

Off-target Possibility Prediction of Efficient siRNA: A Workflow

Reena Murali, David Peter S

Abstract - For knocking down targeted genes, small interfering RNAs (siRNAs) are used as an important tool in biological and biomedical research, which is called gene silencing by RNA interference. This is done by degradation of the target mRNA. These target mRNAs must be selected such that their corresponding siRNAs are likely to be efficient against that target and unlikely to accidentally silence other transcripts due to sequence similarity. Hence while designing efficient siRNAs, the ability to knock down target genes and off-target effect on non-target genes must be considered. In this work we present the work flow of our design that enables to predict efficient siRNAs that can identify the off target possibility on non-target genes. Hence we hope that the risk of “off target effect“ while doing gene silencing in various bioinformatics fields can be overcome, by carefully examining the inhibition efficiency and off target possibility of each siRNA.

Keywords - small interfering RNA, messenger RNA, RNA interference, gene silencing

I. INTRODUCTION

RNAi is an important biological process which can control the gene regulation by a sequence specific post transcriptional gene silencing mechanism [1] mediated by double stranded RNA (dsRNA). The RNAi pathway was discovered by Fire and Mello in 1998 [2]. RNAi has been successfully used to target diseases such as AIDS [3], neurodegenerative diseases [4], cholesterol [5] and cancer [6] on mice with the hope of extending these approaches to treat humans. In functional genomic research, the discovery of RNA interference has become much helpful in drug design and therapeutic applications because of its ability to perform gene silencing. The RNAi can be endogenous or exogenous. Recently a considerable amount of work has been done to understand the gene silencing mediated by siRNA. Hence, designing of siRNA which has good target silencing efficiency with good specificity of selected target mRNA [7,8] or gene is a research issue to be addressed up on.

Reena Murali is Associate Professor with the Department of Computer Science and Engineering, Rajiv Gandhi Institute of Technology, Kerala, India (corresponding author e-mail: reena.rajesh@rit.ac.in).

David Peter S. is Professor with the Department of Computer Science and Engineering, Cochin University of Science and Technology, Kerala, India.

There exist several good scoring siRNA design models like Biopredsi[9], DSIR[10], ThermoComposition21[11], i-Score[12], Scales[13], My siRNA-Designer[14] and MysiRNA[15] to predict possible siRNAs targeting the mRNAs. But recent studies reveals that among all siRNAs that can be generated against a target mRNA, only a fraction are successful in causing degradation and all siRNAs do not perform equal knockdown effects [16]. The efficiency of siRNA differs in different target sites of same mRNA. Thus the aim of siRNA efficiency prediction is to design siRNA sequences that are highly capable of inhibiting their target mRNA sequences. Earlier it was understood that full complementary siRNA is needed to silence a target gene. But later, many studies shown that siRNA behaves like miRNA and suppress protein synthesis even though it is not fully complementary to the target. This shows that mismatches are allowed during target selection by siRNA [17-18]. This may cause a very serious problem of “off-target effect” where unintended target genes may be suppressed by selected siRNA [19-21]. In this paper, we present the work flow of our design OpsiD [24], that enables to identify the efficacy and off target effect of siRNAs against target genes. The tool lists all siRNAs against a particular target with their inhibition efficacy and number of matches or sequence similarity with other genes in the data base. Hence we hope that the risk of “ off target effect “ while doing gene silencing in various bioinformatics fields can be overcome, by carefully examining the inhibition efficiency and off target possibility of each siRNA using this siRNA design method. The comparison and validation of our previous models were done with existing good scoring siRNA design models [22-24].

II. MATERIALS AND METHODS

A. Software Used:

For computing the metrics to be used as input parameters to our neural network model, we used the tools DSIR, i-Score Designer, ThermoComposition21 and MysiRNA designer, all of which must be downloaded from their respective sources before using our model, OpsiD. All the input parameters except MysiRNA score are read from the i-Score designer Excel file. The MysiRNA scores are read from the output file of the MysiRNA designer tool.

