.015591
hsa-mir-124-3	-40.299999	-22.5	50	-0.009856	-0.463218	-0.013441
hsa-mir-133a-2	-47.5	-23.5	54	-0.008956	-0.465686	-0.013537
hsa-mir-135a-2	-35.099998	0	2	-0.010969	-0.351	0
hsa-mir-153-2	-41.299999	-24.5	55	-0.012492	-0.474713	-0.017181
hsa-mir-181a-1	-38.299999	-22	55	-0.007569	-0.348182	-0.014194
hsa-mir-181b-2	-35	0	2	-0.009592	-0.393258	0
hsa-mir-196a-2	-51.200001	-28.5	59	-0.008782	-0.465455	-0.018387
hsa-mir-199a-2	-43.200001	-25.5	56	-0.006656	-0.392727	-0.014182
hsa-mir-376a-1	-18.200001	-16.4	57	-0.013382	-0.267647	-0.014695
hsa-mir-1285-2	-31.700001	-23.299999	74	-0.007831	-0.360227	-0.007659

cloop length	bulge ratio stem	bulge ratio mir	GAsurplusCU	Adjusted base pairing propensity (dP)	hpolyA	hpolyU
5	0.121212	0.045455	0.181818	0.385542	2	3
5	0.060606	0	0.181818	0.369048	2	3
-1	0.212121	0.045455	0.272727	4.714286	0	0
4	0.151515	0.045455	0.272727	0.35443	2	3
5	0.272727	0	0	0.390805	2	4
4	0.212121	0.045455	0.090909	0.392857	2	3
9	0.272727	0.045455	0	0.369048	2	3
8	0.030303	0	0	0.3125	2	2
5	0.060606	0.086957	0.304348	0.395062	4	5
6	0.121212	0.095238	0.238095	0.365854	3	2
16	0.030303	0.043478	-0.130435	0.309091	3	4
4	0.060606	0.043478	-0.043478	0.372727	2	3
4	0.030303	0	0.181818	0.376471	3	3
5	0.272727	0.190476	0.428571	0.342466	1	1

4	0.181818	0.083333	-0.25	0.337209	2	3
4	0.030303	0	-0.238095	0.352941	3	4
5	0.090909	0.043478	0.130435	0.404762	3	2
5	0.090909	0.045455	0.090909	0.371134	3	2
5	0.030303	0	-0.363636	0.347826	4	3
4	0	0	-0.363636	0.370787	4	2
4	0.30303	0.095238	-0.52381	0.353659	3	3
4	0.030303	0	-0.272727	0.316832	2	3
5	0.121212	0.045455	0.272727	0.356164	2	2
6	0.090909	0.043478	-0.304348	0.371134	4	4
9	0.030303	0.043478	-0.130435	0.378049	3	4
7	0.121212	0.086957	-0.217391	0.352941	1	2
6	0.121212	0.045455	-0.090909	0.4	2	4
7	0.151515	0.136364	-0.090909	0.368421	3	2
6	0.030303	0	-0.090909	0.330189	2	2
6	0.060606	0.045455	0.090909	0.395349	2	3
13	0.030303	0	-0.304348	0.340909	3	4
7	0.090909	0.045455	0.090909	0.373737	3	2
4	0.090909	0.090909	0.090909	0.383562	2	2
7	0.151515	0.181818	0	0.397059	3	3
5	0.060606	0.090909	-0.090909	0.313131	3	3
4	0.121212	0.130435	-0.565217	0.370787	1	2
4	0.090909	0.045455	-0.363636	0.392857	2	2
4	0.151515	0.086957	-0.130435	0.37931	2	2
9	0.090909	0.136364	-0.272727	0.369048	2	3
4	0.090909	0.086957	0.130435	0.430769	2	5
5	0	0	0.090909	0.313253	4	4
7	0.030303	0	0	0.336735	2	3
7	0.090909	0.086957	0.217391	0.369048	3	2
11	0.181818	0.136364	0	0.318182	3	3
4	0.181818	0.130435	-0.043478	0.328467	2	2
14	0.090909	0.125	0	0.318182	2	3
6	0.181818	0.181818	0	0.345455	2	2
6	0.060606	0.045455	0.363636	0.378049	3	3
14	0.060606	0.045455	-0.181818	0.383721	3	6
5	0.090909	0.043478	0.217391	0.309859	2	2
3	0.121212	0.086957	0.130435	0.309859	3	2
4	0.060606	0.043478	0.217391	0.369565	4	3
16	0.090909	0.095238	-0.047619	0.290909	2	2
9	0	0	0.363636	0.337349	3	4
14	0.121212	0.047619	0.238095	0.367816	3	2
10	0.060606	0	-0.090909	0.369048	1	3
4	0.030303	0	0.652174	0.386667	2	2
4	0.151515	0.043478	-0.304348	0.336364	3	3
7	0.151515	0.227273	-0.090909	0.402439	4	4
5	0.363636	0.227273	0	0.355556	2	3
7	0.151515	0.136364	0	0.336842	2	2
8	0.060606	0.045455	-0.181818	0.367647	2	3
15	0	0	-0.636364	0.036364	3	3
13	0.121212	0	-0.363636	0.327273	2	3
6	0.090909	0.086957	0.217391	0.408451	3	3
7	0.060606	0.086957	0.217391	0.362319	3	4
5	0.090909	0.090909	0	0.388889	2	2
3	0.090909	0.090909	-0.272727	0.345455	1	1
5	0.090909	0.045455	-0.727273	0.318182	3	3

3	0.166667	0.217391	-0.304348	0.072727	2	2
13	0	0	-0.363636	0.354545	2	2
11	0.030303	0.047619	0.142857	0.309091	2	3
4	0.121212	0.363636	0	0.345455	2	2
9	0.121212	0.090909	-0.090909	0.4	3	3
5	0.181818	0.045455	0	0.363636	2	3
11	0.030303	0.045455	-0.090909	0.345455	2	3
9	0.181818	0.181818	0	0.336364	3	3
5	0.090909	0.047619	-0.142857	0.358025	4	3
10	0.030303	0.045455	0.181818	0.369863	3	3
14	0.121212	0.181818	-0.181818	0.340206	4	3
7	0.151515	0.142857	0.333333	0.345238	2	2
6	0.060606	0.047619	0.142857	0.376623	2	2
8	0.030303	0.045455	0.090909	0.346154	2	2
3	0.121212	0.090909	0.181818	0.381443	2	3
11	0.166667	0.045455	-0.272727	0.359375	3	4
5	0.181818	0.045455	-0.363636	0.386364	3	5
3	0.212121	0.136364	-0.363636	0.360465	3	3
4	0.212121	0.090909	0.090909	0.408451	3	3
12	0.090909	0.090909	-0.181818	0.352273	3	3
4	0.060606	0	0	0.314607	3	3
6	0.060606	0.090909	0.090909	0.314286	3	3
4	0.121212	0.090909	0.090909	0.358696	3	3
8	0.030303	0	0.238095	0.366197	2	2
4	0.060606	0.045455	-0.090909	0.4	2	3
4	0.242424	0.043478	0.652174	0.333333	3	2
13	0.151515	0.181818	-0.272727	0.329787	3	2
6	0	0	0.391304	0.351064	5	5
4	0	0	-0.304348	0.319149	1	1
5	0.121212	0	-0.047619	0.376812	2	3
7	0.060606	0.045455	0.181818	0.347368	3	3
4	0.212121	0.190476	0.142857	0.333333	3	2
4	0.090909	0.136364	-0.181818	0.336735	2	1
4	0.090909	0.090909	-0.090909	0.363636	2	2
5	0.090909	0.086957	0.043478	0.357143	2	2
4	0.030303	0.043478	0.130435	0.38961	2	2
5	0.060606	0.045455	0.090909	0.306667	5	4
7	0.121212	0	-0.181818	0.32	2	2
11	0.121212	0.130435	-0.130435	0.343284	3	3
14	0.030303	0.045455	0	0.376812	4	5
4	0	0	0.090909	0.430556	2	2
13	0.166667	0.047619	-0.238095	0.348485	3	3
4	0.060606	0	0.333333	0.358209	2	3
4	0.030303	0.045455	-0.090909	0.386667	2	3
8	0.151515	0.045455	0.272727	0.356164	2	3
5	0	0.5	0	0.326531	4	4
7	0	0	-0.043478	0.287356	2	2
6	0.242424	0.045455	-0.181818	0.344086	2	2
11	0.060606	0.090909	0.272727	0.364706	3	3
10	0	0	-0.181818	0.897059	1	2
4	0.030303	0	-0.181818	0.393258	2	3
9	0.121212	0.130435	-0.391304	0.395062	3	3
9	0.090909	0.090909	-0.181818	0.390244	3	4
6	0.242424	0.285714	-0.047619	0.357143	3	3
9	0.151515	0.181818	-0.090909	0.357143	2	3

5	0.060606	0.043478	0.217391	0.408451	2	3
4	0.060606	0.090909	0.181818	0.351351	3	2
6	0.030303	0	-0.333333	0.367816	4	7
8	0.060606	0.08	-0.12	0.386364	3	4
5	0.060606	0	0	0.402299	4	4
8	0.060606	0.045455	-0.181818	0.371134	4	3
11	0	0	0.238095	0.435294	3	3
4	0	0	0.454545	0.371134	4	4
10	0.030303	0.043478	0.304348	0.369863	1	2
5	0	0	0.130435	0.320513	2	2
13	0.090909	0.090909	0.454545	0.375	2	6
6	0.121212	0.173913	0.130435	0.3875	3	2
6	0.030303	0.047619	-0.047619	0.382716	3	3
6	0.121212	0.043478	-0.304348	0.346154	2	5
11	0.272727	0.045455	0.090909	0.394958	2	3
8	0.030303	0	-0.090909	0.37037	2	2
5	0.272727	0.045455	0.181818	0.347222	2	3
5	0.121212	0.045455	0.181818	0.337349	2	4
5	0.030303	0	-0.181818	0.352273	3	3
4	0.030303	0	-0.181818	0.325843	2	4
4	0.060606	0.086957	0.130435	0.365854	2	3
8	0.090909	0.086957	0.304348	0.357143	2	3
4	0.060606	0.047619	-0.333333	0.347222	5	5
10	0.121212	0.095238	-0.333333	0.377778	5	5
5	0.060606	0.086957	0.130435	0.373737	2	2
6	0.121212	0.086957	0.130435	0.333333	2	3
10	0.181818	0.090909	0.181818	0.370787	4	2
8	0.090909	0.090909	0.181818	0.358025	3	3
11	0	0.173913	0.043478	0.072727	6	3
13	0.121212	0.173913	-0.130435	0.390805	4	4
4	0.060606	0.090909	-0.181818	0.40625	4	4
9	0.030303	0	-0.238095	0.354545	4	2
4	0.242424	0	-0.238095	0.318182	3	2
13	0.030303	0.045455	-0.090909	0.338235	2	4
9	0.181818	0.136364	0.090909	0.356164	3	3
7	0.090909	0	0.090909	0.38961	2	2
14	0	0	0.090909	0.333333	2	3
13	0.090909	0.125	0	0.382716	3	3
14	0.060606	0.086957	-0.217391	0.305556	3	3
5	0.090909	0.047619	-0.238095	0.4125	4	4
14	0.090909	0.090909	-0.090909	0.382716	3	5
4	0.121212	0.086957	0.043478	0.409091	3	5
12	0	0	0.043478	0.036364	3	4
6	0.060606	0.043478	0.130435	0.358974	2	3
3	0.060606	0.045455	0.090909	0.386667	3	3
3	0.272727	0	0.181818	0.378378	2	3
6	0.060606	0.043478	-0.217391	0.388889	3	3
12	0.181818	0	-0.090909	0.367089	3	5
7	0.060606	0.043478	-0.130435	0.37037	3	3
12	0.060606	0.045455	-0.090909	0.356322	2	3
8	0	0	0.363636	0.333333	4	3
-1	0.060606	0.043478	-0.304348	0.35	0	0
7	0.060606	0.045455	-0.272727	0.37931	4	5
7	0.090909	0.086957	0.043478	0.345455	4	3
-1	0.090909	0.086957	-0.043478	0.337079	0	0

