Vertebrate microRNA genes and CpG-islands

KA-LOK NG [a*], CHIEN-HUNG HUANG [b], MING-CHENG TSAI [a†]
[a] Department of Bioinformatics
Asia University
500 Lioufeng Road, Wufeng Shiang, Taichung, Taiwan 41354
[b] Department of Computer Science and Information Engineering
National Formosa University
64,Wen-Hwa Road, Hu-wei, Yun-Lin, Taiwan 632
[*]Corresponding author, Email: ppiddi@gmail.com, [†]moved to ScinoPharm Taiwan, Ltd. August, 2007

Abstract: - A portion of five vertebrate species microRNA (miRNA) genes are found to associate with CpG-islands. It is calculated that 74 of 462 (16.0%), 37 of 373 (9.9%), 25 of 234 (10.7%), 25 of 149 (16.8%) and 11 of 177 (6.2%) of the miRNA genes are located within 1000 bp of a CpG-island for human, mouse, rat, chicken and frog respectively. A statistical test is proposed to quantify the probability distribution of the relative distance between a miRNA gene and a CpG-island, the results suggested that there could be some spatial association between them. In the human and mouse samples the Fisher's exact test indicated that some of the promoter-related miRNA genes are CpG-island-associated. These findings suggest that a potential role of these miRNAs related to CpG-island methylation, which deserve further investigation. A web-based interface has been implemented to facilitate data display, in which the vertebrate CpG-associated miRNA genes results are publicly available.

Key-Words: - microRNA; promoter regions; CpG-islands; methylation; gene regulation

1 Introduction

MicroRNA (miRNA) genes encode small RNA molecules involved in mRNA translation and degradation by the RNA interference (siRNA) machinery [1,2]. MicroRNAs (miRNAs) are a large evolutionarily conserved class of noncoding RNAs (ncRNAs) which are about 18–22 nucleotides long and mediate posttranscriptional silencing of genes [3]. MiRNAs were first discovered in *Caenorhabditis elegans* [4] with the identification of the *lin-4* and *let-7* miRNA genes, which act as posttranscriptional repressors of target genes by antisense binding to their 3' untranslated regions. Shortly thereafter, hundreds of other miRNAs were found in worms as well as in flies, plants, and vertebrates (for review, see [1,5,6]). Rapid progress has begun to uncover the genetic roles of miRNAs in development and other biological processes. For example, in *C. elegans, let-7* and *lin-4* miRNAs function as heterochronic genes, and mutations in either disrupt proper specification of cell fates [4]. In *Drosophila*, a mutation in *miR-14* leads to a disruption in normal patterns of cell death and to defects in fat metabolism. In mammals, approximately 230 miRNAs have been identified from a vast array of tissues and cell types [3,6,7,8].

There are several studies suggesting that miRNAs may have a role in the localized expression of genes in their neighborhood. In *C. elegans*, Inaoka et al. [9] found that there is a decrease in mRNA abundance, localized within a 10kb of chromosomal distance of a majority of miRNAs. The same group also found decreased expression of coding gene mRNA that is localized to and centered on the position of miRNA throughout the *Mus musculus* genome [10]. Sevignani et al. [11] reported that there is a significant association between the chromosomal location of miRNAs and the mouse cancer susceptibility loci.

Recently, it is reported that some miRNAs have been implicated in human cancers. For example, the expression of the human miR-127 is specifically downregulated in multiple solid cancers. This miRNA is located within a CpG-island, which is specifically hypermethylated in cancers [12]. Another work also reported that four miRNAs located around CpG islands (i.e., miR-34b, miR-137, miR-193a, and miR-203) are silenced by DNA hypermethylation in oral cancer [13]. Furthermore, it is demonstrated that miR-124a undergoes transcriptional inactivation by CpG island hypermethylation in different cancer cell types [[14]]. The above works demonstrated that miRNA

could be silenced by epigenetic mechanisms [[15]]. Among the six cancer-related miRNA genes reported above, five are located within the promoters at no more than 10 kbp far (except miR-193a). Since human cancer is tightly linked to the presence of CpG-island promoter hypermethylation [[14]], it is interesting to know if the same mechanism could regulated miRNA expression, therefore, promoter-related miRNA genes are identified as well. Being motivated by this reason, we proposed to investigate if there is a distance-association of miRNA with other genetic components, such as the CpG-island, and promoter.