B. NCBI BLAST

The NCBI BLAST tool (blastall) is used to filter out siRNAs with high off-target effect by running a BLAST search on the NCBI RefSeq database. The blastall tool is

bundled along with OpsiD, but the NCBI RefSeq database must be downloaded separately. Website: <ftp://ftp.ncbi.nlm.nih.gov/blast/executables/release/LATEST>

C. Encog and Encog Workbench IDE

The Encog machine learning framework for Java is used to create and use the siRNA designer neural network model. Encog is an advanced, lightweight Java machine learning framework which can be used to develop common neural network and other machine learning models (including Support Vector Machines, Genetic Algorithms, Bayesian Networks, Hidden Markov Models etc). The Encog Workbench IDE was used to easily design and test the model. The IDE provides an easy-to-use graphical interface to design various neural network configurations, and also to train as well as test the neural network using various neural network training algorithms.

Website: <http://www.heatonresearch.com/encog>

D. Apache POI

The Apache POI library was used to read and write Microsoft Excel files, such as the i-Score designer Excel file. The Apache POI library is an open-source library developed by the Apache Software Foundation which provides a set of Java APIs for creating and manipulating Microsoft Office Documents (both the new OOXML formats and the old OLE2 Compound Document Format). It is available under version 2.0 of the Apache License. Website: <http://poi.apache.org/>

E. Data Sets

Our neural network was trained using the experimental inhibition capacity values of the siRNAs in the Huesken data set [9]. We named this data set as Data Set 1. This data set contains a total of 2431 siRNAs derived from 30 genes, and their corresponding experimental inhibition capacity values. We used entire Huesken data set for training our neural network model. For testing our neural network, we used some other data sets as well. Our second data set consists of 419 siRNAs (named as Data Set 2, which is mutually exclusive with Data Set1) taken from various sources such as Reynolds, Vickers, Haborth, Ui-Tei and Khovorova [18-21]. This data set was compiled by Ichihara et al for their i-Score designer tool [13] and was used by Mysara et al for testing their MysRNA model [15]. For conducting validations, we used another data set of 476 siRNAs (named as Data Set 3, which is entirely different from Data Set 1 and Data Set 2) presented by Mysara et al. [15] in their work. These 476 siRNAs were originally taken from a larger data set of 18,593 siRNAs introduced by Fellman et al. [27].

F. Neural Network Model

The final efficacy score (inhibition capacity) of siRNA was computed using a multi-layer perceptron, feed-forward neural network trained using the Scaled Conjugate Gradient training algorithm [25] (provided by Encog), which is based upon a class of optimization techniques well known in numerical analysis as the Conjugate Gradient Methods. Our Neural Network Model has 5 neurons in the input layer, a single hidden layer of 12 neurons and 1 neuron in the output layer (5*12*1 neural network). The training began from a

randomized state and was done for 136,000 iterations. The neural network model was built and trained using the Encog Workbench IDE and later integrated into our siRNA designer tool (OpsiD).

G. Input values

The final inhibition capacity score of each siRNA was computed with the following five different metrics: whole ΔG (or stacking energy), DSIR [10], ThermoComposition21 [11], i-Score [12], and MysRNA [15]. These metrics were found to work well and gave good results with the Huesken data set in our experiments.

H. Normalization of Input values

The input values given to the neural network, i.e. the 5 metrics described above, are normalized using the range normalization method, also known as min-max normalization. That is, the normalized input values are given by:

$$A_i^N = \frac{A_i - \min_A}{\max_A - \min_A} (\max_R - \min_R) + \min_R$$

Where A_i^N is the normalized value of the metric A for the i th tuple of the training data, A_i is the actual value, \max_A is the maximum value of the metric A for the entire training set of siRNAs and \min_A is the minimum value. The values are normalized to the range [minR, maxR].

The minimum and maximum values of each input metric, obtained from the training data set, are supplied along with our program, and are used for normalizing the input values. The input values for the neural network are normalized to the range [-1, 1] before being given as input. The experimental inhibition values from the training data set were also normalized to the range [0, 1] before training the neural network.

I. Output Values

The neural network gives a single output value in the range [0,1], which is then multiplied by 100 to give the final score which is displayed for each siRNA in our OpsiD tool.

III. WORKFLOW

A. Prerequisites

The siRNA designer software (OpsiD) requires the following software/files to be downloaded and installed:

- Java 7
- Any Perl Distribution (tested with ActivePerl)
- i-Score Designer Excel file
- ThermoComposition19 and ThermoComposition21
- MysRNA designer and model file
- NCBI BLAST tool (blastall) †
- NCBI RefSeq RNA BLAST database †
- Encog Workbench ‡

† Required only for BLAST search filtering (optional)

‡ Required only for testing the neural network model

B. Input

The software takes an mRNA or cDNA gene sequence as input. The nucleotides may be specified using the uppercase

letters A, T, G, C and U or the corresponding lowercase letters, and any spaces or newlines within the sequence are ignored. The user can also enter the RefSeq number of the gene if it has been taken from the NCBI RefSeq RNA database. This is only required if the user selects to perform BLAST search.

