

CHAPTER 2: LITERATURE REVIEW

2.1 Types of Ribonucleic Acid (RNA)

In cells, there are vast amount of RNA which can be classified as ribosomal RNA (rRNA), messenger RNA (mRNA), transfer RNA (tRNA), heterogeneous nuclear RNA (hnRNA) and other small non-coding RNAs. Ribosomal RNA is a component of ribosomes (a protein synthesis factory in cell); mRNA comprises RNA sequences that are transcribed from DNA and contain codons to be translated into amino acid residues; tRNA is a small RNA that carries anticodon and amino acid residue for protein synthesis; hnRNA comprises RNA molecules of various sizes found in the nucleus (Russell, 2006). Small non-coding RNAs do not encode proteins but have regulatory functions in biological systems. In bacteria, they are designated as small RNAs (sRNAs) with biological functions including regulation of intrinsic catalytic activity, protein synthesis, mRNA transcription, translation and stability in bacteria (Massé *et al.*, 2003; Tjaden *et al.*, 2006). In eukaryotes, they are denoted as non-coding RNAs (ncRNAs) and are widely available in the nucleus and cytoplasm. MicroRNAs have many similarities with short interference RNAs. The length of miRNAs and siRNAs are also the same. Both are processed by Dicer-like1 enzyme (Lim *et al.*, 2003^b) and therefore possess protruding ends. The functioning single strand of small RNA will be recognized by RISC protein complex and directs this to the mRNA target for further process. Mature miRNAs are 20-24 nucleotides in length (Lim *et al.*, 2003^b; Bartel, 2004). According to Ghosh *et al.* (2007), the miRNA precursors have a well-predicted stable extended stem loop hairpin structure with continuous helical pairing and a few internal bulges. Ambros *et al.* (2003^a) mentioned that the animal miRNAs precursors are about 50-

60 nucleotides and they observed that different animal miRNAs are often grouped within a single RNA precursor. Both sense and antisense miRNA genes reside in the stem loop precursor (derived from miRBase, 2008). Therefore, miRNA genes can derive from a single hairpin-look endogenous transcript. In addition, Lim *et al.* (2003^b) reviewed that both mature miRNA and the miRNA precursor are evolutionarily conserved across species. MicroRNA genes can be found in the genome as intergenic and intronic miRNAs (Lee *et al.*, 1993; Wightman *et al.*, 1993; Reinhart *et al.*, 2000; Ambros, 2003^b; Lin *et al.*, 2006; Tang and Maxwell, 2008). A single miRNA can target at different mRNAs of genes and at the same time, more than one miRNA gene can target at the same protein coding transcripts (Lim *et al.*, 2003^b; Kim, 2005; Giraldez *et al.*, 2006). Another 21-23 nucleotide RNA duplexes anchored in RNA silencing is known as short interference RNA (siRNA) (McManus *et al.*, 2002; Reddy *et al.*, 2007; Yang and Mattes, 2008). siRNA can be synthesized and introduced to cell using vectors. Only one strand of the siRNAs will serve as the “guide” RNA to target perfectly with mRNA of a gene and this RNAi process is taken in the cytoplasm (Yang and Mattes, 2008). Elbashir *et al.* (2001) discovered 21-nucleotides siRNAs could direct gene silencing through mRNA cleavage and degradation and specifically suppressed the heterogeneous and endogenous genes in different mammalian cells.

The difference between miRNA and siRNA is: miRNA is processed from hairpin-like structure where its primary transcript is transcribed from one strand of the genome locus (Kim, 2005; Lu *et al.* 2005; Carthew and Sontheimer, 2009) whereas siRNA is formed from the Dicer cleavage of longer double stranded RNAs (Kim, 2005; Carthew and Sontheimer, 2009).