CpG-islands are regions where CpGs are present at significantly higher levels than is typical for the genome as a whole [[16]]. CpG-islands are associated with genes, particularly housekeeping genes, in vertebrates. CpG-islands are typically common near transcription start sites (TSS), are often associated with promoter regions, and some are hyper-methylated in cancer [[17]]. It is suggested that the effect of a CpG-island near the TSS could be more important than the global GC content of the region where the gene resides [[18],[19]].

Based on this information, we used the UCSC Genome Browser [[20],[21]] to find whether vertebrates (human, mouse, rat, chicken and frog) miRNA genes are associated with CpG-islands. This was followed by a simple sequence homology search using NCBI-BLAST [[22]], to find whether human miRNA genes are located within promoter regions (promoter data were obtained from DBTSS [[18],[19]]). We found that a proportion (~6% or more) of miRNA genes were located within 1000 bp of a CpG-island.

A subset of miRNA genes are associated with CpG-islands and promoter regions as well, suggesting a potential role for these miRNAs in CpG-island methylation and gene regulation. In plants, siRNAs targeted to CpG-islands within a promoter regions could induce RNA directed DNA-methylation [[23],[24]]. It would be logical to suspect that miRNAs close to or overlapping with CpG-islands could be involved in autoregulation of the methylation of their respective CpG-islands.

2 Materials and Methods

The vertebrate miRNA precursor sequences were downloaded from the miRNA Registry [[25],[26],[27]] release 8.1 (http://microrna.sanger.ac.uk/sequences/) and the vertebrate promoter sequences were retrieved from DBTSS release 5.1 (http://dbtss.hgc.jp/), where each promoter sequence contains a 1 kb upstream sequence and 200 bp downstream sequence from each Transcription Start Site (TSS) described. A total of 30964 TSS were collected in our analysis. The vertebrate miRNA precursor sequences were aligned using NCBI-BLAST against the vertebrate promoter sequences (i.e. 100% identity), consequently, it locates precursor sequences originating from the promoter regions. The CpG-island information associated with miRNA genes was obtained from the UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway, March 2006 draft).

To characterize the spatial association between miRNA genes and CpG-islands, we introduce a uniform distance distribution model to describe the relative distance between these two genetic components. The method is based on the assumption that miRNA genes and CpG-islands are evenly distributed along the chromosome. Let , n and L be the total number of known miRNA genes, number of CpG-islands along a chromosome, and the size of the chromosome respectively. The relative distance between a miRNA gene and a CpG-island, D, is defined by the following expression,

$$D = \left| \frac{\alpha}{\mu} - \frac{j}{n} \right| L \qquad (1)$$

where $1 \le \alpha \le \mu$, $1 \le j \le n$ and $|x|$ denotes the absolute value of x. Given , n and L one can obtained a probability distribution function of D, and compute the cumulative probability that D is less than , i.e. $prob(D < \delta)$. A small probability value means that it is highly unlikely that a miRNA gene is located near a CpG-island with a distance less than . For approximation, we assume 10% of the known miRNA genes are uniformly distributed along chromosome one for the species under study. It is noted that all the species have at least 20 pairs of chromosomes, therefore, the 10% estimation is a reasonable upper limit. Statistics of the total number of CpG-islands for the vertebrate species are available from NCBI Genome biology web site, http://www.ncbi.nlm.nih.gov/Genomes/.

To test the association between miRNA genes and promoter regions, the Fisher's exact test [[28]] is employed. Let X and Y be the number of successes from two independent Bernoulli samples, and let x and y be the corresponding observed values. Let $n = n_1 + n_2$ be the total sample size, and $m = x +$

y be the total observed number of successes. Table 1 presents the data in the form of a 2×2 table,

Table 1. Table for data from two independent Bernoulli samples.