C. Processing

The siRNA designer software first enumerates each possible 19-mer siRNA sequences from the input nucleotide sequence. Then, it computes some basic parameters for each siRNA sense strand, such as GC content percentage, whole delta G, delta H, delta S, melting temperature etc. For calculating whole delta G, we use the nearest-neighbor model from Sugimoto et al [26], which is also used by other tools such as i-Score and MysiRNA. If the user has selected the option to filter siRNAs based on GC content, the software then removes all siRNAs which do not have GC content % between the minimum and maximum values specified by the user. Then, it gives each siRNA sequence as input to various pre-existing second generation siRNA designer tools such as i-Score, ThermoComposition21, DSIR and MysiRNA to get their scores for the strand's inhibition capacity. The siRNA strand's i-Score, ThermoComposition21, DSIR, MysiRNA scores and the initially calculated whole delta G value are given as input to our neural network model. The model gives an output value in the range [0, 1] which is then multiplied by 100 to give the displayed final score (in the range 0 – 100) for the siRNA strand. If the user has selected the option to filter siRNAs based on their score, the software will remove all siRNAs whose final scores lie below a certain threshold value specified by the user.

D. Off Target Possibility Prediction

When the user selects BLAST search option in our tool for off target finding, it will run BLAST for each siRNA sequence on the NCBI RefSeq RNA database. This database is not included in our tool but must be downloaded and placed in the suitable location. The output will display the greatest BLAST score of each siRNA sequence. BLAST score is the number of nucleotide matches of each siRNA with the database ie., Blast score 19 for an siRNA denotes a perfect match of that siRNA with any other genes in the database. For example, in Table.1 the siRNA with best inhibition efficacy against a given target site is 93%, but the BLAST score 18, means even though the siRNA is best efficient to degrade the target mRNA, there is a high risk of 'off target effect' with 18 nucleotide matches to any other sequences in the data base. So according to our tool, instead of selecting siRNA with best inhibition capacity alone, we can consider both efficiency and number of matches of BLAST score to select siRNAs for gene silencing.

There is also an exclude list option, which is necessary if the input mRNA is from the RefSeq RNA database itself. In that case, the user can specify that matches to that particular mRNA need not be considered while finding the greatest BLAST score. This is done by specifying the RefSeq numbers of the mRNAs to exclude. After all these

processes, the siRNA designer software, Opsid, displays the output. The software displays the output in tabular form. All the siRNAs after the selected filtrations will be displayed, along with the various values such as whole delta G, GC content, Biopredsi, i-Score, ThermoCompositon21, DSIR, MysiRNA, BLAST score and final score (Opsid Score). The software also provides the option to export the tabular data to a CSV format file, which can then be read by various statistical analysis tools such as Microsoft Excel and IBM SPSS etc. Fig.1 and Fig.2 shows the sample screen shot for user interface and output showing inhibition efficacy with BLAST score of Opsid respectively.

TABLE I
SAMPLE siRNAs WITH INHIBITION CAPACITY AND BLAST SCORE IN OPSID

siRNA	siRNA Inhibition Capacity	BLAST Score (No of nucleotide matches with other sequence in Data Base)
AGGGUUAUUUUUCUUUGGC	75	15
GAAAAAAACCAAAGGGUUA	67	5
AACCACUGUAGAAAAU AAC	55	0
UCUUUAUGUUUUUGGCGUC	93	18
UUCUUUAUGUUUUUGGCGU	66	9
GGGCCUUUCUUUAUGUUUU	55	7
UUAAAAUGUCGUUCGCGG	80	12
UAAUUUUUUGGAUGAUUGG	45	7
UUAAAAUCGCAGUAUCCGG	77	5