3	0	0	0	0.381818	3	4
6	0.090909	0.043478	-0.304348	0.327273	2	2
5	0.090909	0.045455	-0.181818	0.338235	4	3
4	0.333333	0.190476	-0.238095	0.329545	3	3

hpolyG	hpolyC	mipolyA	mipolyU	mipolyG	mipolyC	pre-maxdibase-ratio	pre-minbase-occurance	pri-max-dibase-ratio
3	4	1	2	2	0	0.551282	0.205128	0.590361
3	3	1	2	2	0	0.571429	0.171429	0.607143
0	0	1	2	2	1	0	0	0
2	3	1	2	2	0	0.528571	0.2	0.56962
4	3	0	0	0	0	0.488372	0.151163	0.494253
3	3	1	3	2	1	0.551282	0.205128	0.5
3	3	1	3	2	1	0.653846	0.153846	0.642857
2	3	2	2	1	3	0.516129	0.193548	0.55
2	1	4	2	2	1	0.580645	0.177419	0.580247
3	2	3	1	1	1	0.564516	0.177419	0.609756
4	3	2	3	1	3	0.619048	0.142857	0.609091
4	3	2	3	1	3	0.564516	0.177419	0.6
2	2	2	3	2	1	0.596774	0.16129	0.494118
5	3	1	1	5	1	0.760563	0.070423	0.753425
3	3	2	3	1	3	0.548387	0.193548	0.593023
2	2	1	4	2	1	0.548387	0.193548	0.529412
3	3	2	1	3	2	0.516129	0.225806	0.535714
3	2	2	1	3	1	0.483871	0.209677	0.463918
4	3	1	2	3	3	0.528571	0.157143	0.576087
3	2	1	2	1	1	0.5	0.193548	0.573034
3	3	1	3	1	3	0.483871	0.225806	0.573171
3	2	1	3	2	2	0.548387	0.193548	0.653465
4	3	1	2	2	2	0.564516	0.193548	0.60274
3	3	1	4	2	2	0.532258	0.209677	0.546392
2	2	1	4	2	2	0.564516	0.145161	0.585366
2	3	1	1	1	2	0.596774	0.129032	0.602941
3	3	1	4	2	3	0.532258	0.209677	0.55
2	2	1	2	2	2	0.457143	0.171429	0.578947
3	1	1	1	2	1	0.618182	0.181818	0.575472
2	1	2	1	2	1	0.596774	0.129032	0.523256
4	3	2	4	2	3	0.516129	0.209677	0.556818
3	2	2	2	3	2	0.483871	0.177419	0.575758
2	3	2	2	2	2	0.419355	0.209677	0.561644
1	2	3	2	1	2	0.564516	0.193548	0.544118
2	1	2	2	2	1	0.548387	0.209677	0.565657
5	3	1	2	2	2	0.677419	0.096774	0.719101
5	3	2	2	1	3	0.580645	0.16129	0.666667
3	2	1	2	2	2	0.516129	0.209677	0.632184
2	2	1	2	2	2	0.564516	0.177419	0.595238
4	2	2	2	4	1	0.612903	0.16129	0.6
2	2	2	3	2	1	0.548387	0.16129	0.46988
2	1	1	3	2	1	0.580645	0.177419	0.489796
2	1	3	2	2	1	0.435484	0.16129	0.440476
3	2	2	2	2	2	0.516129	0.209677	0.545455
5	3	2	2	3	1	0.528571	0.185714	0.59854
2	2	2	3	2	1	0.442857	0.185714	0.572727
3	2	2	2	2	1	0.5	0.193548	0.554545

4	4	3	2	2	2	0.564516	0.129032	0.634146
3	3	3	4	3	2	0.564516	0.145161	0.581395
2	3	2	1	2	1	0.457143	0.185714	0.521127
2	4	2	2	2	1	0.485714	0.214286	0.521127
2	4	4	1	2	3	0.514286	0.3	0.597826
2	4	2	2	1	2	0.557143	0.314286	0.572727
4	3	1	4	4	1	0.564516	0.177419	0.626506
3	2	3	2	2	1	0.532258	0.209677	0.563218
3	2	1	3	2	2	0.580645	0.145161	0.607143
3	3	1	1	3	1	0.596774	0.177419	0.626667
2	3	1	3	1	3	0.564516	0.193548	0.545455
1	1	1	4	1	1	0.612903	0.129032	0.573171
2	3	1	2	2	2	0.548387	0.258065	0.555556
3	2	1	2	3	1	0.532258	0.177419	0.6
3	3	1	3	1	3	0.629032	0.177419	0.617647
3	3	1	3	1	3	0.548387	0.177419	0.554545
2	3	1	2	2	2	0.548387	0.177419	0.527273
2	1	3	2	2	1	0.612903	0.112903	0.619718
3	2	3	1	2	1	0.548387	0.193548	0.507246
3	2	1	2	1	1	0.532258	0.193548	0.569444
2	4	1	1	1	4	0.677419	0.129032	0.727273
4	3	1	3	1	3	0.564516	0.16129	0.590909
3	3	1	2	2	2	0.612903	0.16129	0.736364
2	2	1	2	1	2	0.555556	0.206349	0.563636
2	2	2	2	1	2	0.629032	0.16129	0.654545
3	3	2	1	2	1	0.548387	0.209677	0.527273
2	3	2	3	2	1	0.516129	0.225806	0.482353
3	2	2	1	2	2	0.587302	0.206349	0.554545
3	2	1	1	1	2	0.532258	0.193548	0.563636
2	4	2	3	1	1	0.457143	0.185714	0.490909
2	3	2	2	2	2	0.571429	0.171429	0.567901
4	2	1	2	4	2	0.370968	0.16129	0.575342
3	2	1	2	3	2	0.548387	0.177419	0.536082
3	3	2	2	3	1	0.580645	0.145161	0.630952
2	2	2	2	2	1	0.548387	0.16129	0.571429
3	2	1	2	3	1	0.564516	0.193548	0.602564
2	2	2	2	2	1	0.634921	0.174603	0.56701
2	2	1	4	2	1	0.645161	0.16129	0.671875
2	2	1	3	2	2	0.596774	0.177419	0.602273
1	1	1	3	1	1	0.580645	0.177419	0.593023
3	2	3	1	2	2	0.539683	0.222222	0.535211
3	2	3	1	1	2	0.483871	0.177419	0.568182
3	2	0	0	0	0	0.423077	0.205128	0.573034
2	4	3	1	2	4	0.532258	0.225806	0.542857
2	2	3	2	2	2	0.528571	0.157143	0.576087
3	2	2	1	2	1	0.5	0.209677	0.535211
2	2	2	2	1	1	0.634921	0.15873	0.614286
7	4	1	1	7	1	0.68254	0.301587	0.68
3	2	1	2	3	2	0.564516	0.193548	0.606383
1	2	5	1	1	1	0.5	0.129032	0.617021
2	3	1	1	2	3	0.629032	0.129032	0.734043
2	3	1	2	1	1	0.580645	0.193548	0.594203
3	2	3	2	1	1	0.571429	0.142857	0.610526
4	3	1	2	4	1	0.442857	0.171429	0.494949
4	3	1	1	3	2	0.612903	0.129032	0.602041

2	3	1	2	2	1	0.4	0.157143	0.6
2	2	1	2	2	1	0.532258	0.209677	0.535714
2	2	1	2	2	1	0.545455	0.218182	0.597403
3	2	2	1	3	1	0.542857	0.142857	0.453333
4	2	1	2	2	1	0.548387	0.16129	0.573333
2	2	3	2	2	2	0.435484	0.193548	0.552239
5	2	4	3	5	1	0.677419	0.145161	0.666667
1	2	2	2	1	2	0.722222	0.111111	0.680556
2	2	1	2	2	2	0.645161	0.112903	0.712121
2	2	2	1	2	1	0.564516	0.177419	0.58209
3	3	1	3	2	3	0.587302	0.174603	0.6
3	2	2	2	2	1	0.483871	0.177419	0.547945
2	1	2	4	1	1	0.548387	0.16129	0.591837
3	3	2	2	1	3	0.603175	0.174603	0.574713
3	2	1	2	2	2	0.576923	0.179487	0.591398
2	1	3	3	2	1	0.571429	0.157143	0.494118
4	4	1	1	1	4	0.634921	0.15873	0.632353
2	2	1	3	2	1	0.634921	0.174603	0.595506
3	2	1	2	2	2	0.6	0.2	0.617284
2	3	2	4	1	3	0.580645	0.177419	0.621951
2	2	1	3	2	1	0.548387	0.274194	0.669643
3	3	2	3	3	3	0.532258	0.193548	0.428571
4	2	2	2	3	1	0.619048	0.15873	0.605634
2	2	2	1	2	1	0.564516	0.129032	0.581081
1	1	1	4	1	1	0.612903	0.112903	0.609195
3	3	3	3	3	2	0.580645	0.193548	0.590909
3	2	2	3	3	1	0.564516	0.145161	0.574713
3	3	4	2	1	3	0.516129	0.193548	0.597938
5	5	3	2	5	1	0.483871	0.193548	0.505882
2	2	4	1	2	1	0.580645	0.177419	0.649485
3	3	1	1	3	3	0.636364	0.163636	0.684932
3	4	2	1	3	4	0.66129	0.112903	0.692308
3	5	1	1	3	1	0.619048	0.126984	0.75
2	3	3	2	2	1	0.532258	0.209677	0.6
3	1	3	2	1	1	0.612903	0.112903	0.493827
2	2	1	5	2	1	0.557143	0.171429	0.564103
4	4	2	2	2	0	0.569767	0.139535	0.588235
3	3	2	2	1	3	0.516129	0.193548	0.555556
3	2	1	2	2	0	0.571429	0.157143	0.597222
3	4	1	2	2	0	0.564103	0.166667	0.518072
3	3	2	2	1	3	0.564516	0.193548	0.579545
3	3	2	2	1	3	0.516129	0.209677	0.483146
4	2	1	1	4	2	0.542373	0.169492	0.609756
5	2	1	1	5	2	0.516129	0.209677	0.571429
3	4	0	5	3	1	0.542857	0.257143	0.569444
3	4	0	5	3	1	0.5	0.185714	0.488889
3	2	2	2	2	2	0.548387	0.193548	0.585859
3	3	2	2	2	2	0.557143	0.2	0.571429
2	2	3	2	2	1	0.585714	0.157143	0.595506
2	2	3	2	2	1	0.628571	0.142857	0.62963
3	2	2	2	2	1	0.532258	0.209677	0.536364
2	2	1	4	2	2	0.564516	0.193548	0.609195
2	2	1	4	2	1	0.585714	0.142857	0.65625
2	2	2	2	1	2	0.528571	0.171429	0.490909
2	2	2	2	1	2	0.457143	0.185714	0.554545

2	2	1	1	1	2	0.516129	0.193548	0.529412
2	2	3	1	1	2	0.467742	0.209677	0.547945
3	3	2	2	2	2	0.548387	0.209677	0.571429
2	2	2	2	2	2	0.580645	0.193548	0.630952
2	2	1	3	2	1	0.6	0.1	0.641975
3	2	3	1	1	2	0.564516	0.193548	0.583333
2	2	1	4	2	2	0.596774	0.129032	0.6625
2	2	1	3	2	1	0.571429	0.171429	0.617284
2	2	2	4	2	1	0.614286	0.142857	0.581818
2	3	2	4	2	1	0.596774	0.177419	0.527273
3	3	2	2	3	1	0.66129	0.112903	0.679487
4	3	1	3	4	1	0.66129	0.16129	0.626667
4	3	1	2	2	0	0.614286	0.171429	0.635135
2	2	1	3	2	1	0.564516	0.193548	0.488889
2	2	1	2	1	1	0.6	0.171429	0.493671
2	2	3	1	2	2	0.548387	0.193548	0.617284
2	2	1	2	2	2	0.5	0.209677	0.54023
2	4	4	1	2	2	0.516129	0.209677	0.509804
0	0	1	5	2	2	0	0	0.47
1	3	1	5	1	1	0.596774	0.16129	0.574713
2	2	2	2	2	1	0.596774	0.193548	0.581818
0	0	2	2	3	1	0	0	0.539326
3	2	1	3	2	1	0.580645	0.129032	0.618182
2	3	1	2	1	3	0.532258	0.241935	0.536364
2	2	1	2	1	2	0.645161	0.145161	0.705882
3	3	1	3	1	3	0.525641	0.192308	0.522727

pri-min-base occurance	preAcnt	preUcnt	preGcnt	preCcnt	preGCcnt	preG/C	priAcnt	priUcnt
0.192771	0.205128	0.282051	0.269231	0.217949	0.487179	1.235294	0.192771	0.277108
0.190476	0.214286	0.328571	0.242857	0.171429	0.414286	1.416667	0.202381	0.321429
0	0	0	0	0	0	0	0	0
0.202532	0.242857	0.228571	0.285714	0.2	0.485714	1.428571	0.227848	0.202532
0.149425	0.244186	0.337209	0.244186	0.151163	0.395349	1.615385	0.252874	0.344828
0.214286	0.24359	0.269231	0.282051	0.205128	0.487179	1.375	0.25	0.25
0.154762	0.153846	0.320513	0.333333	0.192308	0.525641	1.733333	0.154762	0.309524
0.225	0.241935	0.274194	0.209677	0.193548	0.403226	1.083333	0.225	0.3125
0.209877	0.274194	0.306452	0.193548	0.177419	0.370968	1.090909	0.259259	0.320988
0.158537	0.177419	0.274194	0.290323	0.258065	0.548387	1.125	0.158537	0.231707
0.181818	0.285714	0.333333	0.238095	0.142857	0.380952	1.666667	0.254545	0.354545
0.172727	0.258065	0.306452	0.241935	0.177419	0.419355	1.363636	0.290909	0.309091
0.211765	0.306452	0.290323	0.177419	0.16129	0.33871	1.1	0.282353	0.282353
0.082192	0.070423	0.169014	0.450704	0.309859	0.760563	1.454545	0.082192	0.164384
0.162791	0.193548	0.225806	0.290323	0.258065	0.548387	1.125	0.162791	0.244186
0.223529	0.241935	0.306452	0.193548	0.209677	0.403226	0.923077	0.223529	0.247059
0.22619	0.209677	0.209677	0.290323	0.225806	0.516129	1.285714	0.238095	0.22619
0.206186	0.209677	0.225806	0.258065	0.209677	0.467742	1.230769	0.206186	0.268041
0.173913	0.157143	0.242857	0.257143	0.271429	0.528571	0.947368	0.173913	0.25
0.191011	0.225806	0.274194	0.193548	0.209677	0.403226	0.923077	0.191011	0.303371
0.207317	0.241935	0.241935	0.258065	0.225806	0.483871	1.142857	0.207317	0.219512
0.168317	0.193548	0.258065	0.290323	0.193548	0.483871	1.5	0.168317	0.178218
0.178082	0.193548	0.225806	0.290323	0.274194	0.564516	1.058824	0.178082	0.219178
0.195876	0.225806	0.322581	0.225806	0.209677	0.435484	1.076923	0.195876	0.278351