2.2 Biogenesis of MicroRNA Genes

Like cell transcriptional process, the DNA containing miRNAs sequences can reside within intergenic or intronic regions of coding sequence, untranslated region or exonic regions of non-coding sequence. These are transcribed into long miRNA primary transcripts (pri-miRNAs) by RNA polymerase II (Ambros, 2003^a; Kim, 2005). The animal pri-miRNA can reach from the length of 1kb up to 4kb (Megraw *et al.*, 2006; Sethupathy and Collins, 2008). The pri-miRNA contains more than one stem-loop structures. Then, the hairpin-like structure in the pri-miRNA will be cleaved by RNase-III-type enzyme Dicer and nuclear microprocessor complex (Drosha and its cofactor) to form a stem-loop structure miRNA precursor (pre-miRNA; Yang and Mattes, 2008). For human, the microprocessor complex is consisted of Drosha and DiGeorge syndrome critical region gene 8 protein (DGCR8) (Kim, 2005; Lee *et al.*, 2006) whereas in *C. elegans* and *D. melanogaster*, Drosha and Pasha are the microprocessor (Gregory *et al.*, 2006; Lee *et al.*, 2006). Drosha contains two tandem RNase III domains (RIIIDs) and a double-stranded RNA binding domain (dsRBD) (Kim, 2005; Pillai, 2005). Both DGCR8 and Pasha contain two dsRBDs which help Drosha to recognize substrate (Kim, 2005). The length of miRNA precursor is diverse in different eukaryotes. For example, Zhang *et al.* (2006) reviewed that plant miRNA precursors can vary from 60 up to 400 nucleotides in length. For animals, most scientists found that miRNA precursors of 60-120 nucleotides (Ambros *et al.*, 2003^a; Lim *et al.*, 2003^b; Tang and Maxwell, 2008). Kim (2005) reviewed that the length of miRNA precursor was usually 60-80 nucleotides. Precursor miRNA will be transported out from the nucleus to the cytoplasm by exportin-5 and Ran-GTP (Bartel, 2004; Pillai, 2005). The pre-miRNA export is interceded by exportin-5 (a Ran-dependent importin- β -related transport receptor) because the pre-miRNA formed a minihelix-like structure (Kim, 2004). Gwizdek *et al.* (2003) found

that exportin-5 mediated the export of minihelix-like structure RNA in adenovirus VA1 mutant. Following that, Lund *et al.* (2004) found that exportin-5 was the candidate that for transporting pre-miRNA. When they employed the RNAi pathway to silence the exportin-5, the miRNA levels in *Xenopus* oocyte nucleic were reduced. Once in the cytoplasm, other RNase III endonuclease, Dicer will cleave precursor to produce the miRNA: miRNA* duplexes (Bartel, 2004; Kim, 2005; Hammond, 2006; Yang and Mattes, 2008). The cytoplasmic dicer protein belongs to RNase III nuclease family and catalyses cleaving of double stranded RNAs (Grishok *et al.*, 2001; Kim, 2005). It has been found as different homologs in various organisms ranging from plant, fungi, worms, flies and mammals (Bernstein *et al.*, 2001). The miRNA duplexes do not stay long in the cell. Next, the helicase will unwind the duplexes to release the mature miRNA and antisense miRNA (Yang and Mattes, 2008). During this time, RNA induced silencing complex (RISC) composed of Dicer, TRBP, and the argonaute-2 protein will bind to the mature miRNA (Yang and Mattes, 2008). Following that, the complexes will complement at the target side of mRNA to induce post translational gene silencing. In animals, several studies have found that animal miRNA genes complemented imperfectly to mRNA targets and mediated translation inhibition. The strongest base-pairing between the miRNA and mRNA was found to happen at especially the first 8 or 9 nucleotides of 5' end of the miRNA (Nilsen, 2005; Palkodeti *et al.*, 2006; Norden-Krichmar *et al.*, 2007). Therefore, location at 2-7 nucleotides or 2-8 nucleotides of 5' miRNA sequences was known as the well conserved “seed” or “nucleus” regions (Palkodeti *et al.*, 2006; Tang and Maxwell, 2008).

2.3 miRNA-mediated Gene Regulation

MicroRNAs regulate temporal and tissue specific development, apoptosis, cell proliferation, neuronal patterning, metabolic regulation, diseases and viral infections via transcriptional and posttranscriptional gene silencing (Reinhart *et al.*, 2002; Wienholds and Plasterk, 2005). Animal miRNAs are often found to inhibit translational processes due the imperfect matches. This post translational activity differs for plant miRNA which often binds with perfectly complementarity to their target mRNAs and induces alteration of translation (Kloosterman *et al.*, 2006; Lindow and Gorodkin, 2007). Generally, miRNA can function via 3 different mechanisms: by cleavage of the target mRNA, by blocking the process of translation of the encoded protein and transcriptional silencing by DNA modification.