IV. CONCLUSION

Many models have been developed earlier to predict and evaluate the siRNA inhibition efficacy. We have analyzed the best scoring models and algorithms currently available, and developed a neural network model by combining the results of selected previous models to improve the prediction accuracy. In addition, we have also included modules to reduce the consequence of 'off target' effect. The proposed artificial neural network model 'Opsid' can predict siRNA efficiency over a targeted mRNA sequence and can identify their similarity score in the data base, and able to predict the possibility of 'off target' effect. Some research area of drug design in cancer therapy is concentrating to artificially inject exogenous siRNA capable of degrading the mRNA responsible for tumor development. The proposed soft computing model may be useful in finding exogenous siRNAs capable of degrading the disease causing target mRNA "without doing off target silencing" and may help in drug design for cancer treatment and other areas of bioinformatics by 'gene silencing'. One of the importance of the study include, as it clearly address the off target effect which was a very serious problem to be dealt with while designing siRNA for therapeutic applications.

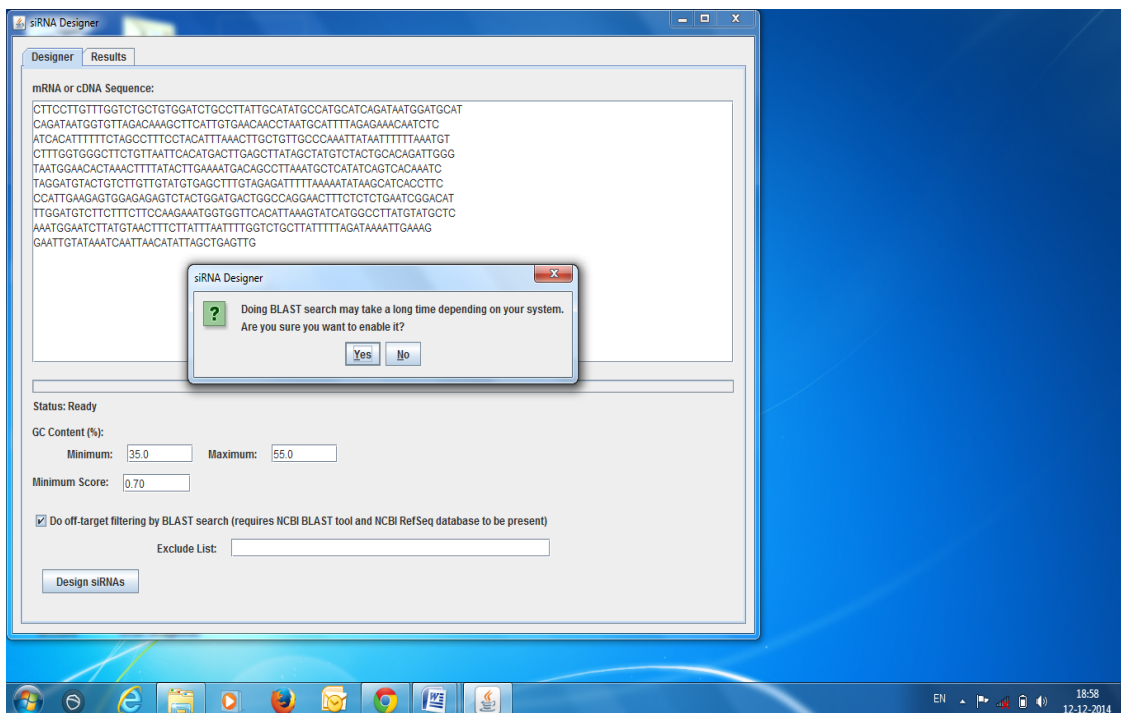


Fig 1. Sample Screen Shot showing the user interface of Opsid

	A	B	C	D	E	F	G	H	I	J	K	L	M
133	Position	Sense strand	Antisense strand	GC Content	Whole deltaG	s-Biopredsi	i-Score	DSIR	ThermoComposition21	MySiRNA	BLAST Score	Opsid Score (%)	
134	134	CAAUUAUCGCAUGAUUU	AAUAUCGCAUGAUUUUggu	26.32%	-29.2	0.836	69.792	85.603	0.81	89.568	16	65.971	
135	135	AAUAUAUCGCAUGAUUUU	AAUAUAUCGCAUGAUUUUggg	21.05%	-28	0.781	58.619	74.756	0.72	70.861	16	51.272	
136	136	AUAUAUCGCAUGAUUUUU	AAAUAUAUCGCAUGAUUUUg	21.05%	-28	0.75	57.598	72.546	0.76	71.853	17	49.407	
137	137	UAUAUCGCAUGAUUUUUU	CAAAAUAUCGCAUGAUUUuu	26.32%	-29	0.61	49.249	61.127	0.69	59.279	17	37.108	
138	138	AUAUCGCAUGAUUUUUUG	CCAAAAUAUCGCAUGAUUau	31.58%	-31	0.448	40.712	54.663	0.54	49.599	17	36.694	