	Success	Failure	Row total
Sample 1	x	$n_1 - x$	n_1
Sample 2	y	n_2 - y	n_2
Column total	m	n-m	n

Fisher's exact test uses X, the number of successes from sample 1, as the test statistics. When null hypothesis H0 is true, the successes are equally likely from the two samples. Consider the upper one-sided testing problem: H0: p1 = p2 vs. H1: p1 > p2. The P-value is the probability that at least x of the total m successes come from sample 1 when H0 is true:

$$P-value = P(X \geq x \mid X+Y=m) = \sum_{i \geq x} \frac{\binom{n_1}{i}\binom{n_2}{m-i}}{\binom{n}{m}} \quad (2)$$

where the summation extends over i = x, x+1, … to an upper limit which is either n1 or m, whichever is smaller. In the present study, success and failure columns correspond to near and far from CpG-island events respectively, whereas sample 1 and 2 denote miRNA gene that is near and far from promoter respectively.

3 Results

3.1 miRNA, CpG-islands and promoters

A portion of human miRNA genes, 74 of 462 (16.0%), were found either internal to or to lie within 1000 bp of a CpG-island (Table 2). This result is consistent with the work of Holstebroe and Tommerup [[29]]. In this study, we extended their work to a larger set of human miRNA genes, and four more species; i.e. mouse, rat, chicken, and frog, miRNA genes are studied. Table 2 shows the results for the distances of human miRNA genes measured from genes and CpG-islands. In Table 2, the first column denotes the distances of genes and CpG-islands from human miRNA genes. The word 'internal' means an miRNA gene overlaps with the CpG-island, and the word 'far' (near) means the distances of genes or CpG-islands from human miRNA genes are longer (shorter) than 1000 bp. For instance, 14 miRNA genes are located internal to a gene and a CpG-island.

Among the 462 human miRNA genes, 239 of them are located in intergenic region, the rest can be grouped into exonic, intronic or mixed type of miRNA genes. From Table 2, it can be seen that 42 of the miRNA genes are associated with CpG-islands as well as genes. For coding genes, *in vivo* data indicates that around 70% of human genes are CpG-associated [[30]].

Table 2. Distances of human miRNA genes measured from genes and CpG islands.

Distances of genes and CpG-islands from human miRNA genes	Number of miRNA genes
gene [far] and CpG-island [internal]	19
gene [far] and CpG-island [near]	13
gene [near (upstream)] and CpG-island [internal]	8
gene [near (upstream)] and CpG-island [near]	3
gene [near (downstream)] and CpG-island [internal]	2
gene [near (downstream)] and CpG-island [near]	0
gene [internal] and CpG-island [internal]	14
gene [internal] and CpG-island [near]	15
Total	74

Table 3 shows the subset of human miRNA genes which are located near the promoter region of a gene, and may be sitting near a CpG-island. It was found that seven out of the ten promoter-related miRNA genes are CpG-island-associated. Also, we have identified eight miRNA genes where each one is located near a promoter on the opposite strand. This information can be found at our web site, http://bioinfo.csie.nfu.edu.tw/html/dna.php. A recent study reported that miRNA could target promoter sequence and induce gene expression [[31]]. The Fisher's exact test is applied to test the association between miRNA genes and promoter regions. In the human sample, we have i=x=7, y=35, m=42, n_1=10, n_2=213, and n=223, so according to Eq. (3), $P(x = 7 \mid X + Y = 42)$ =3.2*10^{-5}. The small *P-value* (less than 0.01) suggested that promoter-related miRNA gene could possibly associate with CpG-island region.