MicroRNAs can control gene expression by directing the endonuclease cleavage of target mRNA. This activity resembles “gene silencing” or RNA interference pathway. There are 2 major mechanisms for mRNA degradation in eukaryotic cells (Parker and Song, 2004; Valencia-Sanchez *et al.*, 2006). One starts with the removal of 3’ poly(A) tails (deadenylation), then, 3’to 5’ exonuclease activity occurs at the 3’ end of mRNA (Parker and Song, 2004). The other mechanism, also starts with deadenylation, then, mRNA will be decapped by Dcp1/Dcp2 enzymes, followed by 5’to 3’ exonucleolytic decay by Xrn1p exoribunuclease (Sheth and Parker, 2003; Wang *et al.*, 2002). miR-196 was found to target complementary to 3’ UTR of *HOXB8* mRNA, which encoded transcription factor in developmental regulation; and subsequently mediated *HOXB8* mRNA cleavage in mouse embryo (Yekta *et al.*, 2004). However, for cis-regulatory elements at 3’ UTR such as AU rich elements, specific regulatory binding proteins and others were found to work closely with miRNA in enhancing mRNA degradation (Valencia-Sanchez *et al.*, 2006).

Earlier studies on *C. elegans*, protein synthesis repression by miRNA were observed where the number of targeted mRNA remained abundant but the protein product of the targeted gene was remarkably reduced. miRNA *lin-4* targeted partially to multiple sites of 3' UTR of *lin-14* mRNA and *lin-28* mRNA and subsequently inhibited the *lin-14* and *lin-28* protein synthesis which regulated the developmental timing of embryonic events (Lee *et al.*, 1993; Wightman *et al.*, 1993). Animal miRNAs have partial complementarity (usually seed region of 5' miRNA: nucleotide 2-8) to the mRNA target and only lead to a halt to translation without producing any gene product (Kloosterman *et al.*, 2006; Gu *et al.*, 2007; Lindow and Gorodkin, 2007). Humphreys *et al.* (2005) found that *CXCR4* miRNA repressed the translation initiation of *Renilla*-luciferase mRNA (contained 4 imperfect binding sites for *CXCR4* miRNA at 3' UTR) by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function when they introduced *CXCR4* miRNA into HeLa cells.

The third potential mechanism for miRNA-mediated gene regulation is transcriptional silencing via DNA modification. In nucleus, RNA can be base paired to DNA sequence and induce RNAi pathway to affect gene function at genomic DNA level such as RNA-directed DNA methylation and RNAi-directed heterochromatin modification (Matzke and Birchler, 2005; Muhonen and Holthofer, 2009). The precursor miRNA will be cleaved by Dicer to yield miRNA: miRNA* duplex. Only one strand (mature miRNA) will be the “guide” to bind to DNA. RNA-directed DNA methylation and RNAi-directed heterochromatin modification involved modification of cytosines in DNA or histones in an epigenetic process (Russell, 2006). Mette *et al.* (2000) reported that dsRNA with complementary sequences to a promoter region, which was processed into small RNA could trigger promoter methylation in plants. Methylation at a promoter prevents transcription from occurring and results in gene silencing. RNAi-directed heterochromatin

modification has been observed in fission yeast and *Drosophila* (Mathieu and Bender, 2004). Tandem repeats or transposon elements available in the heterochromatin can generate dsRNA by RNA-dependent RNA polymerase. The dsRNA later will be cleaved by Dicer and produce siRNA to guide histone methyltransferase to modify the lysine 9 on histone 3 in the chromatin (Matzke and Birchler, 2005). Once the modification occurs, the transcription process will be inhibited.