0.195122	0.193548	0.370968	0.145161	0.193548	0.33871	0.75	0.195122	0.353659
0.161765	0.129032	0.225806	0.354839	0.241935	0.596774	1.466667	0.161765	0.25
0.18	0.209677	0.274194	0.258065	0.225806	0.483871	1.142857	0.18	0.28
0.168421	0.228571	0.285714	0.228571	0.171429	0.4	1.333333	0.168421	0.273684
0.198113	0.181818	0.272727	0.345455	0.2	0.545455	1.727273	0.198113	0.226415
0.232558	0.306452	0.290323	0.193548	0.129032	0.322581	1.5	0.244186	0.255814
0.204545	0.225806	0.306452	0.225806	0.209677	0.435484	1.076923	0.204545	0.318182
0.20202	0.225806	0.306452	0.225806	0.177419	0.403226	1.272727	0.20202	0.343434
0.191781	0.209677	0.258065	0.258065	0.209677	0.467742	1.230769	0.191781	0.246575
0.220588	0.306452	0.258065	0.209677	0.193548	0.403226	1.083333	0.294118	0.25
0.212121	0.241935	0.306452	0.209677	0.209677	0.419355	1	0.262626	0.30303
0.101124	0.096774	0.225806	0.419355	0.258065	0.677419	1.625	0.101124	0.179775
0.154762	0.16129	0.177419	0.274194	0.306452	0.580645	0.894737	0.154762	0.178571
0.333333	0.209677	0.225806	0.290323	0.225806	0.516129	1.285714	0.183908	0.183908
0.190476	0.177419	0.370968	0.193548	0.193548	0.387097	1	0.22619	0.369048
0.184615	0.274194	0.33871	0.16129	0.209677	0.370968	0.769231	0.261538	0.338462
0.156627	0.241935	0.290323	0.258065	0.16129	0.419355	1.6	0.313253	0.26506
0.173469	0.274194	0.306452	0.177419	0.209677	0.387097	0.846154	0.316327	0.316327
0.154762	0.258065	0.258065	0.274194	0.16129	0.435484	1.7	0.285714	0.285714
0.236364	0.241935	0.209677	0.274194	0.241935	0.516129	1.133333	0.227273	0.227273
0.189781	0.2	0.214286	0.314286	0.185714	0.5	1.692308	0.211679	0.189781
0.2	0.242857	0.242857	0.257143	0.185714	0.442857	1.384615	0.2	0.227273
0.218182	0.306452	0.241935	0.241935	0.193548	0.435484	1.25	0.290909	0.218182
0.146341	0.129032	0.225806	0.290323	0.274194	0.564516	1.058824	0.146341	0.219512
0.151163	0.209677	0.419355	0.209677	0.145161	0.354839	1.444444	0.209302	0.430233
0.211268	0.228571	0.271429	0.228571	0.185714	0.414286	1.230769	0.239437	0.309859
0.225352	0.228571	0.257143	0.214286	0.228571	0.442857	0.9375	0.239437	0.295775
0.173913	0.2	0.214286	0.2	0.3	0.5	0.666667	0.173913	0.228261
0.2	0.185714	0.242857	0.185714	0.314286	0.5	0.590909	0.2	0.227273
0.180723	0.177419	0.290323	0.274194	0.209677	0.483871	1.307692	0.180723	0.325301
0.241379	0.258065	0.209677	0.274194	0.225806	0.5	1.214286	0.218391	0.218391
0.166667	0.145161	0.306452	0.193548	0.274194	0.467742	0.705882	0.166667	0.357143
0.32	0.209677	0.177419	0.274194	0.322581	0.596774	0.85	0.186667	0.186667
0.209091	0.241935	0.322581	0.193548	0.225806	0.419355	0.857143	0.209091	0.263636
0.182927	0.33871	0.274194	0.193548	0.129032	0.322581	1.5	0.268293	0.304878
0.2	0.209677	0.290323	0.209677	0.258065	0.467742	0.8125	0.2	0.244444
0.189474	0.225806	0.274194	0.258065	0.177419	0.435484	1.454545	0.189474	0.210526
0.161765	0.177419	0.290323	0.33871	0.193548	0.532258	1.75	0.161765	0.279412
0.218182	0.258065	0.241935	0.290323	0.177419	0.467742	1.636364	0.227273	0.290909
0.218182	0.209677	0.274194	0.177419	0.274194	0.451613	0.647059	0.254545	0.263636
0.15493	0.274194	0.33871	0.209677	0.112903	0.322581	1.857143	0.28169	0.338028
0.202899	0.225806	0.290323	0.258065	0.193548	0.451613	1.333333	0.246377	0.304348
0.208333	0.225806	0.290323	0.241935	0.193548	0.435484	1.25	0.222222	0.291667
0.127273	0.16129	0.129032	0.258065	0.419355	0.677419	0.615385	0.145455	0.127273
0.172727	0.16129	0.209677	0.306452	0.258065	0.564516	1.1875	0.172727	0.236364
0.127273	0.16129	0.209677	0.225806	0.387097	0.612903	0.583333	0.136364	0.127273
0.209091	0.238095	0.206349	0.238095	0.31746	0.555556	0.75	0.227273	0.209091
0.163636	0.290323	0.33871	0.177419	0.16129	0.33871	1.1	0.318182	0.336364
0.236364	0.241935	0.306452	0.209677	0.225806	0.435484	0.928571	0.245455	0.281818
0.2	0.241935	0.274194	0.225806	0.225806	0.451613	1	0.2	0.258824
0.218182	0.206349	0.349206	0.238095	0.206349	0.444444	1.153846	0.227273	0.3
0.190909	0.193548	0.290323	0.225806	0.241935	0.467742	0.933333	0.190909	0.3
0.209091	0.228571	0.271429	0.228571	0.185714	0.414286	1.230769	0.209091	0.254545
0.197531	0.242857	0.328571	0.228571	0.171429	0.4	1.333333	0.234568	0.320988
0.178082	0.16129	0.274194	0.274194	0.209677	0.483871	1.307692	0.178082	0.246575

0.216495	0.241935	0.306452	0.225806	0.177419	0.403226	1.272727	0.216495	0.278351
0.142857	0.145161	0.241935	0.33871	0.241935	0.580645	1.4	0.142857	0.22619
0.168831	0.16129	0.33871	0.209677	0.209677	0.419355	1	0.168831	0.272727
0.192308	0.193548	0.241935	0.306452	0.258065	0.564516	1.1875	0.192308	0.205128
0.195876	0.190476	0.349206	0.285714	0.174603	0.460317	1.636364	0.237113	0.28866
0.15625	0.258065	0.387097	0.16129	0.16129	0.322581	1	0.265625	0.40625
0.193182	0.177419	0.387097	0.209677	0.193548	0.403226	1.083333	0.204545	0.352273
0.197674	0.258065	0.322581	0.177419	0.209677	0.387097	0.846154	0.302326	0.290698
0.225352	0.238095	0.253968	0.285714	0.222222	0.507937	1.285714	0.225352	0.239437
0.193182	0.241935	0.306452	0.241935	0.177419	0.419355	1.363636	0.261364	0.306818
0.202247	0.205128	0.282051	0.282051	0.217949	0.5	1.294118	0.202247	0.292135
0.228571	0.274194	0.258065	0.225806	0.241935	0.467742	0.933333	0.257143	0.285714
0.206522	0.228571	0.271429	0.257143	0.157143	0.414286	1.636364	0.217391	0.271739
0.225352	0.258065	0.225806	0.241935	0.209677	0.451613	1.153846	0.253521	0.239437
0.157143	0.238095	0.396825	0.206349	0.15873	0.365079	1.3	0.242857	0.371429
0.146667	0.15873	0.15873	0.380952	0.301587	0.68254	1.263158	0.146667	0.173333
0.191489	0.241935	0.193548	0.274194	0.290323	0.564516	0.944444	0.191489	0.202128
0.180851	0.306452	0.306452	0.129032	0.193548	0.322581	0.666667	0.255319	0.361702
0.095745	0.129032	0.177419	0.274194	0.354839	0.629032	0.772727	0.095745	0.170213
0.188406	0.193548	0.33871	0.241935	0.225806	0.467742	1.071429	0.188406	0.333333
0.168421	0.228571	0.342857	0.2	0.142857	0.342857	1.4	0.231579	0.378947
0.212121	0.242857	0.242857	0.271429	0.171429	0.442857	1.583333	0.252525	0.252525
0.183673	0.129032	0.193548	0.387097	0.225806	0.612903	1.714286	0.183673	0.214286
0.190909	0.242857	0.257143	0.257143	0.157143	0.414286	1.636364	0.209091	0.281818
0.22619	0.225806	0.306452	0.209677	0.209677	0.419355	1	0.25	0.285714
0.194805	0.309091	0.218182	0.236364	0.236364	0.472727	1	0.337662	0.25974
0.16	0.285714	0.257143	0.242857	0.142857	0.385714	1.7	0.293333	0.293333
0.266667	0.16129	0.209677	0.290323	0.258065	0.548387	1.125	0.213333	0.213333
0.208955	0.241935	0.258065	0.258065	0.193548	0.451613	1.333333	0.238806	0.253731
0.144928	0.145161	0.370968	0.306452	0.16129	0.467742	1.9	0.144928	0.347826
0.152778	0.333333	0.388889	0.12963	0.111111	0.240741	1.166667	0.305556	0.375
0.106061	0.258065	0.387097	0.193548	0.112903	0.306452	1.714286	0.30303	0.409091
0.179104	0.258065	0.306452	0.225806	0.177419	0.403226	1.272727	0.283582	0.298507
0.16	0.238095	0.333333	0.253968	0.174603	0.428571	1.454545	0.266667	0.333333
0.205479	0.258065	0.258065	0.306452	0.177419	0.483871	1.727273	0.260274	0.246575
0.173469	0.258065	0.290323	0.225806	0.16129	0.387097	1.4	0.255102	0.234694
0.206897	0.222222	0.174603	0.285714	0.31746	0.603175	0.9	0.206897	0.218391
0.204301	0.192308	0.282051	0.294872	0.179487	0.474359	1.642857	0.204301	0.258065
0.2	0.3	0.271429	0.214286	0.157143	0.371429	1.363636	0.294118	0.294118
0.176471	0.15873	0.206349	0.285714	0.349206	0.634921	0.818182	0.176471	0.191176
0.146067	0.174603	0.365079	0.269841	0.190476	0.460317	1.416667	0.146067	0.314607
0.17284	0.236364	0.363636	0.2	0.2	0.4	1	0.246914	0.37037
0.182927	0.241935	0.33871	0.177419	0.177419	0.354839	1	0.268293	0.353659
0.160714	0.225806	0.225806	0.274194	0.274194	0.548387	1	0.169643	0.160714
0.178571	0.193548	0.274194	0.258065	0.225806	0.483871	1.142857	0.178571	0.285714
0.15493	0.15873	0.253968	0.365079	0.222222	0.587302	1.642857	0.15493	0.253521
0.162162	0.306452	0.241935	0.258065	0.129032	0.387097	2	0.297297	0.256757
0.172414	0.258065	0.354839	0.209677	0.112903	0.322581	1.857143	0.275862	0.333333
0.193182	0.209677	0.33871	0.193548	0.241935	0.435484	0.8	0.193182	0.340909
0.183908	0.274194	0.290323	0.258065	0.145161	0.403226	1.777778	0.241379	0.310345
0.185567	0.322581	0.241935	0.241935	0.193548	0.435484	1.25	0.319588	0.278351
0.235294	0.274194	0.274194	0.193548	0.209677	0.403226	0.923077	0.247059	0.270588
0.175258	0.193548	0.354839	0.225806	0.177419	0.403226	1.272727	0.298969	0.350515
0.136986	0.163636	0.2	0.345455	0.290909	0.636364	1.1875	0.136986	0.178082
0.371795	0.177419	0.112903	0.290323	0.370968	0.66129	0.782609	0.153846	0.153846