Table 3. Human miRNA genes that are located within promoter regions. MiRNA and promoter regions are located on the same strand.

pre-miRNA	internal or near CpG-islands	DBTSS promoter ID (gene group ; alternative)	Gene ID	Gene Name
hsa-mir-30c-1	No	4439|13 ; 3	4802	NFYC
hsa-mir-103-2	No	10812|13 ; 3	80025	PANK2
hsa-mir-186	No	3779|12 ; 2	9406	ZNF265
hsa-mir-137	Yes	2951|12 ; 2	400765	FLJ35409
		2951|11 ; 1		
hsa-mir-219-1	Yes	615|11 ; 1	6015	RING1
hsa-mir-594	Yes	15894|11 ; 1	154791	HSPC268
hsa-mir-637	Yes	5214|13 ; 3	1613	DAPK3
hsa-mir-639	Yes	5770|11 ; 1	9524	GPSN2
hsa-mir-611	Yes	14517|11 ;1	746	C11orf10
hsa-mir-632	Yes	8401|12 ; 2	7756	ZNF207
		8401|13 ; 3		

In the mouse sample, it was found that 37 of 373 miRNA genes (9.9%) lie within 1000 bp of a CpG-island (Table 4). The total number of miRNA genes that are associated with CpG-islands is reported as 38 because one of the miRNA genes is located within 1000 bp of two CpG-islands. Among the 373 miRNA genes, 200 of them are located in intergenic region. From Table 4, it can be seen that 17 of the miRNA genes are associated with CpG-islands as well as genes. It is interesting to note that experimental data indicates that around 60% of all mouse genes are CpG-associated [[32]].

Table 4. Distances of mouse miRNA genes measured from genes and CpG-islands.

Distances of neighboring genes and CpG-islands from mouse miRNA genes	Number of miRNA genes
gene [far] and CpG-island [internal]	13
gene [far] and CpG-island [near]	8
gene [near (upstream)] and CpG-island [internal]	4
gene [near (upstream)] and CpG-island [near]	1
gene [near (downstream)] and CpG-island [internal]	2
gene [near (downstream)] and CpG-island [near]	0
gene [internal] and CpG-island [internal]	6
gene [internal] and CpG-island [near]	4
Total	38

Table 5 shows the subset of mouse miRNA genes which are promoter-related and may be located near CpG-islands. It was found that a portion of promoter-related miRNA genes are CpG-related as well, that is four out of seven (sense) and four out of four (anti-sense).

Table 5. Mouse miRNA genes that are located within promoter regions. In the strand column, 'senses' and 'anti-sense' denote that the miRNA gene is located on the same and opposite side of the gene's promoter strand respectively.

Strand	pre-miRNA	internal or near CpG-island	DBTSS promoter ID (gene group; alternative)	Gene ID	Gene Name
sense	mmu-mir-199a-2	No	4691,1	18152	Npn1
	mmu-mir-9-2	No	1941,1	320203	C130071C03Rik
	mmu-mir-208	No	11686,2	17888	Myh6
	mmu-mir-707	Yes	1189,2	67605	Akt1s1
	mmu-mir-688	Yes	8307,1	67942	Atp5g2
	mmu-mir-718	Yes	6899,1	16179	Irak1
	mmu-mir-219-1	Yes	12254,1	19763	Ring1
anti-sense	mmu-mir-685	Yes	12105,1	11546	Parp2
	mmu-mir-707	Yes	201,1	233204	Tbc1d17
	mmu-mir-762	Yes	284,2	12055	Bcl7c
			284,1	12055	Bcl7c
	mmu-mir-320	Yes	11838,1	67065	Polr3d

For the Fisher's exact test, we have i=x=4, y=13, m=17, n_1=7, n_2=159, and n=166, so $P(x=4 \mid X+Y=17)=2.1*10^{-3}$. Since the *P-value*

is smaller than 0.01, it is suggesting that promoter-related miRNA gene and CpG-island could possibly associated.

In the rat sample, it was found that 25 of 234 rat miRNA genes (10.7%) lie within 1000 bp of a CpG-island (Table 6). Among the 234 miRNA genes, 140 of them are located in intergenic regions. From Table 6, it can be seen that 5 of the miRNA genes are associated with CpG-islands as well as genes. The promoter information for the rat sample is not available, therefore, we cannot provide a promoter-proximal miRNA gene result. It has previously been reported that in the rat around 47% of genes are CpG-associated [[32]].

Table 6. Distances of rat miRNA genes measured from genes and CpG-islands.