2.4 Discovery of miRNAs Genes and Their Roles in Various Organisms

In 1993, the discovery of a 22-nucleotide RNA, *lin-4* by both Lee *et al.* (1993) and Wightman *et al.* (1993) yielded a remarkable finding where this miRNA targeted partially to multiple site of 3' untranslated region (UTR) of *lin-14* mRNA and subsequently inhibited the expression of *lin-14* protein in *C. elegans*. Seven years later, Reinhart *et al.* (2000) discovered that a 21-nucleotide RNA known as *let-7* regulates temporally at the 3' UTR of the *lin-14*, *lin-28*, *lin-41*, *lin-42* and *daf-12* genes. They also reported that *lin-4* regulated negatively by complementary to 3' UTR of *lin-14* and *lin-28* messenger RNAs. Both *lin-4* and *let-7* are involved in the normal temporal control of *Caenorhabditis elegans* different postembryonic developmental events; they are the first members known as microRNAs (miRNAs). Wienholds and Plasterk (2005) mentioned that miRNAs were predicted to control 30% of the human genome despite most biological functions of identified miRNAs being unknown. Esau and Monia (2007) reviewed that miRNA genes have biological functions in development, metabolism, neurobiology, cancer and even in virus infections. The discovery of microRNAs (*lin-4* and *let 7*) in the nematode *Caenorhabditis elegans* (*C. elegans*) has shown a breakthrough in reverse genetic and gene regulation in biological systems (Lee *et al.*, 1993; Wightman *et al.*, 1993; Reinhart *et al.*,

2000). Subsequently, the miRNA study has enhanced the search of miRNAs in more eukaryotes (Reinhart *et al.*, 2002; Lim *et al.*, 2003^a; Legendre *et al.*, 2005; Kloosterman *et al.*, 2006). *Bantam* miRNA being expressed in temporally tissue-specific manner in *Drosophila melanogaster*, was found to bind to *hid* mRNA to stimulate cell proliferation and repress apoptosis (Brennecke *et al.*, 2003). *Lsy-6* miRNA was detected to bind to 3' UTR of *cog-1* mRNA (Nkx-type homeobox gene) to regulate the neuronal left/right asymmetry of chemosensory receptor expression in nematode (Johnston and Hobert, 2003). Yekta *et al.* (2004) discovered miR-196 mediated cleavage of transcription factor, *HOXB* gene that involved in mouse embryo developmental regulation. Poy *et al.* (2004) has discovered that miR-375 targeted to myotrophin (*Mtpn*) gene to repress the glucose-induced insulin secretion in *Mus musculus*. Lai *et al.* (2005^a) reported that the miRNA genes negatively regulated *Notch* signaling by targeting to conserved motifs at 3' UTR of *Notch* genes in *Drosophila* (miR-7 was found to bind to GY -box; miR-4 and miR-79 targeted to Brd-box; miR-2 and miR-11 bound to K-box). Esau *et al.* (2006) found the function of liver-specific miR-122 in regulating the lipid metabolism in mice where inhibition of miR-122 reduced the cholesterol synthesis rate and hepatic fatty-acid level. This could serve as a potential therapy for metabolic diseases.

Up to date, relatively few reports on miRNAs works in aquatic animals have been published, except for zebrafish (*Danio rerio*) (Chen *et al.*, 2005; Kloosterman *et al.*, 2006). Zebrafish is a model to study the development of vertebrates (Grunwald and Eisen, 2002) and cancer (Amatruda *et al.*, 2002) since the complete zebrafish genome sequence is available. Giraldez *et al.* (2005) reported the role of zebrafish microRNAs in regulating brain morphogenesis. Giraldez *et al.* (2006) found that zebrafish miR-430 is expressed during zygotic transcription and it regulates early development morphogenesis. They also

showed that zebrafish miR-430 directs maternal mRNA deadenylation and clearance during early embryogenesis. In miRBase, miRNAs in pufferfish (*Fugu rubripes* and *Tetraodon nigroviridis*) were predicted based on homology from established zebrafish sequences (miRBase, 2008). The Dicer-1 in zebrafish has also been identified and its importance in vertebrate development is discovered (Wienholds *et al.*, 2003). Su *et al.* (2008) reported the identification and functional characterization of Dicer-1 in tiger prawn (*Penaeus monodon*).

2.5 Methods of miRNA Gene Identification

Generally, there are 3 approaches that are used for the identification of miRNA genes. They are: (1) cloning and sequencing of small RNA libraries; (2) computational prediction; and (3) High throughput data analysis.

After the discoveries of *lin-4* and *let-7* in *C. elegans*, efforts in searching for miRNA genes were conducted with small RNA cloning and validation with Northern blotting. However, this method is time consuming and small numbers of novel miRNAs candidates are generated. Therefore, the number of miRNAs available initially was limited. The discovery of next generation sequencing technique developed by Illumina(Solexa), 454/Roche and Applied Biosystems, will allow, scientists to identify more novel miRNA genes. By using this method, Hafner *et al.* (2007) described that sequencing the cDNA library created from pooled small RNAs not only discovered new classes of small RNAs, but was also useful in creating miRNA expression profiles based on clone count frequencies.