0.083333	0.126984	0.253968	0.285714	0.333333	0.619048	0.857143	0.083333	0.166667
0.175	0.209677	0.258065	0.241935	0.274194	0.516129	0.882353	0.175	0.225
0.197531	0.290323	0.322581	0.177419	0.112903	0.290323	1.571429	0.296296	0.296296
0.205128	0.171429	0.328571	0.2	0.228571	0.428571	0.875	0.205128	0.320513
0.159664	0.267442	0.302326	0.22093	0.139535	0.360465	1.583333	0.252101	0.319328
0.197531	0.193548	0.241935	0.274194	0.241935	0.516129	1.133333	0.197531	0.271605
0.152778	0.271429	0.3	0.228571	0.157143	0.385714	1.454545	0.277778	0.319444
0.180723	0.230769	0.333333	0.192308	0.166667	0.358974	1.153846	0.240964	0.337349
0.159091	0.193548	0.322581	0.241935	0.241935	0.483871	1	0.159091	0.306818
0.235955	0.241935	0.274194	0.209677	0.225806	0.435484	0.928571	0.247191	0.258427
0.158537	0.169492	0.288136	0.254237	0.237288	0.491525	1.071429	0.158537	0.341463
0.202381	0.209677	0.209677	0.306452	0.209677	0.516129	1.461538	0.202381	0.238095
0.208333	0.2	0.285714	0.2	0.257143	0.457143	0.777778	0.208333	0.305556
0.211111	0.228571	0.242857	0.185714	0.257143	0.442857	0.722222	0.211111	0.255556
0.181818	0.258065	0.193548	0.274194	0.274194	0.548387	1	0.232323	0.181818
0.190476	0.2	0.228571	0.271429	0.285714	0.557143	0.95	0.190476	0.238095
0.179775	0.285714	0.3	0.171429	0.157143	0.328571	1.090909	0.292135	0.303371
0.185185	0.328571	0.3	0.157143	0.142857	0.3	1.1	0.308642	0.320988
0.227273	0.274194	0.258065	0.225806	0.209677	0.435484	1.076923	0.263636	0.236364
0.183908	0.241935	0.322581	0.225806	0.193548	0.419355	1.166667	0.206897	0.356322
0.145833	0.257143	0.328571	0.214286	0.142857	0.357143	1.5	0.28125	0.375
0.190909	0.228571	0.271429	0.257143	0.171429	0.428571	1.5	0.254545	0.3
0.209091	0.228571	0.228571	0.271429	0.185714	0.457143	1.461538	0.209091	0.236364
0.264706	0.241935	0.274194	0.193548	0.241935	0.435484	0.8	0.235294	0.264706
0.205479	0.241935	0.241935	0.258065	0.209677	0.467742	1.230769	0.205479	0.246575
0.181818	0.209677	0.290323	0.258065	0.225806	0.483871	1.142857	0.181818	0.25974
0.178571	0.193548	0.387097	0.193548	0.193548	0.387097	1	0.178571	0.345238
0.123457	0.228571	0.371429	0.214286	0.1	0.314286	2.142857	0.234568	0.37037
0.194444	0.306452	0.258065	0.241935	0.193548	0.435484	1.25	0.305556	0.277778
0.125	0.241935	0.354839	0.225806	0.129032	0.354839	1.75	0.275	0.3875
0.17284	0.185714	0.385714	0.171429	0.171429	0.342857	1	0.209877	0.395062
0.190909	0.314286	0.3	0.157143	0.142857	0.3	1.1	0.263636	0.318182
0.272727	0.274194	0.322581	0.177419	0.225806	0.403226	0.785714	0.236364	0.254545
0.128205	0.112903	0.354839	0.306452	0.177419	0.483871	1.727273	0.128205	0.384615
0.186667	0.16129	0.387097	0.274194	0.16129	0.435484	1.7	0.186667	0.36
0.175676	0.2	0.357143	0.257143	0.171429	0.428571	1.5	0.189189	0.351351
0.211111	0.290323	0.274194	0.225806	0.193548	0.419355	1.166667	0.277778	0.277778
0.151899	0.257143	0.342857	0.214286	0.171429	0.385714	1.25	0.253165	0.341772
0.17284	0.193548	0.290323	0.258065	0.225806	0.483871	1.142857	0.17284	0.320988
0.264368	0.241935	0.258065	0.225806	0.209677	0.435484	1.076923	0.229885	0.229885
0.235294	0.274194	0.225806	0.209677	0.241935	0.451613	0.866667	0.245098	0.245098
0.13	0	0	0	0	0	0	0.34	0.34
0.16092	0.274194	0.322581	0.209677	0.16129	0.370968	1.3	0.264368	0.298851
0.2	0.322581	0.274194	0.193548	0.209677	0.403226	0.923077	0.272727	0.309091
0.224719	0	0	0	0	0	0	0.280899	0.258427
0.190909	0.241935	0.306452	0.274194	0.129032	0.403226	2.125	0.190909	0.327273
0.227273	0.225806	0.290323	0.225806	0.241935	0.467742	0.933333	0.227273	0.236364
0.147059	0.306452	0.33871	0.145161	0.145161	0.290323	1	0.338235	0.367647
0.204545	0.24359	0.192308	0.282051	0.217949	0.5	1.294118	0.272727	0.204545

priGcnt	priCcnt	priGCcnt	pri G/C	wobber vs all matches(stem)	miAcnt	miUcnt	miGcnt	miCcnt
0.313253	0.216867	0.53012	1.444444	0.172414	0.136364	0.409091	0.454545	0

0.285714	0.190476	0.47619	1.5	0.208333	0.181818	0.409091	0.409091	0
0	0	0	0	0.227273	0.272727	0.318182	0.363636	0.045455
0.316456	0.253165	0.56962	1.25	0.130435	0.227273	0.363636	0.409091	0
0.252874	0.149425	0.402299	1.692308	0.166667	0	0	0	0
0.285714	0.214286	0.5	1.333333	0.125	0.227273	0.409091	0.318182	0.045455
0.333333	0.202381	0.535714	1.647059	0.217391	0.136364	0.454545	0.363636	0.045455
0.2375	0.225	0.4625	1.055556	0.217391	0.272727	0.227273	0.227273	0.272727
0.209877	0.209877	0.419753	1	0.037037	0.347826	0.217391	0.304348	0.130435
0.317073	0.292683	0.609756	1.083333	0.166667	0.333333	0.238095	0.285714	0.142857
0.209091	0.181818	0.390909	1.15	0.107143	0.217391	0.347826	0.217391	0.217391
0.227273	0.172727	0.4	1.315789	0.111111	0.26087	0.304348	0.217391	0.217391
0.223529	0.211765	0.435294	1.055556	0.068966	0.181818	0.363636	0.409091	0.045455
0.452055	0.30137	0.753425	1.5	0.086957	0.047619	0.190476	0.666667	0.095238
0.313953	0.27907	0.593023	1.125	0.181818	0.208333	0.291667	0.166667	0.333333
0.258824	0.270588	0.529412	0.956522	0.08	0.190476	0.428571	0.190476	0.190476
0.297619	0.238095	0.535714	1.25	0.066667	0.217391	0.173913	0.347826	0.26087
0.268041	0.257732	0.525773	1.04	0.115385	0.227273	0.227273	0.318182	0.227273
0.293478	0.282609	0.576087	1.038462	0.28	0.090909	0.272727	0.227273	0.409091
0.269663	0.235955	0.505618	1.142857	0.24	0.136364	0.409091	0.181818	0.272727
0.292683	0.280488	0.573171	1.043478	0.157895	0.142857	0.380952	0.095238	0.380952
0.287129	0.366337	0.653465	0.783784	0.12	0.136364	0.409091	0.227273	0.227273
0.30137	0.30137	0.60274	1	0.125	0.181818	0.227273	0.454545	0.136364
0.268041	0.257732	0.525773	1.04	0.178571	0.173913	0.478261	0.173913	0.173913
0.231707	0.219512	0.451219	1.055556	0.071429	0.217391	0.434783	0.217391	0.130435
0.352941	0.235294	0.588235	1.5	0.166667	0.173913	0.304348	0.217391	0.304348
0.27	0.27	0.54	1	0.107143	0.181818	0.363636	0.272727	0.181818
0.305263	0.252632	0.557895	1.208333	0	0.227273	0.318182	0.227273	0.227273
0.330189	0.245283	0.575472	1.346154	0.08	0.090909	0.318182	0.363636	0.227273

0.267442	0.232558	0.5	1.15	0.142857	0.363636	0.318182	0.181818	0.136364
0.238636	0.238636	0.477273	1	0.074074	0.173913	0.304348	0.173913	0.347826
0.232323	0.222222	0.454545	1.045455	0.111111	0.272727	0.318182	0.272727	0.136364
0.287671	0.273973	0.561644	1.05	0.153846	0.318182	0.272727	0.227273	0.181818
0.235294	0.220588	0.455882	1.066667	0.076923	0.318182	0.227273	0.181818	0.272727
0.222222	0.212121	0.434343	1.047619	0.192308	0.272727	0.318182	0.181818	0.227273
0.438202	0.280899	0.719101	1.56	0.25	0.043478	0.347826	0.173913	0.434783
0.309524	0.357143	0.666667	0.866667	0.038462	0.181818	0.272727	0.136364	0.409091
0.298851	0.333333	0.632184	0.896552	0.038462	0.217391	0.304348	0.217391	0.26087
0.214286	0.190476	0.404762	1.125	0.16	0.090909	0.409091	0.272727	0.227273
0.184615	0.215385	0.4	0.857143	0.111111	0.26087	0.347826	0.304348	0.086957
0.26506	0.156627	0.421687	1.692308	0.190476	0.272727	0.318182	0.272727	0.136364
0.193878	0.173469	0.367347	1.117647	0.125	0.318182	0.272727	0.181818	0.227273
0.27381	0.154762	0.428571	1.769231	0.076923	0.304348	0.217391	0.304348	0.173913
0.309091	0.236364	0.545455	1.307692	0.125	0.272727	0.272727	0.227273	0.227273
0.335766	0.262774	0.59854	1.277778	0.136364	0.173913	0.391304	0.304348	0.130435
0.3	0.272727	0.572727	1.1	0.08	0.291667	0.291667	0.208333	0.208333
0.263636	0.227273	0.490909	1.16	0.041667	0.272727	0.318182	0.227273	0.181818
0.329268	0.304878	0.634146	1.08	0.230769	0.318182	0.181818	0.363636	0.136364
0.209302	0.151163	0.360465	1.384615	0.133333	0.227273	0.363636	0.181818	0.227273
0.239437	0.211268	0.450704	1.133333	0.181818	0.304348	0.26087	0.304348	0.130435
0.239437	0.225352	0.464789	1.0625	0.136364	0.26087	0.304348	0.304348	0.130435
0.26087	0.336957	0.597826	0.774194	0.12	0.391304	0.086957	0.217391	0.304348
0.245455	0.327273	0.572727	0.75	0.074074	0.285714	0.238095	0.190476	0.285714
0.301205	0.192771	0.493976	1.5625	0.24	0.181818	0.227273	0.5	0.090909
0.321839	0.241379	0.563218	1.333333	0.04	0.380952	0.190476	0.238095	0.190476
0.22619	0.25	0.47619	0.904762	0.259259	0.090909	0.454545	0.363636	0.090909
0.306667	0.32	0.626667	0.958333	0.192308	0.217391	0.086957	0.608696	0.086957