Distances of neighboring genes and CpG-islands from rat miRNA genes	Number of miRNA genes
gene [far] and CpG-island [internal]	14
gene [far] and CpG-island [near]	6
gene [near (upstream)] and CpG-island [internal]	1
gene [near (upstream)] and CpG-island [near]	1
gene [near (downstream)] and CpG-island [internal]	1
gene [near (downstream)] and CpG-island [near]	0
gene [internal] and CpG-island [internal]	1
gene [internal] and CpG-island [near]	1
Total	25

A portion of the chicken miRNA genes, 25 of 149 (16.8%) were found to lie within 1000 bp of a CpG-island (Table 7). Among the 149 miRNA genes, 97 of them are located in intergenic regions. It has been reported that in chicken around 48% of genes are CpG-associated [[33],[34]]. Microchromosomes are richer in CpG-islands and replicate earlier than do macrochromosomes [[35],[36],[37]].

Table 7. Distances of chicken miRNA genes measured from genes and CpG-islands.

Distances of neighboring genes and CpG-islands from chicken miRNA genes	Number of miRNA genes
gene [far] and CpG-island [internal]	10
gene [far] and CpG-island [near]	10
gene [near (upstream)] and CpG-island [internal]	0
gene [near (upstream)] and CpG-island [near]	0
gene [near (downstream)] and CpG-island [internal]	1
gene [near (downstream)] and CpG-island [near]	2
gene [internal] and CpG-island [internal]	2
gene [internal] and CpG-island [near]	0
Total	25

A portion of frog miRNA genes, 11 of 177 (6.2%), are found to lie within 1000 bp of a CpG-island (Table 8). It has been reported that the frog genome is CG-poor compared to vertebrate [[16]]. The majority of frog promoters lack the CpG-island structure that is typical of mice and humans [[38]].

Among the 177 miRNA genes, 135 of them are located in the intergenic regions.

Table 8. Distances of frog miRNA genes measured from genes and CpG-islands.

Distances of neighboring genes and CpG-islands from frog miRNA genes	Number of miRNA genes
gene [far] and CpG-island [internal]	2
gene [far] and CpG-island [near]	9
gene [near (upstream)] and CpG-island [internal]	0
gene [near (upstream)] and CpG-island [near]	0
gene [near (downstream)] and CpG-island [internal]	0
gene [near (downstream)] and CpG-island [near]	0
gene [internal] and CpG-island [internal]	0
gene [internal] and CpG-island [near]	0
Total	11

3.2 Statistical test of the relative distance between CpG-islands and miRNA genes

Table 9 shows the results of the probability of the relative distance between a miRNA gene and CpG-island is less than 1000 bp, i.e. $prob(D < 1000)$ for human, mouse, rat and chicken. Since the CpG-islands data for frog is currently not available, therefore, its relative distance calculation is not reported. From Table 9, the small *p-value* result indicated that it did not support the uniform distance distribution model, so there could be some spatial association between a miRNA gene and a CpG-island.

Table 9. Results of the probability of relative distance between miRNA and CpG-island is less than 1000 bp for human, mouse, rat and chicken. Columns *L*, *n*, 0.1 , and *p-value* denote the size of chromosome one, total number of CpG-islands on chromosome one, 10% of known miRNA genes respectively.

Species (genome assembly)	L (Mbp)	n	0.1	p-value
Human (Build 36.3)	247	27139	46	$8.8*10^{-6}$
Mouse (Build 37.1)	197.1	11189	37	$1.2*10^{-5}$
Rat (RGSC v3.4)	267.9	21530	23	$6.1*10^{-6}$
Chicken (Build 2.1)	201	11133	149	$1.0*10^{-5}$

A web site has been set up for the public to view the data, in which the human, mouse, rat, chicken and frog CpG-associated or promoter-associated miRNA genes results are available. Genetic location information, such as the exonic, intronic and intergenic locations information, of the miRNA genes are also provided. The URL address of the web site is http://bioinfo.csie.nfu.edu.tw/960412/html/dna.php.