139	139	UAUCGCAUGAUUUUUUGG	ACCAAAAUAUCGCAUGAUua	31.58%	-32.1	0.444	39.777	57.695	0.67	54.67	16	35.957	
140	140	AUCGCAUGAUUUUUUGGU	AACCAAAAUAUCGCAUGau	31.58%	-31.7	0.688	55.155	71.555	0.82	81.268	15	56.44	
141	141	UCGCAUGAUUUUUUGGUA	UAACCAAAAUAUCGCAua	31.58%	-31.9	0.75	59.964	76.95	0.87	88.035	14	60.962	
142	142	CGCAUGAUUUUUUGGUAA	UUAACCAAAAUAUCGCAu	31.58%	-30.4	0.871	82.355	94.758	1.01	96.492	13	79.938	
143	143	GCAUGAUUUUUUGGUAAA	UUUAACCAAAAUAUCGga	26.32%	-28.9	0.877	81.739	96.537	0.98	96.092	12	81.02	
144	144	CAUGAUUUUUUGGUAAAAG	CUUUAAACCAAAAUAUCGcg	26.32%	-27.6	0.709	55.257	69.435	0.72	65.111	0	44.979	
145	145	AUGAUUUUUUGGUAAAAGA	UCUUAAACCAAAAUAUCgc	21.05%	-27.9	0.755	63.11	74.752	0.77	79.519	12	54.015	
146	146	UGAUUUUUUGGUAAAAGAU	AUCUUAAACCAAAAUAUCug	21.05%	-27.9	0.837	65.59	84.243	0.86	86.327	13	61.61	
147	147	GAUUAUUUUUGGUAAAAGAU	AAUCUUAAACCAAAAUAUcau	21.05%	-26.7	0.806	70.367	86.453	0.89	89.308	14	65.9	
148	148	AUAUUUUUGGUAAAAGAUUU	AAAUUUUAAACCAAAAUAuca	15.79%	-25.2	0.768	66.389	78.147	0.76	76.177	14	54.596	
149	149	UAUUUUUGGUAAAAGAUUUU	UAUUUUUAAACCAAAAUAuuc	15.79%	-25.4	0.831	73.878	81.84	0.8	87.011	14	62.437	
150	150	AUUUUUGGUAAAAGAUUUU	AUAAUUUUUAAACCAAAAua	15.79%	-25.2	0.739	57.655	74.125	0.67	59.301	15	44.008	
151	151	UUUUUGGUAAAAGAUUUUUC	GAUAAUUUUUAAACCAAAAua	21.05%	-26.5	0.525	43.586	60.306	0.59	48.721	16	36.622	
152	152	UUUUUGGUAAAAGAUUUUCG	CGAUAAUUUUUAAACCAAAu	26.32%	-28	0.402	40.069	54.765	0.51	47.275	17	40.968	
153	153	UUUUUGGUAAAAGAUUUUCGG	CCGAUAAUUUUUAAACCAaaa	31.58%	-30.4	0.353	36.692	53.449	0.53	47.525	18	38.217	
154	154	UGGUUUAAAAGAUUUUUCGGA	UCGGAUAAUUUUUAAACCAaa	31.58%	-31.9	0.719	58.957	76.582	0.79	83.934	19	58.081	
155	155	GGUUAAAAGAUUUUUCGGAC	GUCCGAUAAUUUUUAAACcaa	36.84%	-32	0.711	59.581	77.188	0.84	86.732	19	59.869	
156	156	GUUAAAAGAUUUUUCGGACA	UGUCCGAUAAUUUUUAAACca	31.58%	-30.8	0.776	66.993	83.638	0.8	89.081	19	64.046	
157	157	UUAAAAGAUUUUUCGGACAG	CUGUCCGAUAAUUUUUAAacc	31.58%	-30.7	0.372	41.524	52.962	0.51	48.999	19	37.713	