0.245455	0.281818	0.527273	0.870968	0.08	0.173913	0.391304	0.173913	0.26087
0.243902	0.182927	0.426829	1.333333	0.111111	0.272727	0.363636	0.181818	0.181818
0.255556	0.3	0.555556	0.851852	0.047619	0.227273	0.227273	0.272727	0.272727
0.326316	0.273684	0.6	1.192308	0.16	0.227273	0.272727	0.272727	0.227273
0.338235	0.220588	0.558824	1.533333	0.208333	0.136364	0.318182	0.272727	0.272727
0.263636	0.218182	0.481818	1.208333	0	0.090909	0.454545	0.090909	0.363636
0.218182	0.263636	0.481818	0.827586	0.192308	0.136364	0.318182	0.181818	0.363636
0.225352	0.15493	0.380282	1.454545	0.142857	0.304348	0.304348	0.304348	0.086957
0.246377	0.202899	0.449275	1.214286	0.125	0.304348	0.217391	0.304348	0.173913
0.277778	0.208333	0.486111	1.333333	0.259259	0.272727	0.363636	0.227273	0.136364
0.309091	0.418182	0.727273	0.73913	0.043478	0.181818	0.090909	0.181818	0.545455
0.3	0.290909	0.590909	1.03125	0.2	0.045455	0.454545	0.090909	0.409091
0.281818	0.454545	0.736364	0.62	0	0.173913	0.347826	0.173913	0.304348
0.254545	0.309091	0.563636	0.823529	0.068966	0.090909	0.363636	0.227273	0.318182
0.181818	0.163636	0.345455	1.111111	0.038462	0.380952	0.238095	0.190476	0.190476
0.236364	0.236364	0.472727	1	0.08	0.272727	0.272727	0.227273	0.227273
0.258824	0.282353	0.541176	0.916667	0.142857	0.227273	0.363636	0.227273	0.181818
0.254545	0.218182	0.472727	1.166667	0.166667	0.318182	0.318182	0.181818	0.181818
0.263636	0.245455	0.509091	1.074074	0.076923	0.227273	0.272727	0.227273	0.272727
0.254545	0.281818	0.536364	0.903226	0.043478	0.227273	0.363636	0.272727	0.136364
0.246914	0.197531	0.444444	1.25	0.130435	0.190476	0.333333	0.238095	0.238095
0.30137	0.273973	0.575342	1.1	0.153846	0.090909	0.318182	0.5	0.090909
0.247423	0.257732	0.505155	0.96	0.074074	0.090909	0.409091	0.318182	0.181818
0.369048	0.261905	0.630952	1.409091	0.166667	0.238095	0.190476	0.428571	0.142857
0.298701	0.25974	0.558442	1.15	0.148148	0.333333	0.333333	0.238095	0.095238
0.307692	0.294872	0.602564	1.043478	0.230769	0.181818	0.272727	0.363636	0.181818
0.278351	0.195876	0.474227	1.421053	0.2	0.272727	0.272727	0.318182	0.136364
0.171875	0.15625	0.328125	1.1	0	0.136364	0.5	0.227273	0.136364

0.25	0.193182	0.443182	1.294118	0.166667	0.090909	0.409091	0.227273	0.272727
0.197674	0.209302	0.406977	0.944444	0.043478	0.181818	0.409091	0.136364	0.272727
0.295775	0.239437	0.535211	1.235294	0.076923	0.318182	0.227273	0.227273	0.227273
0.238636	0.193182	0.431818	1.235294	0.153846	0.318182	0.272727	0.090909	0.318182
0.280899	0.224719	0.505618	1.25	0.318182	0	0	0	0
0.228571	0.228571	0.457143	1	0.136364	0.318182	0.181818	0.227273	0.272727
0.304348	0.206522	0.51087	1.473684	0.074074	0.318182	0.272727	0.227273	0.181818
0.28169	0.225352	0.507042	1.25	0.08	0.285714	0.190476	0.333333	0.190476
0.228571	0.157143	0.385714	1.454545	0.259259	0.318182	0.363636	0.136364	0.181818
0.373333	0.306667	0.68	1.217391	0.217391	0.130435	0.043478	0.695652	0.130435
0.276596	0.329787	0.606383	0.83871	0	0.045455	0.363636	0.318182	0.272727
0.180851	0.202128	0.382979	0.894737	0.068966	0.521739	0.173913	0.173913	0.130435
0.329787	0.404255	0.734043	0.815789	0.125	0.130435	0.217391	0.217391	0.434783
0.26087	0.217391	0.478261	1.2	0.12	0.190476	0.380952	0.285714	0.142857
0.221053	0.168421	0.389474	1.3125	0.2	0.409091	0.318182	0.181818	0.090909
0.282828	0.212121	0.494949	1.333333	0.1	0.190476	0.333333	0.380952	0.095238
0.336735	0.265306	0.602041	1.269231	0.178571	0.136364	0.227273	0.272727	0.363636
0.318182	0.190909	0.509091	1.666667	0.208333	0.090909	0.409091	0.363636	0.136364
0.238095	0.22619	0.464286	1.052632	0.041667	0.217391	0.347826	0.304348	0.130435
0.207792	0.194805	0.402597	1.066667	0.038462	0.217391	0.304348	0.347826	0.130435
0.253333	0.16	0.413333	1.583333	0.086957	0.227273	0.272727	0.318182	0.181818
0.306667	0.266667	0.573333	1.15	0.217391	0.181818	0.272727	0.227273	0.318182
0.298507	0.208955	0.507463	1.428571	0.086957	0.26087	0.304348	0.173913	0.26087
0.318841	0.188406	0.507246	1.692308	0.24	0.227273	0.227273	0.272727	0.272727
0.166667	0.152778	0.319444	1.090909	0.133333	0.409091	0.318182	0.136364	0.136364
0.181818	0.106061	0.287879	1.714286	0	0.142857	0.47619	0.238095	0.142857
0.238806	0.179104	0.41791	1.333333	0.173913	0.285714	0.238095	0.380952	0.095238
0.24	0.16	0.4	1.5	0.071429	0.181818	0.363636	0.272727	0.181818

0.287671	0.205479	0.493151	1.4	0.166667	0.272727	0.272727	0.363636	0.090909
0.336735	0.173469	0.510204	1.941176	0.142857	0.318182	0.318182	0.181818	0.181818
0.287356	0.287356	0.574713	1	0.083333	0.304348	0.217391	0.173913	0.304348
0.333333	0.204301	0.537634	1.631579	0.181818	0.181818	0.363636	0.227273	0.227273
0.211765	0.2	0.411765	1.058824	0	0.363636	0.227273	0.272727	0.136364
0.279412	0.352941	0.632353	0.791667	0	0.136364	0.227273	0.272727	0.363636
0.280899	0.258427	0.539326	1.086957	0.296296	0.181818	0.454545	0.227273	0.136364
0.209877	0.17284	0.382716	1.214286	0.148148	0.043478	0.478261	0.26087	0.217391
0.195122	0.182927	0.378049	1.066667	0.074074	0.181818	0.409091	0.227273	0.181818
0.330357	0.339286	0.669643	0.973684	0.045455	0.190476	0.285714	0.285714	0.238095
0.285714	0.25	0.535714	1.142857	0.041667	0.181818	0.318182	0.272727	0.227273
0.352113	0.239437	0.591549	1.470588	0.178571	0.26087	0.173913	0.347826	0.217391
0.283784	0.162162	0.445946	1.75	0.2	0.318182	0.181818	0.272727	0.227273
0.218391	0.172414	0.390805	1.266667	0.12	0.095238	0.428571	0.238095	0.238095
0.215909	0.25	0.465909	0.863636	0.111111	0.2	0.28	0.24	0.28
0.264368	0.183908	0.448276	1.4375	0.148148	0.227273	0.272727	0.272727	0.227273
0.216495	0.185567	0.402062	1.166667	0.068966	0.318182	0.272727	0.090909	0.318182
0.247059	0.235294	0.482353	1.05	0.030303	0.285714	0.238095	0.333333	0.142857
0.175258	0.175258	0.350515	1	0.068966	0.409091	0.181818	0.318182	0.090909
0.356164	0.328767	0.684932	1.083333	0.076923	0.173913	0.130435	0.478261	0.217391
0.320513	0.371795	0.692308	0.862069	0.16	0.26087	0.043478	0.304348	0.391304
0.354167	0.395833	0.75	0.894737	0.107143	0.181818	0.090909	0.545455	0.181818
0.2875	0.3125	0.6	0.92	0.08	0.217391	0.26087	0.347826	0.173913
0.209877	0.197531	0.407407	1.0625	0.142857	0.333333	0.428571	0.142857	0.095238
0.24359	0.230769	0.474359	1.055556	0.08	0.173913	0.434783	0.173913	0.217391
0.268908	0.159664	0.428571	1.684211	0.190476	0.227273	0.454545	0.318182	0
0.283951	0.246914	0.530864	1.15	0.178571	0.227273	0.272727	0.227273	0.272727
0.25	0.152778	0.402778	1.636364	0.136364	0.227273	0.409091	0.363636	0

0.240964	0.180723	0.421687	1.333333	0.148148	0.272727	0.409091	0.318182	0
0.261364	0.272727	0.534091	0.958333	0.24	0.227273	0.272727	0.181818	0.318182
0.258427	0.235955	0.494382	1.095238	0.125	0.227273	0.272727	0.181818	0.318182
0.268293	0.231707	0.5	1.157895	0.24	0.173913	0.173913	0.391304	0.26087
0.333333	0.22619	0.559524	1.473684	0.2	0.26087	0.130435	0.391304	0.217391
0.222222	0.263889	0.486111	0.842105	0.043478	0	0.428571	0.333333	0.238095
0.255556	0.277778	0.533333	0.92	0.076923	0	0.428571	0.333333	0.238095
0.30303	0.282828	0.585859	1.071429	0.037037	0.173913	0.26087	0.391304	0.173913
0.27381	0.297619	0.571429	0.92	0.125	0.173913	0.26087	0.391304	0.173913
0.224719	0.179775	0.404494	1.25	0.08	0.318182	0.227273	0.272727	0.181818
0.185185	0.185185	0.37037	1	0.038462	0.318182	0.227273	0.272727	0.181818
0.272727	0.227273	0.5	1.2	0	0.26087	0.26087	0.26087	0.217391
0.252874	0.183908	0.436782	1.375	0.137931	0.173913	0.347826	0.26087	0.217391
0.197917	0.145833	0.34375	1.357143	0.148148	0.181818	0.454545	0.227273	0.136364
0.254545	0.190909	0.445455	1.333333	0.148148	0.190476	0.428571	0.190476	0.190476
0.290909	0.263636	0.554545	1.103448	0.157895	0.190476	0.428571	0.190476	0.190476
0.235294	0.264706	0.5	0.888889	0.090909	0.227273	0.318182	0.227273	0.227273
0.273973	0.273973	0.547945	1	0.136364	0.318182	0.181818	0.227273	0.272727
0.311688	0.246753	0.558442	1.263158	0.148148	0.318182	0.272727	0.227273	0.181818
0.285714	0.190476	0.47619	1.5	0.192308	0.318182	0.272727	0.227273	0.181818
0.271605	0.123457	0.395062	2.2	0.206897	0.166667	0.416667	0.333333	0.083333
0.222222	0.194444	0.416667	1.142857	0.047619	0.304348	0.26087	0.086957	0.347826
0.2125	0.125	0.3375	1.7	0.074074	0.142857	0.47619	0.238095	0.142857
0.222222	0.17284	0.395062	1.285714	0.172414	0.136364	0.363636	0.318182	0.181818
0.227273	0.190909	0.418182	1.190476	0.071429	0.217391	0.434783	0.304348	0.043478
0.236364	0.272727	0.509091	0.866667	0	0.217391	0.434783	0.304348	0.043478
0.294872	0.192308	0.487179	1.533333	0.285714	0.173913	0.304348	0.391304	0.130435
0.266667	0.186667	0.453333	1.428571	0.230769	0.136364	0.363636	0.409091	0.090909