4 Conclusion

A portion of vertebrates (human, mouse, rat, chicken and frog) miRNA genes are located near CpG-islands. It was found that 74 of 462 (16.0%), 37 of 373 (9.9%), 25 of 234 (10.7%), 25 of 149 (16.8%) and 11 of 177 (6.2%) of the miRNA genes are located within 1000 bp of a CpG-island for human, mouse, rat, chicken and frog respectively. A statistical test is proposed to describe the probability of the relative distance distribution of a miRNA gene and a CpG-island less than 1000 bp, the results suggested that there could be some spatial association between them. Furthermore, in the human and mouse samples, the Fisher's exact test indicated that promoter-related miRNA genes are CpG-related as well. These findings suggest a potential relationship between those miRNAs and the CpG-island methylation events, which deserve further *in-vitro* investigation.

Acknowledgement

Drs. Ka-Lok Ng and Chien-Hung Huang work are supported by the National Science Council of R.O.C. under the grants of NSC98-2221-E-468-013 and. NSC98-2221-E-150-062 respectively. Our gratitude goes to Dr. Timothy Williams, Asia University, for his help on proof reading the manuscript.

References:

[1] Bartel D.P., MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116, (2004), 281–297.

[2] Bernstein E., Caudy A.A., Hammond S.M., and Hannon G.J., Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, (2001), 295–296.

[3] Lagos-Quintana M., Rauhut R., Lendeckel W., and Tuschl T. Identification of novel genes coding for small expressed RNAs. *Science* 294, (2001),853–858.

[4] Ambros V., Control of developmental timing in *Caenorhabditis elegans*. *Curr. Opin. Genet. Dev*. 10, (2000), 428–433.

[5] Carrington J.C., and Ambros V. Role of microRNAs in plant and animal development. *Science* 301, (2003),336–338.

[6] Lagos-Quintana M., Rauhut R., Yalcin A., Meyer J., Lendeckel W., and Tuschl T., Identification of tissue-specific microRNAs from mouse. *Curr. Biol.* 12, (2002) ,735–739.

[7] Lee Y. et al., The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, (2003),415–419.

[8] Lagos-Quintana M., Rauhut R., Meyer J., Borkhardt A., and Tuschl T., New microRNAs from mouse and human. *RNA* 9, (2003),175–179.

[9] Inaoka H., Fukuoka Y., and Kohane I.S.. Lower expression of genes near microRNA in C. elegans germline. *BMC Bioinformatics*. 7, (2006),112.

[10] Inaoka H., Fukuoka Y., and Kohane I.S., Evidence of spatially bound gene regulation in Mus musculus: decreased gene expression proximal to microRNA genomic location. *Proc Natl Acad Sci*.104(12), (2007),5020-5.

[11] Sevignani C., Calin G.A., Nnadi S.C., Shimizu M., Davuluri R.V., Hyslop T., Demant P., Croce C.M., and Siracusa L.D., MicroRNA genes are frequently located near mouse cancer susceptibility loci. *Proc Natl Acad Sci.*, 104(19), (2007),8017-22.

[12] Saito Y., Liang G., Egger G., Friedman J.M., Chuang J.C., Coetzee G.A., and Jones P.A., Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell.* 9(6), (2006),435-43.

[13] Kozaki K., Imoto I., Mogi S., Omura K., and Inazawa J., Exploration of tumor-suppressive microRNAs silenced by DNA hypermethylation in oral cancer. *Cancer Res.* 68(7), (2008),2094-105.

[14] Lujambio A., Ropero S., Ballestar E., Fraga M.F., Cerrato C., Setién F., Casado S., Suarez-Gauthier A., Sanchez-Cespedes M., Git A., Spiteri I., Das P.P., Caldas C., Miska E., and Esteller M., Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res*. 67(4), (2007),1424-9

[15] Lujambio A. and Esteller M., CpG island hypermethylation of tumor suppressor microRNAs in human cancer. *Cell Cycle* 6(12), (2007),1455-9.

[16] Gardiner-Garden M., and Frommer M., CpG-islands in vertebrate genomes. *J Mol Biol*. 196(2), (1987), 261-82.