Fig 2. Sample Screen Shot showing the output with BLAST Score of Opsid

OpsID Work Flow

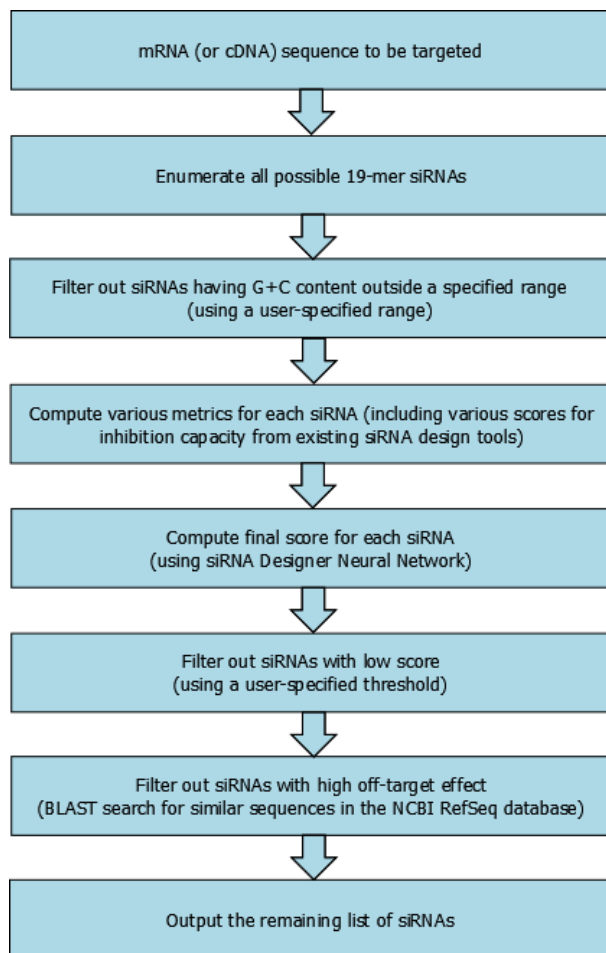


Fig 3: OpsID Work Flow

Acknowledgments

The authors acknowledge the support extended by TEQIP-Phase II (MHRD world bank project) for the research work. The authors thank the Central Library, Cochin University of Science and Technology, Indian Institute of Technology Kharagpur for providing access to necessary literature articles.

REFERENCES

- [1] S.M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, and T. Tuschl, "Duplexes of 21 nucleotide RNAs mediate RNA interference in cultured mammalian cells", *Nature* 411 (2001) 494-498.
- [2] A. Fire, S. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver, and C.C. Mello, "Potent and specific genetic interference by double stranded RNA in *c. elegans*", *Nature*, 391 (1998) 806-811.
- [3] A. Martínez Miguel, Gutiérrez Arantxa, Armand-Ugón Mercedes, Blanco Julià, Parera Mariona, Gómez Jordib, Clotet Bonaventura, and A. Esté José, "Suppression of chemokine receptor expression by rna interference allows for inhibition of hiv1 replication", *AIDS*. 16 (2002) 2385-2390.
- [4] H. Xia, Q. Mao, S.L. Eliason, S.Q. Harper, I.H. Martins, H.T. Orr, H.L. Paulson, L. Yang, R.M. Kotin, and B.L. Davidson, "Rnai suppresses poly glutamine-induced neuro degeneration in a model of spinocerebellar ataxia", *Nature Medicine* 10 (2004) 816 -820.