0.283784	0.175676	0.459459	1.615385	0.227273	0.227273	0.409091	0.363636	0
0.233333	0.211111	0.444444	1.105263	0.035714	0.173913	0.478261	0.217391	0.130435
0.253165	0.151899	0.405063	1.666667	0.230769	0.227273	0.318182	0.227273	0.227273
0.296296	0.209877	0.506173	1.411765	0.178571	0.217391	0.304348	0.217391	0.26087
0.275862	0.264368	0.54023	1.043478	0.111111	0.181818	0.272727	0.272727	0.272727
0.27451	0.235294	0.509804	1.166667	0.04	0.454545	0.181818	0.227273	0.136364
0.19	0.13	0.32	1.461538	0.259259	0.173913	0.521739	0.173913	0.130435
0.275862	0.16092	0.436782	1.714286	0.178571	0.136364	0.5	0.227273	0.136364
0.2	0.218182	0.418182	0.916667	0.041667	0.26087	0.26087	0.26087	0.217391
0.224719	0.235955	0.460674	0.952381	0.12	0.173913	0.347826	0.304348	0.173913
0.290909	0.190909	0.481818	1.52381	0.310345	0.136364	0.454545	0.363636	0.045455
0.290909	0.245455	0.536364	1.185185	0.043478	0.173913	0.304348	0.173913	0.347826
0.147059	0.147059	0.294118	1	0	0.227273	0.409091	0.181818	0.181818
0.272727	0.25	0.522727	1.090909	0.058824	0.142857	0.333333	0.238095	0.285714

miGCcnt	pri lenght		
0.454545	83		
0.409091	84		
0.409091	7		
0.409091	79		
0	87		
0.363636	84		
0.409091	84		
0.5	80		
0.434783	81		
0.428571	82		
0.434783	110		
0.434783	110		
0.454545	85		
0.761905	73		
0.5	86		
0.380952	85		
0.608696	84		
0.545455	97		
0.636364	92		
0.454545	89		
0.47619	82	0.5	110
0.454545	101	0.409091	110
0.590909	73	0.47619	81
0.347826	97	0.590909	73
0.347826	82	0.5	97

0.521739	68
0.454545	100
0.454545	95
0.590909	106
0.318182	86
0.521739	88
0.409091	99
0.409091	73
0.454545	68
0.409091	99
0.608696	89
0.545455	84
0.478261	87
0.5	84
0.391304	65
0.409091	83
0.409091	98
0.478261	84
0.454545	110
0.434783	137
0.416667	110
0.409091	110
0.5	82
0.409091	86
0.434783	71
0.434783	71
0.521739	92
0.47619	110
0.590909	83
0.428571	87
0.454545	84
0.695652	75
0.434783	110
0.363636	82
0.545455	90
0.5	95
0.545455	68
0.454545	110
0.545455	110
0.391304	71
0.478261	69
0.363636	72
0.727273	110
0.5	110
0.478261	110
0.545455	110
0.380952	110
0.454545	110
0.409091	85
0.363636	110
0.47619	85
0.409091	97
0.695652	73
0.695652	78
0.727273	96
0.571429	84
0.333333	77
0.545455	78
0.454545	97
0.363636	64
0.5	88
0.409091	86
0.454545	71
0.409091	88
0	89
0.5	70
0.409091	92
0.52381	71
0.318182	70
0.826087	75
0.590909	94
0.304348	94
0.652174	94
0.428571	69
0.272727	95
0.47619	99
0.636364	98
0.5	110
0.434783	84
0.478261	77
0.5	75
0.545455	75
0.434783	67
0.545455	69
0.272727	72
0.380952	66
0.47619	67
0.454545	75
0.454545	73
0.363636	98
0.478261	87
0.454545	93
0.409091	85
0.636364	68
0.363636	89
0.478261	81
0.409091	82
0.52381	112
0.5	84
0.565217	71
0.5	74
0.47619	87
0.52	88
0.5	87
0.409091	97

0.521739	80
0.238095	81
0.391304	78
0.318182	119
0.5	81
0.363636	72
0.318182	83
0.5	88
0.5	89
0.652174	82
0.608696	84
0.571429	72
0.571429	90
0.565217	99
0.565217	84
0.454545	89
0.454545	81
0.478261	110
0.478261	87
0.363636	96
0.380952	110
0.380952	110
0.454545	68
0.5	73
0.409091	77
0.409091	84
0.416667	81
0.434783	72
0.380952	80
0.5	81
0.347826	110
0.347826	110
0.521739	78
0.5	75
0.363636	74
0.347826	90
0.454545	79
0.478261	81
0.545455	87
0.363636	102
0.304348	100
0.363636	87
0.478261	110
0.478261	89
0.409091	110
0.521739	110
0.363636	68
0.52381	88

Discussion

We have proposed a computational strategy to classify miRNA on the basis of their involvements in cancer or normal cells. Strategy involves the identification of two classes of miRNAs, first class of miRNAs is the one whose expression level are affected during cancer and second class of miRNAs is whose expression level remains unaffected. These miRNA were extracted from mir2disease database and their sequences were downloaded from miRBase. Total 183 validated sequences of cancer related miRNAs were identified. So the data was assigned in such a way that both the classes have 183 sequences. Descriptors have crucial role in prediction of putative miRNA and further classification of its different classes. Broadly these descriptors were categorised in three classes based on their biophysical characteristics.

a) Sequence based descriptors

Different parameters of nucleotide sequence were considered in the study. These descriptors are based on sequence composition, complexity, variation, tandem repeats, contents of nucleic acids and variation in length. Composition and content of nucleotides were studied in ‘pri-seq’, ‘pre-seq’ and ‘miRNAseq’ descriptors. Repeats were considered in ‘ploy A, U, G, C of hairpin and miRNA’. Sequence complexity were extracted in ‘GAsurplusCU’,’ Max di base ratio’, and ‘Min base occurrence’.

b) Structure based descriptors

Primary miRNA has a complex secondary structure which consists of a) central loop, b) upper stem, c) lower stem and d) flanking sequences. Descriptors were identified according to these four components of secondary structure. Descriptor ‘Cloop’ considers the length of central loop in pri-miRNA. ‘Bulge ratio stem’ and ‘bulge ratio miR’ extracts the ratio of asymmetry in stem and miRNA. ‘Wobber matches’, which cannot be considered in primary sequence can also be studied in secondary structure.

c) Energy based descriptors

Mfold and RNAfold software were used to calculate minimum free energy. The core algorithm predicts a minimum free energy, ΔG, as well as minimum free energies for folding that must contain any particular base pair.

We have developed a tool to extract the values of these descriptors. Data was transformed into training set and test set. Most of the classifier showed high accuracy at training data set. MultilayerPerceptron had shown 99.27% accuracy on training set with kappa statistic value 0.9854 and 67% accuracy on test set with kappa statistic value .34. Randomforest had given 99.63% accuracy with kappa statistic value 0.9927 on training set and 71 % accuracy with 0.42 kappa statistic value on test set. J48 algorithm had shown 96.36% accuracy on training set with kappa statistic value 0.972 and 70 % accuracy on test set with kaapa statistic values 0.4. Our results suggest that physical character alone cannot responsible for the classification. Furthermore, gene transcriptional regulators may play an important role in correlation with involvement of miRNAs in cancer. There is a need of the development of improved classification algorithm for more meaning full result. The number of miRNA sequences that we had obtained from experimentally validated data are insufficient. Accuracy could be improved if data size could be increased.

Future prospective

The study of the function of miRNA is based on micro array analysis. Thus, finding can be limited to known miRNAs. Although up to 1000 miRNAs have been identified in human genome but only 200-300 have recognised role in cancer (Griffiths-Jones *et al*., 2006). Majority of miRNAs are still need to be identified. Their role in cancer can also be explored. Computational based *in-silico* identification of miRNA can be used in the high throughput technology to further explore the human genome. Furthermore, the investigation of their role in cancer will provide new insight for diagnosis and therapeutics. Although evidence shows that some miRNAs play an essential role in human cancers, the molecular mechanisms of miRNA regulated pathogenesis are unclear. How miRNAs regulate the multiple stages (initiation, promotion, malignant conversion, progression, and metastasis) of cancers also awaits discovery. Expression pattern of many miRNAs is unique and differential in certain cancer as compared to normal cell. The collective study of these miRNAs with their differential expression pattern can improve the accuracy of diagnostic tools. miRNAs function as tumor suppressor and/or oncogenes, it might be possible to use miRNA/anti mRNA as a tool to promote RNAi. RNAi can be used to study gene function and gene therapy. Artificial miRNAs could be synthesized to down-regulate oncogenes and prevent the formation of cancer. To achieve this goal, effective delivery system needs to be developed. How miRNA can be delivered to its target and its continuous activity is a challenge. After these obstacles have been overcome, miRNA-related techniques will have a brilliant future and become new cancer therapeutic tools.

References

Adam, J. A. (1997). General aspects of modeling tumor growth and immune response. In *A survey of models for tumor-immune system dynamics* (pp. 15-87). Birkhäuser Boston.

Holman, A. G., & Gabuzda, D. (2012). A machine learning approach for identifying amino acid signatures in the HIV env gene predictive of dementia. *PloS one*, *7*(11), e49538.

Omer, A., Singh, S., & Duhan, N. (2012). Identification of miRNAs in C. roseus and their potential targets. *Bioinformation*, *8*(2), 75.

Strillacci, A., Griffoni, C., Valerii, M. C., Lazzarini, G., Tomasi, V., & Spisni, E. (2010). RNAi-based strategies for cyclooxygenase-2 inhibition in cancer. *BioMed Research International*, *2010*.

van der Burgt, A., Fiers, M. W., Nap, J. P., & van Ham, R. C. (2009). In silico miRNA prediction in metazoan genomes: balancing between sensitivity and specificity. *BMC genomics*, *10*(1), 204.

Zhang, B., Pan, X., Cobb, G. P., & Anderson, T. A. (2007). microRNAs as oncogenes and tumor suppressors. *Developmental biology*, *302*(1), 1-12.

Batuwita, R., & Palade, V. (2009). microPred: effective classification of pre-miRNAs for human miRNA gene prediction. *Bioinformatics*, *25*(8), 989-995.

Bonnet, E., Wuyts, J., Rouzé, P., & Van de Peer, Y. (2004). Detection of 91 potential conserved plant microRNAs in Arabidopsis thaliana and Oryza sativa identifies important target genes. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(31), 11511-11516.

Cai, X., Hagedorn, C. H., & Cullen, B. R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *Rna*, *10*(12), 1957-1966

Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., ... & Croce, C. M. (2002). Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences*, *99*(24), 15524-15529.

Calin, G. A., Ferracin, M., Cimmino, A., Di Leva, G., Shimizu, M., Wojcik, S. E., ... & Croce, C. M. (2005). A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *New England Journal of Medicine*, *353*(17), 1793-1801.

Carrano, A. C., Eytan, E., Hershko, A., & Pagano, M. (1999). SKP2 is required for ubiquitin-mediated degradation of the CDK inhibitor p27. *Nature cell biology*, *1*(4), 193-199.

Chan, J. A., Krichevsky, A. M., & Kosik, K. S. (2005). MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer research*, *65*(14), 6029-6033.

Chen, X. (2005). MicroRNA biogenesis and function in plants. *FEBS letters*, *579*(26), 5923-5931.

Chendrimada, T. P., Gregory, R. I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., & Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature*, *436*(7051), 740-744.

Chiarle, R., Fan, Y., Piva, R., Boggino, H., Skolnik, J., Novero, D., ... & Inghirami, G. (2002). S-phase kinase-associated protein 2 expression in non-Hodgkin's lymphoma inversely correlates with p27 expression and defines cells in S phase. *The American journal of pathology*, *160*(4), 1457-1466.
Ciafre, S. A., Galardi, S., Mangiola, A., Ferracin, M., Liu, C. G., Sabatino, G., ... & Farace, M. G. (2005). Extensive modulation of a set of microRNAs in primary glioblastoma. *Biochemical and biophysical research communications*, *334*(4), 1351-1358.

Doench, J. G., & Sharp, P. A. (2004). Specificity of microRNA target selection in translational repression. *Genes & development*, *18*(5), 504-511.

Fang, L., Hu, Q., Hua, Z., Li, S., & Dong, W. (2008). Growth inhibition of a tongue squamous cell carcinoma cell line (Tca8113) in vitro and in vivo via siRNA-mediated down-regulation of skp2. *International journal of oral and maxillofacial surgery*, *37*(9), 847-852.

Freyhult, E., Gardner, P. P., & Moulton, V. (2005). A comparison of RNA folding measures. *BMC bioinformatics*, *6*(1), 241.

Takeshita, F., & Ochiya, T. (2006). Therapeutic potential of RNA interference against cancer. *Cancer science*, *97*(8), 689-696.
Gregory, R. I., Yan, K. P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., & Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. *Nature*, *432*(7014), 235-240.

Griffiths-Jones, S., Grocock, R. J., Van Dongen, S., Bateman, A., & Enright, A. J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic acids research*, *34*(suppl 1), D140-D144.

Hayashita, Y., Osada, H., Tatematsu, Y., Yamada, H., Yanagisawa, K., Tomida, S., ... & Takahashi, T. (2005). A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. *Cancer research*, *65*(21), 9628-9632.

He, H., Jazdzewski, K., Li, W., Liyanarachchi, S., Nagy, R., Volinia, S., ... & de La Chapelle, A. (2005). The role of microRNA genes in papillary thyroid carcinoma. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(52), 19075-19080.