[17] Beier V., Mund C., and Hoheisel J.D., Monitoring methylation changes in cancer. *Adv Biochem Eng Biotechnol*., 104, (2007),1-11.

[18] Yamashita R., Suzuki, Y. Sugano S., and Nakai K., Genome-wide analysis reveals strong correlation between CpG-islands with nearby transcription start sites of genes and

their tissue specificity. *Gene* 350(2), (2005),129-36.

[19] Yamashita R., Suzuki Y., Wakaguri H., Tsuritani K., Nakai K., and Sugano S., DBTSS: DataBase of Human Transcription Start Sites, progress report 2006. *Nucleic Acids Res.* 34 (Database Issue), (2006),D86-D89.

[20] Karolchik D. et al., The UCSC Genome Browser Database. *Nucleic Acids Res*. 31(1), (2003),51–4.

[21] Kent W.J., Sugnet C.W., Furey T.S., Roskin K.M., Pringle T.H., Zahler A.M. and Haussler D., The human genome browser at UCSC. *Genome Res*. 12(6), (2002),996–1006.

[22] Altschul S.F., Gish W., Miller W., Myers E.W., and Lipman D.J., Basic local alignment search tool. *J Mol Biol.* 215(3), (1990),403-10.

[23] Kawasaki H., and Taira K., Transcriptional gene silencing by short interfering RNAs. *Curr. Opin. Mol. Ther.* 7(2), (2005),125–31.

[24] Mathieu O., and Bender J., RNA-directed DNA methylation. *J Cell Sci.* 117, (2004),4881-8.

[25] Ambros V., et al., A uniform system for microRNA annotation. *RNA* 9, (2003), 277–279.

[26] Griffiths-Jones S., The MicroRNA Registry. *Nucleic Acids Res.* 32 (Database Issue), (2004), D109–D111.

[27] Griffiths-Jones S., Grocock R.J., van Dongen S., Bateman A., and Enright A.J., miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* 34 (Database Issue), (2006) , D140-D144.

[28] Tamhane A. and D. Dunlop, *Statistics and Data Analysis: From Elementary to Intermediate*, Prentice Hall, NJ, USA, (1999).

[29] Holstebroe S. and Tommerup N., *A subset of miRNA genes are associated with CpG-islands*. European Human Genetics Conference 2005 (ESHG 2005) Prague, (2005).

[30] Jiang C., Han L., Su B., Li W.H., and Zhao Z., Features and trend of loss of promoter-associated CpG-islands in the human and mouse genomes. *Mol Biol Evol*. 24(9), (2007), 1991-2000

[31] Place R.F., Li L.C., Pookot D., Noonan E.J., and Dahiya R., MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proc Natl Acad Sci*. 105(5), (2008),1608-13.

[32] Xuan Z., Zhao F., Wang J., Chen G., and Zhang M.Q., Genome-wide promoter extraction and analysis in human, mouse, and rat. *Genome Biol.* 6(8), (2005), R72

[33] Gao F. and Zhang C.T., Isochore structures in the chicken genome. *FEBS J*. 273(8), (2006), 1637-48.

[34] ICGSC, International Chicken Genome Sequencing Consortium, Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432, (2004),695-716.

[35] Axelsson E., Webster M.T., Smith N.G., Burt D.W., and Ellegren H., Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. *Genome Res.* 15(1), (2005),120-5.

[36] McQueen H.A., Fantes J., Cross S.H., Clark V.H., Archibald A.L., and Bird A.P., CpG-islands of chicken are concentrated on microchromosomes. *Nat Genet.* 12(3), (1996),321-4.

[37] McQueen H.A., Siriaco G., and Bird A.P., Chicken microchromosomes are hyperacetylated, early replicating, and gene rich. *Genome Res*. 8(6), (1998), 621-30.

[38] Stancheva I., El-Maarri O., Walter J., Niveleau A., and Meehan R.R., DNA methylation at promoter regions regulates the timing of gene activation in Xenopus laevis embryos. *Dev Biol*. 243(1), (2002),155-65.