- [5] J. Soutschek, A.Akinc, B. Bramlage, K. Charisse, R.Constien, and M.Donoghue, et al., "Therapeutic silencing of an endogenous gene by systemic administration of modified sirnas", *Nature* 432 (2004) 173-178.
- [6] A. Borkhardt, "Blocking oncogenes in malignant cells by RNA interference new hope for a highly specific cancer treatment?", *Cancer Cell* 2(3) (2002) 167-168.
- [7] G.J. Hannon, and J.J. Rossi, "Unlocking the potential of the human genome with RNA nterference", *Nature* 431(2004) 371-378.
- [8] G. Meister, and T. Tuschl, "Mechanisms of gene silencing by double stranded RNA", *Nature* 431 (2004) 343-349.
- [9] D. Huesken, J. Lange, C. Mickanin, J. Weiler, F. Asselbergs, J.Warner, and B.Meloon, et al., "Design of a genome-wide siRNA library using an artificial neural network", *Nature Biotechnology*. 23 (2005) 995-1001.
- [10] J.P. Vert, N. Foveau, C. Lajaunie, and Y. Vandenbrouck, "An accurate and interpretable model for siRNA efficacy prediction", *BMC Bioinformatics*. 7:520 (2006).
- [11] S.A. Shabalina, A.N. Spiridonov, and A.Y. Ogurtsov, "Computational models with thermodynamic and composition features improve siRNA design", *BMC Bioinformatics*. 7:65 (2006)
- [12] M. Ichihara, Y.Murakumo, A. Masuda, T. Matsuura, N. Asai, M. Jijiwa, and M. Ishida, et al., "Thermodynamic instability of siRNA duplex is a prerequisite for dependable prediction of siRNA activities", *Nucleic Acids Research*. 35 (2007) e123.
- [13] O. Matveeva, Y. Nechipurenko, L. Rossi, B. Moore, A.Y. Ogurtsov, and J.F. Atkins, "Comparison of approaches for rational siRNA design leading to a new efficient and transparent method", et al. *Access* 35 (2007) 1-10.
- [14] M. Mysara, J.M. Garibaldi, and M. Elhefnawi, "MysiRNA-Designer a workflow for efficient siRNA design", *PLoS One* 6 (2011) e25642,
- [15] M. Mysara, M. Elhefnawi, and J. M. Garibaldi, "MysiRNA: Improving siRNA efficacy prediction using a machine-learning model combining multi-tools and whole stacking energy (ΔG)", *Journal of Biomedical Informatics*. 45 (2012) 528-534
- [16] T. Holen, M. Amarzguioui, M.T. Wiiger, E. Babaie, and H. Prydz, "Positional effects of short interfering rnas targeting the human coagulation trigger tissue factor", *Nucleic Acids Research*. 30(2002) 1757-1766.
- [17] S.L Ameres, J.Martinez, and R. Schroeder, "Molecular basis for target RNA recognition and cleavage by human RISC.", *Cell* 130(2007), 101-112.
- [18] J.G. Doench, C.P. Petersen, and P.A. Sharp, "siRNAs can function as miRNAs.", *Genes Dev* 17(2003) 438-442.
- [19] J.Burchard , A.L. Jackson, V. Malkov, R.H. Needham, and Y. Tan, et al, "MicroRNA-like off-target transcript regulation by siRNAs is species specific.", *RNA* 15(2009) 308-315.
- [20] A.L. Jackson, S.R. Bartz, J. Schelter, S.V. Kobayashi, and J. Burchard, et al. "Expression profiling reveals off-target gene regulation by RNAi.", *Nat Biotechnol*. 21(2003) 635-637.
- [21] A.L. Jackson, J. Burchard, J. Schelter, B.N. Chau, and M. Cleary M, et al. "Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity.", *RNA* 12(2006) 1179-1187.
- [22] R.Murali, and D. Peter S, "Computational Model for Predicting Effective siRNA Sequences for Gene Silencing ", Eight International Conference on Data Mining and Warehousing (ICDMW-2014), 2014, pp 138-142.
- [23] R. Murali, and D. Peter S, "Computational Model for Predicting Effective siRNA Sequences Using Whole Stacking Energy (ΔG) for Gene Silencing", World Academy of Science, Engineering and Technology, International Science Index 97, International Journal of Biological, Veterinary, Agricultural and Food Engineering, 9(1)(2015), 6 - 12.
- [24] R. Murali, P. G.John, and D. Peter S, "Soft computing model for optimized siRNA design by identifying off target possibilities using artificial neural network model", *GENE-Elsevier (Article in press)*, 2015, doi:10.1016/j.gene.2015.02.067

- [25] M. Fodslette Moller, "A scaled conjugate gradient algorithm for fast supervised learning", *Neural Networks*. 6(1993) 525-533.
- [26] N. Sugimoto, S.Nakano, M. Katoh, A.Matsumura, H. Nakamuta, T. Ohmichi, M. Yoneyama, and M.Sasaki, "Thermodynamic Parameters To Predict Stability of RNA/DNA Hybrid Duplexes", 34 (1995) 11211-11216
- [27] C.Fellmann, J .Zuber, K .McJunkin, K .Chang, and C.D Malone, et al., "Functional identification of optimized RNAi triggers using a massively parallel sensor assay", *Molecular cell*. 41(2011), 733–746.