He, L., Thomson, J. M., Hemann, M. T., Hernando-Monge, E., Mu, D., Goodson, S., ... & Hammond, S. M. (2005). A microRNA polycistron as a potential human oncogene. *nature*, *435*(7043), 828-833.

Hofacker, I. L. (2003). Vienna RNA secondary structure server. *Nucleic acids research*, *31*(13), 3429-3431.

Hu, Y., Cherton-Horvat, G., Dragowska, V., Baird, S., Korneluk, R. G., Durkin, J. P., ... & LaCasse, E. C. (2003). Antisense oligonucleotides targeting XIAP induce apoptosis and enhance chemotherapeutic activity against human lung cancer cells in vitro and in vivo. *Clinical cancer research*, *9*(7), 2826-2836.

Huse, J. T., Brennan, C., Hambardzumyan, D., Wee, B., Pena, J., Rouhanifard, S. H., ... & Holland, E. C. (2009). The PTEN-regulating microRNA miR-26a is amplified in high-grade glioma and facilitates gliomagenesis in vivo. *Genes & development*, *23*(11), 1327-1337.

Iorio, M. V., Ferracin, M., Liu, C. G., Veronese, A., Spizzo, R., Sabbioni, S., ... & Croce, C. M. (2005). MicroRNA gene expression deregulation in human breast cancer. *Cancer research*, *65*(16), 7065-7070.

Jiang, P., Wu, H., Wang, W., Ma, W., Sun, X., & Lu, Z. (2007). MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. *Nucleic acids research*, *35*(suppl 2), W339-W344.

Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., ... & Slack, F. J. (2005). < i> RAS</i> Is Regulated by the< i> let-7</i> MicroRNA Family. *Cell*, *120*(5), 635-647.

Kadri, S., Hinman, V., & Benos, P. V. (2009). HHMMiR: efficient de novo prediction of microRNAs using hierarchical hidden Markov models. *BMC bioinformatics*, *10*(Suppl 1), S35.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., ... & Kim, V. N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *nature*, *425*(6956), 415-419.

Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H., & Kim, V. N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *The EMBO journal*, *23*(20), 4051-4060.

Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., & Burge, C. B. (2003). Prediction of mammalian microRNA targets. *Cell*, *115*(7), 787-798.

Ma, J. B., Ye, K., & Patel, D. J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature*, *429*(6989), 318-322.

MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V., & Doudna, J. A. (2008). In vitro reconstitution of the human RISC-loading complex. *Proceedings of the National Academy of Sciences*, *105*(2), 512-517.

Meltzer, P. S. (2005). Cancer genomics: small RNAs with big impacts. *Nature*, *435*(7043), 745-746.

Michael, M. Z., O'Connor, S. M., van Holst Pellekaan, N. G., Young, G. P., & James, R. J. (2003). Reduced Accumulation of Specific MicroRNAs in Colorectal Neoplasia11Note: Susan M. O'Connor and Nicholas G. van Holst Pellekaan contributed equally to this work. *Molecular Cancer Research*, *1*(12), 882-891.

Murakami, Y., Yasuda, T., Saigo, K., Urashima, T., Toyoda, H., Okanoue, T., & Shimotohno, K. (2005). Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues. *Oncogene*, *25*(17), 2537-2545.

Nam, J. W., Kim, J., Kim, S. K., & Zhang, B. T. (2006). ProMiR II: a web server for the probabilistic prediction of clustered, nonclustered, conserved and nonconserved microRNAs. *Nucleic acids research*, *34*(suppl 2), W455-W458.

Nam, J. W., Shin, K. R., Han, J., Lee, Y., Kim, V. N., & Zhang, B. T. (2005). Human microRNA prediction through a probabilistic co-learning model of sequence and structure. *Nucleic acids research*, *33*(11), 3570-3581.

Loong, S. N. K., & Mishra, S. K. (2007). Unique folding of precursor microRNAs: quantitative evidence and implications for de novo identification. *Rna*, *13*(2), 170-187.

Oulas, A., Boutla, A., Gkirtzou, K., Reczko, M., Kalantidis, K., & Poirazi, P. (2009). Prediction of novel microRNA genes in cancer-associated genomic regions—a combined computational and experimental approach. *Nucleic acids research*, gkp120.

Clote, P., Ferré, F., Kranakis, E., & Krizanc, D. (2005). Structural RNA has lower folding energy than random RNA of the same dinucleotide frequency. *Rna*, *11*(5), 578-591.

Dubois, R. N., Abramson, S. B., Crofford, L., Gupta, R. A., Simon, L. S., Van De Putte, L. B., & Lipsky, P. E. (1998). Cyclooxygenase in biology and disease. *The FASEB journal*, *12*(12), 1063-1073.

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., ... & Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *nature*, *403*(6772), 901-906.

Robins, H., & Press, W. H. (2005). Human microRNAs target a functionally distinct population of genes with AT-rich 3′ UTRs. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(43), 15557-15562.

Sewer, A., Paul, N., Landgraf, P., Aravin, A., Pfeffer, S., Brownstein, M. J., ... & Zavolan, M. (2005). Identification of clustered microRNAs using an ab initio prediction method. *BMC bioinformatics*, *6*(1), 267.

He, S., Zhang, D., Cheng, F., Gong, F., & Guo, Y. (2009). Applications of RNA interference in cancer therapeutics as a powerful tool for suppressing gene expression. *Molecular biology reports*, *36*(8), 2153-2163.

Song, J. J., Liu, J., Tolia, N. H., Schneiderman, J., Smith, S. K., Martienssen, R. A., ... & Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nature Structural & Molecular Biology*, *10*(12), 1026-1032.

Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., ... & Takahashi, T. (2004). Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer research*, *64*(11), 3753-3756.

Thieme, C. J., Gramzow, L., Lobbes, D., & Theißen, G. (2011). SplamiR—prediction of spliced miRNAs in plants. *Bioinformatics*, *27*(9), 1215-1223.

Voorhoeve, P. M., Le Sage, C., Schrier, M., Gillis, A. J., Stoop, H., Nagel, R., ... & Agami, R. (2006). A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. *Cell*, *124*(6), 1169-1181.

Wu, C. Y., Wu, M. S., Chiang, E. P., Wu, C. C., Chen, Y. J., Chen, C. J., ... & Lin, J. T. (2007). Elevated plasma osteopontin associated with gastric cancer development, invasion and survival. *Gut*, *56*(6), 782-789.

Xue, C., Li, F., He, T., Liu, G. P., Li, Y., & Zhang, X. (2005). Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine. *BMC bioinformatics*, *6*(1), 310.

Zeng, Y., & Cullen, B. R. (2004). Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucleic acids research*, *32*(16), 4776-4785.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research*, *31*(13), 3406-3415.

miRBase: microRNA sequences, targets and gene nomenclature.Nucleic Acids Res. 34, D140–D144.