Computational identification has become one of the forward approaches in identifying miRNA genes. This is most probably because the size of miRNA is quite small; and its expression is low and varies at different target tissues (Zhang *et al.*, 2006). In addition to that, it can be due to the ease availability of software and algorithms. Besides that, genomes of various organisms have been completed and can be easily accessed. Tang and Maxwell (2008), using homology search approach against *Xenopus* genome, have discovered 142 miRNA genes in *Xenopus tropicalis* residing in the intergenic and intronic region of coding genes and exonic region of non-coding genes. Zhou *et al.* (2008) reported their finding of 357 miRNA candidates from dog genome, where 300 miRNAs were found homologous to human miRNAs whereas 57 candidates were known as novel miRNAs. Since miRNAs are said to be evolutionarily conserved, many homologous mature miRNAs have been predicted and identified from different species in both plants and animals (Lagos-quintana *et al.*, 2003). For examples, Reinhart *et al.* (2002) reported that eight Arabidopsis miRNAs showed identical matches in the genome of rice (*Oryza sativa* L. ssp. *indica*). Lagos-quintana *et al.* (2003) found that miRNAs of *Fugu rubripes* (pufferfish) and *Danio rerio* (zebrafish) each showed 53% identity to all known human and mouse miRNAs whereas *Caenorhabditis elegans* or *Drosophila melanogaster* each shared almost 50% identity to all identified human and mouse miRNAs. There are several programs designed for miRNA gene finding using different algorithms. For example, miRAlign (Wang *et al.*, 2005) based on the principles of conserved sequences and secondary structure helps to identify miRNA genes in animals; microHARVESTER (Dezulian *et al.*, 2005) with high sensitivity and specificity, helps to find miRNA candidates from complete plant genome sequences based on an miRNA query. However, the limitation of homology search is no novel miRNA genes will be identified.

The next generation sequencing (NGS) is referred to a developed technology that is capable of generating large scale of DNA sequence reads (up to million base pairs) in a single run (Holt and Jones, 2008; Mardis, 2008). The major companies involved in the NGS technology is Illumina/Solexa, 454/Roche, Applied Biosystems and Helicos. All NGS companies have developed different chemistry and approaches in the high throughput sequencing technology. Illumina/Solexa applied synthesis by sequencing technique via reversible terminator-based sequencing chemistry (Illumina, Inc., USA); 454/Roche applied pyrosequencing method through detection of light emitted from pyrophosphates for every incorporated nucleotide (454 Life Sciences, a Roche company, USA); Applied Biosystems incorporates emulsion PCR and followed by sequencing by ligation method (Applied Biosystems, Inc., USA); and for Helicos, employs single molecule sequencing technique by incorporating one type of nucleotides for each cycle (Helicos BioSciences Corporation, UK).

The development of next generation sequencing for example, the high throughput sequencing technique has enhanced the technology of high throughput data analysis. This method enables researchers to identify and validate the complete genome of species of interest. Small RNA library sequencing not only allows the discovery of known and novel miRNA genes, it also helps in quantifying miRNA expression and provides the complete profile of miRNA genes available in the organism (Hafner *et al.*, 2008). The complete miRNA profile can be then produced as miRNA array for mRNA target validation and can also served as potential biomarkers in screening diseases. Galzov *et al.* (2008) has applied sequencing by synthesis approach developed by Illumina (Solexa) in his study on chicken miRNA genes. His group not only has identified most of previously known chicken

miRNA genes and miRNA* sequences, but also identified 361 new miRNA genes and 88 new miRNA candidates from different developmental stages of chicken embryos. Morin *et al.* (2008) who employed the Illumina/Solexa technology on the human embryonic stem cells and embryoid bodies, managed to identify 334 known miRNA genes and 104 novel miRNA genes. Moreover, their group was also able to observe 171 known and 23 new miRNA genes showing significant expression variation between human embryonic stem cells and embryonic bodies. Although next generation sequencing technique provides a good tool for miRNA studies, the cost for the system is very expensive. In addition to that, bioinformatics pipeline is an important criterion that has to be taken into consideration to analyze the mass data input.