APPENDIX

1. Code for extracting values of descriptors

```
#include "stdio.h"
#include "string.h"
#include "fold.h"
#include  "utils.h"
#include  "fold_vars.h"
#include <math.h>

main()
{
char s[400] ,g[30] , h[30], a1[200],mf[100],code[90],mircode[30],hcode[50],seq5h[60],seq5b[90],seq3h[60],seq3b[90];
char fname[20], gname[40] ,hname[30],mfname[20],l2[60] ,l3[60],l4[60],l5[60];
int counts=0,countg=0,j=0,k=0,i=0,a=0 ,counth=0 ,code_cnt=0,v=0,f=0,
bulge=0,line=0,polA=0,pola=0,polG=0,polg=0,polC=0,polc=0,polU=0,mpolA=0,mpola=0,mpolG=0,mpolg=0,mpolC=
0,mpolc=0,mpolU=0,mpolu=0,polu=0,hlen=0,
b=0,countp=0,q=0,p=0,c=0,hbond=0,m=0,o=0,cloop=0,greater=0,lesser=0,stem=0,h_greater=0,h_lesser=0,nn=0,x=0,la
=0;
float energy1 ,add=0,h_add=0,least=0,h_least=0,sblg=0,mblg=0,dp_match=0, energy2
,miA=0,miU=0,miG=0,miC=0,migc=0,GAsurplusCU=0,piA=0,piU=0,piG=0,piC=0,pigc=0,hiA=0,hiU=0,hiG=0,hiC=0
,higc=0,gu_contri,code_match=0,wobber=0;
char hairpin[80],*seq,*pin,*structure1,*structure2;

/*printf("Enter .txt file name\n");*/
scanf("%s",fname);

/*printf("enter mirna5p file\n");*/
scanf("%s",gname);

/*printf("enter 3p\n");*/
scanf("%s",hname);
/*printf("mfoldfile\n");*/
scanf("%s",mfname);


FILE *inputf;
FILE *inputg;
FILE *inputh;
FILE *inputmf;

inputf=fopen(fname,"r");
inputg=fopen(gname,"r");
inputh=fopen(hname,"r");
inputmf=fopen(mfname,"r");

printf("id %s\n",fname);
                                        while(fgets(s,399,inputf)!=NULL)
{
                                         printf("%s",s);
}
fclose(inputf);
counts=strlen(s);
counts--;
printf("string length %d\n",counts);
                                       while(fgets(g,29,inputg)!=NULL)
{
                                      printf("%s",g);
}
fclose(inputg);
```

```
countg=strlen(g);
countg--;
printf("m ir len%d",countg);
                                    while (fgets(h,29,inputh)!=NULL)
{
                                  /*printf("%s",h);*/
}
fclose(inputh);
counth=strlen(h);
counth--;
printf("anti %d",counth);
/*-----------------------------------------------------------------------------------------*/
  for(i=0;i<=counts;i++)
{
nn++;
/*printf("%c",s[i]);*/
if(s[i]==g[j])
{
k++;
j++;
}

if(nn==countg)
{
x++;
if(k==countg)
{
a=x-1;
break;
}
/*printf("%d %d\n",x,k);*/
i=i+1;
i=i-countg;
nn=0;
k=0;
j=0;
}
}

printf("\na=%d\n",a);

/*----------------------------------------------------------------------------------------------------------------------*/
k=0;
j=0;
x=0;
nn=0;
                             for(i=0;i<=counts;i++)
{
nn++;
/*printf("%c",s[i]);*/
if(s[i]==h[j])
{
k++;
j++;
}

if(nn==counth)
{
x++;
if(k==counth)
{
b=x+counth-2;
break;
```

```
}
/*printf("%d %d\n",x,k);*/
i=i+1;
i=i-counth;
nn=0;
k=0;
j=0;
}
}

printf("b=%d\n",b);


/*------------------------------------------------------------------------------------------------------------------*/
  j=0;
printf("\nhairpin :");
                              for(i=a;i<=b;i++)
{
printf("%c",s[i]);
if(s[i]=='A')
{
hiA++;
}
if(s[i]=='U')
{
hiU++;
}
if(s[i]=='G')
{
hiG++;
}
if(s[i]=='C')
{
hiC++;
}

hairpin[j]=s[i];

j++;
}
higc=hiG+hiC;
countp=strlen(hairpin);
/*foreign libraries*/
seq=s;
pin=hairpin;
temperature= 37.;
initialize_fold(strlen(seq));
initialize_fold(strlen(pin));
structure1 = (char* ) space(sizeof(char)*(strlen(seq)+1));
structure2 = (char* ) space(sizeof(char)*(strlen(pin)+1));

energy1=fold(seq , structure1);
energy2=fold(pin , structure2);

printf("\npri-\npri-mfe %f\nMFEpril(dG):%f " ,energy1,energy1/counts);
printf("\n%s",structure1);

printf("\nhairpin-\nharipin-mfe%f\nMFEhail(dG):%f                                                                     \nMFEhai
index%f:",energy2,energy2/countp,energy2/countp*1/higc);

for(i=0;i<=countp-1;i++)
{
/*printf("%c",structure[i]);*/
```

```
a1[i]=structure2[i];

}

if(a1[i]=='.' && a1[i+1]==')')
{

q=i+1;
;
break;
}
}
for(i=q;i>=0;i--)
{
if(a1[i]=='(')
{

p=i;

break;

}
}

cloop=q-p-1;
stem = countp-cloop;
printf("\n\n;;;;;;;;;;p=%d %d q=%d  \ncentral loop length (nt) %d \nhairpin length=%d",p,countp-q,q,cloop,stem);
printf("\nstem details");
 /*MFE Adjusted for Stem Length*/

/*----------------------------------------------------------------------------------------------------*/
  p=0;
  q=0;
                     while (fgets(mf,99,inputmf)!=NULL)
{
                       printf("%s",mf);
                        line++;
                       if(line==8)
{
                      la=strlen(mf);

for(i=la-3;i>=la/2;i--)
{
if(mf[i]==' ')
{
p=i;
break;
}
}
a=0;
for(i=0;i<=p;i++)
{
if((mf[i]==' ') || (mf[i]=='|'))
l2[a]='m';
else if(mf[i]=='.')
l2[a]='*';
else
l2[a]=mf[i];

a++;
}
}
                      if(line==9)
```

```
{
   /*printf("%s",mf);*/
for(i=0;i<=p;i++)
{
l3[i]=mf[i];
}


}

                              if(line==10)
{
for(i=0;i<=p;i++)
l4[i]=mf[i];
}
                              if(line==11)
{

for(i=0;i<=p;i++)
{
if((mf[i]==' ') || (mf[i]=='^'))

l5[v]='m';

else if ((mf[i]=='A')||(mf[i]=='U')||(mf[i]=='G')||(mf[i]=='C')||(mf[i]=='-'))

l5[v]=mf[i];

else
l5[v]='*';
v++;

}
}


}

fclose(inputmf);
/*=====================================================================*/


         for(i=0;i<=p;i++)
{
if((l3[i]=='A')||(l3[i]=='U')||(l3[i]=='G')||(l3[i]=='C'))
seq5b[i]=l3[i];
else
seq5b[i]=l2[i];

}
for(i=0;i<=p;i++)
{
if(((l4[i]=='A')||(l4[i]=='U')||(l4[i]=='G')||(l4[i]=='C')))
seq3b[i]=l4[i];
else
seq3b[i]=l5[i];
}
for(i=0;i<=p;i++)
{
code_cnt++;
if((l2[i]=='m') && (l5[i]=='m'))
{
code[i]='m';
```

```
dp_match++;
}
else if ((l2[i]=='-') && ((l5[i]=='A')||(l5[i]=='U')||(l5[i]=='G')||(l5[i]=='C')))
code[i]='b';
else if((l5[i]=='-') && ((l2[i]=='A')||(l2[i]=='U')||(l2[i]=='G')||(l2[i]=='C')))
code[i]='B';
else if ((l2[i]=='*')&&(l5[i]=='*'))
code[i]='s';
else if (((l2[i]!='m')&&(l5[i]=='m')) || ((l2[i]=='m')&&(l5[i]!='m')))
{
code[i]='s';
seq5b[i]='s';
seq3b[i]='s';
}
else
code[i]='x';

}

printf("%s          %d\n%s\n%s\n%s\n%s\n%s\n%s\nAdjusted     base     pairing     propensity     (dP)     %f
",l2,p,l3,l4,l5,seq5b,seq3b,code,dp_match/counts);

/*---------------------------------------------------------------------------------------------------------------------------------*/


k=0;

j=32;
for(i=code_cnt-1;i>=1;i--)
{

if(code[i]=='s')
f++;
else
{
k++;
hcode[j]=code[i];
seq5h[j]=seq5b[i];
seq3h[j]=seq3b[i];
j--;
}
if(k==33)
{
hcode[j]='\0';
seq5h[j]='\0';
seq3h[j]='\0';
break;
}
}

hlen=strlen(hcode);
for(j=0;j<=hlen-1;j++)
{
if((seq5h[j]=='G') && (seq3h[j]=='U') || (seq5h[j]=='U') && (seq3h[j]=='G'))
{
wobber++;
}
}
for(i=0;i<=hlen-1;i++)
{
if(hcode[i]=='b'||hcode[i]=='B')
sblg++;
else if(hcode[i]=='m')
```

```
code_match++;
}
printf("\nstem code : %s%d %d\nbluge ratio stem:%f \nstem5p:%s\nstem3p:%s\n%d\nwobber vs all matches(stem)%f",hcode,code_cnt,q,sblg/hlen,seq5h,seq3h,j,wobber/code_match);
/*---------------------------------------------------------------------------*/
k=0;
j=0;
a=0;
x=0;
nn=0;
 for(i=0;i<=code_cnt;i++)
{
nn++;
/*printf("%c",s[i]);*/
if(s[i]==g[j])
{
k++;
j++;
}

if(nn==countg)
{
x++;
if(k==countg)
{
a=x-1;
break;
}
/*printf("%d %d\n",x,k);*/
i=i+1;
i=i-countg;
nn=0;
k=0;
j=0;
}
}

printf("\na=%d\n",a);

j=countg-1;
for(i=a+countg-1;i>=a;i--)
{
mircode[j]=code[i];
if((mircode[j]=='b' ) || (mircode[j]=='B'))
{
mblg++;
}
j--;
}
printf("\n miRNA code%s \nbulge ratio mir: %f",mircode,mblg/countg);
/*---------------------------------------------------------------------------*/
for(i=0;i<=counts-1;i++)
{
if(s[i]=='A')
{
piA++;
}
if(s[i]=='U')
{
piU++;
}
if(s[i]=='G')
{
```

```
piG++;
}
if(s[i]=='C')
{
piC++;
}

}
pigc=piG+piC;

if((piA>piU)&&(piA>piG)&&(piA>piC))
{
greater=piA;
            if((piU>piG)&&(piU>piC))
  {
    lesser=piU;
  }

{
lesser=piG;
}
       else
{
lesser=piC;
}
}
else if ((piU>piA)&&(piU>piG)&&(piU>piC))
{
greater=piU;
 if((piA>piG)&&(piA>piC))
  {
    lesser=piA;
  }
else if ((piG>piA)&&(piG>piC))
{
lesser=piG;
}
else
{
lesser=piC;
}
}
else if ((piG>piA)&&(piG>piU)&&(piG>piC))
{
greater=piG;
 if((piU>piA)&&(piU>piC))
  {
    lesser=piU;
  }
else if ((piA>piU)&&(piA>piC))
{
lesser=piA;
}
else
{
lesser=piC;
}
}
else
{
greater=piC;
                if((piU>piG)&&(piU>piA))
  {
```

```
    lesser=piU;
  }
else if ((piG>piU)&&(piG>piA))
{
lesser=piG;
}
else
{
lesser=piA;
}
}
/*----------------------------------------------------------------------------*/
if((hiA>hiU)&&(hiA>hiG)&&(hiA>hiC))
{
h_greater=hiA;
             if((hiU>hiG)&&(hiU>hiC))
  {
    h_lesser=hiU;
  }
     else if ((hiG>hiU)&&(hiG>hiC))
{
h_lesser=hiG;
}
        else
{
h_lesser=hiC;
}
}
else if ((hiU>hiA)&&(hiU>hiG)&&(hiU>hiC))
{
h_greater=hiU;
 if((hiA>hiG)&&(hiA>hiC))
  {
    h_lesser=hiA;
  }
else if ((hiG>hiA)&&(hiG>hiC))
{
h_lesser=hiG;
}
else
{
h_lesser=hiC;
}
}
else if ((hiG>hiA)&&(hiG>hiU)&&(hiG>hiC))
{
h_greater=hiG;
 if((hiU>hiA)&&(hiU>hiC))
  {
    h_lesser=hiU;
  }
else if ((hiA>hiU)&&(hiA>hiC))
{
h_lesser=hiA;
}
else
{
h_lesser=hiC;
}
}
else
{
h_greater=hiC;
```

```
          if((hiU>hiG)&&(hiU>hiA))
  {
    h_lesser=hiU;
  }
else if ((hiG>hiU)&&(hiG>hiA))
{
h_lesser=hiG;
}
else
{
h_lesser=hiA;
}
}
if((piA<piU)&&(piA<piG)&&(piA<piC))
least=piA;
else if((piU<piA)&&(piU<piG)&&(piU<piC))
least=piU;
else if((piG<piU)&&(piG<piA)&&(piG<piC))
least=piG;
else
least=piC;

if((hiA<hiU)&&(hiA<hiG)&&(hiA<hiC))
h_least=hiA;
else if((hiU<hiA)&&(hiU<hiG)&&(hiU<hiC))
h_least=hiU;
else if((hiG<hiU)&&(hiG<hiA)&&(hiG<hiC))
h_least=hiG;
else
h_least=hiC;
printf("\npri-miRNA seq A=%f U=%f G=%f C=%f GC=%f G/C ratio=%f \nMFEpri index
%f",piA/counts,piU/counts,piG/counts,piC/counts,pigc/counts,piG/piC,energy1/counts*1/pigc);
add=greater+lesser;
printf("\n (pri)max di base count=%d,%d ratio=%f \n(pri)minimal base
occurrance=%f",greater,lesser,add/counts,least/counts);
printf("\npre-miRNA seq A=%f U=%f G=%f C=%f GC=%f G/C
ratio=%f",hiA/countp,hiU/countp,hiG/countp,hiC/countp,higc/countp,hiG/hiC);
h_add=h_greater+h_lesser;
printf("\n (pre)max di base count=%d,%d ratio=%f \n(pre)minimal base
occurrance=%f",h_greater,h_lesser,h_add/countp,h_least/countp);

for(i=0;i<=countg-1;i++)
{
if(g[i]=='A')
{
miA++;
}
if(g[i]=='U')
{
miU++;
}
if(g[i]=='G')
{
miG++;
}
if(g[i]=='C')
{
miC++;
}

}
migc=miG+miC;
```

```
printf("\nmiRNA deails A=%f U=%f G=%f C=%f GC=%f
",miA/countg,miU/countg,miG/countg,miC/countg,migc/countg);

GAsurplusCU = ( miG + miA - miC - miU ) / ( miG + miA + miC + miU );
printf("\nGAsurplusCU=%f\n",GAsurplusCU);

for(i=0;i<=countp-1;i++)
{
if(hairpin[i]=='A')
polA++;
else
{
if((polA>pola) || (polA==pola))
{
pola=0;
pola=polA;
polA=0;
}
else if(pola>polA)
polA=0;
}
}

for(i=0;i<=countp-1;i++)
{
if(hairpin[i]=='U')
polU++;
else
{
if((polU>polu) || (polU==polu))
{
polu=0;
polu=polU;
polU=0;
}
else if(polu>polU)
polU=0;
}
}

for(i=0;i<=countp-1;i++)
{
if(hairpin[i]=='G')
polG++;
else
{
if((polG>polg) || (polG==polg))
{
polg=0;
polg=polG;
polG=0;
}
else if(polg>polG)
polG=0;
}
}

for(i=0;i<=countp-1;i++)
{
if(hairpin[i]=='C')
polC++;
else
{
```

```
if((polC>polc) || (polC==polc))
{
polc=0;
polc=polC;
polC=0;
}
else if(polc>polC)
polC=0;
}
}

printf("poly details hairpin polyA =%d polyU=%d polyG=%d polyC=%d\n",pola,polu,polg,polc);

for(i=0;i<=countg-1;i++)
{
if(g[i]=='A')
mpolA++;
else
{
if((mpolA>mpola) || (mpolA==mpola))
{
mpola=0;
mpola=mpolA;
mpolA=0;
}
else if(mpola>mpolA)
mpolA=0;
}
}

for(i=0;i<=countg-1;i++)
{
if(g[i]=='U')
mpolU++;
else
{
if((mpolU>mpolu) || (mpolU==mpolu))
{
mpolu=0;
mpolu=mpolU;
mpolU=0;
}
else if(mpolu>mpolU)
mpolU=0;
}
}

for(i=0;i<=countg-1;i++)
{
if(g[i]=='G')
mpolG++;
else
{
if((mpolG>mpolg) || (mpolG==mpolg))
{
mpolg=0;
mpolg=mpolG;
mpolG=0;
}
else if(mpolg>mpolG)
mpolG=0;
}
}
```

```
for(i=0;i<=countg-1;i++)
{
if(g[i]=='C')
mpolC++;
else
{
if((mpolC>mpolc) || (mpolC==mpolc))
{
mpolc=0;
mpolc=mpolC;
mpolC=0;
}
else if(mpolc>mpolC)
mpolC=0;
}
}
printf("mi_poly details miRNA mpolyA =%d mpolyU=%d mpolyG=%d mpolyC=%d\n",mpola,mpolu,mpolg,mpolc);

}
```

2. SHELL Code for automation

```
#!bin/bash

for list; do
if [ -r $list ]; then

for j in $(strings $list); do

echo "$j"
./abc < ~/Documents/name/$j >> outputfilter


done
fi
done
```

3 mfold automation

```
#!bin/bash

for list; do
if [ -r $list ]; then

for j in $(strings $list); do


mfold SEQ="$j"
rm *.ct
rm *.png
rm *.a*
rm *.g*
rm *.l*
rm *.d*
rm *.h*
rm *.ps
rm *.plot
rm *.pnt
rm *.ps
rm *.ss
```

```
rm *.ss-count
rm *.sav

done
fi
done
```

4 Sequence extraction

```
#!bin/bash
for list; do
if [ -r $list ]; then

for j in $(strings $list); do

grep -i -w "$j" sequence > temp
awk '{print $2}' temp > $j

done
fi
done
```