2.6 miRNA: mRNA Target Identification

Many reported that nucleotide 1-7 or nucleotide 2-8 of 5' end of predicted mature miRNA sequence served as the seed region (Lewis *et al.*, 2005; Gu *et al.*, 2007; Lindow and Gorodkin, 2007; Norden-Krichmar *et al.*, 2007). Most validated miRNA: mRNA targets revealed that this seed region would target at the 3' UTR of protein-coding sequences (Nilsen, 2005; Palakodeti *et al.*, 2006; Norden-Krichmar *et al.*, 2007; deduced from miRBase Targets, 2009). miRNA: mRNA target can be identified via computational approaches using reverse complement BLAST search. Besides that, several programs based on different algorithms such as TargetScan (Friedman *et al.*, 2009), miRDB (Wang, 2008), miRBase Targets (Griffiths-Jones *et al.*, 2006), PicTar (Grün *et al.*, 2005) and MiRscan (Lim *et al.*, 2003^a) have been constructed to ease the search of miRNA: mRNA target in vertebrate, *Drosophila* species and nematoda. However, a serious obstacle for miRNA: mRNA study is the experimental identification of miRNA targets. Up to now, miRNA

targets have been experimentally investigated based on genetic analysis of *C. elegans* and *Drosophila*. Different approaches have been conducted to validate the miRNA: mRNA target through experiments. Lim *et al.* (2005) have utilized microarray approach to identify miRNA-mediated degradation of target mRNAs. Nakamoto *et al.* (2005) applied siRNA technology to knockdown miR-30-3p in a human cell line HepG2 to detect the mRNA targets from polyribosome profiles. Vinther *et al.* (2006) utilized stable isotope labeling by amino acids to culture miR-1 transfected HeLa cell, followed by investigating the repression of protein synthesis using mass spectrophotometer. By 2007, Ørom and Lund managed to come out with another experimental protocol for isolating miRNA target using biotinylated synthetic miRNAs.

2.7 Potential Application of miRNA Genes – RNA Interference Pathway

Fire *et al.* (1998) discovered that injection of double stranded RNA (dsRNA) into *Caenorhabditis elegans* silenced specific genes from expressing proteins. This technology later was named as RNA interference (RNAi). The dsRNA can be present in cells due to the “Dicer-like1” post processing of miRNA hairpin structures or dsRNA virus infection or dsRNA from the genome (Yang and Mattes, 2008). RNAi is considered as a eukaryotic self-defense mechanism to restrict the pathogen such as virus from invading the cells; and also control post transcriptional gene silencing in cells (Yang and Mattes, 2008).

miRNAs serve as the regulators in gene activation and suppression. Once miRNA genes and their targets have been identified, we will be able to control the condition of an organism by manipulating the mechanism of miRNA-mediated gene regulation. It can be done through over expression of certain miRNA gene, inhibition of miRNA by antagomir

or siRNA injection to activate or silence the targeted mRNA. Esau *et al.* (2006) found the level of hepatic fatty acid and cholesterol synthesis rate have reduced in normal and fat mice when they inhibited miR-122 with antagomirs. This suggested that miRNA inhibition could be used as therapeutic treatment of a disease. Copf *et al.* (2006) has shown the use of RNAi in knocking down the *spalt* function to cause derepression of *Hox* genes and homeotic transformations in the crustacean. Other study including induction of antiviral immunity by double stranded RNA in the marine invertebrate *Litopenaeus vannamei* (white shrimp) has also been conducted and positive results indicated that the antiviral response within this invertebrate shared some features of vertebrate antiviral mechanisms (Robalino *et al.*, 2004). Lua *et al.* (2008) reported their finding that short interference RNA could be used to selectively block red seabream iridovirus gene expression and replication in a cell culture system.

Up to date, about 706 miRNA genes have been discovered in human (deduced from miRBase). Studies were furthered into the genetic variants of miRNA sequences especially single nucleotide polymorphisms (SNPs) and how its regulation is associated with human disorders, cancer and tumor diseases (Hu *et al.*, 2008). In addition to that, due to the complete human genome and genome-wide association studies (GWAS) in disease-risk variants, recent research showed that miRNA can be used as a biomarker for various disease and cancer detection (Chen *et al.*, 2008; Gilad *et al.*, 2008; Sethupathy and Collins, 2008).