MicroRNAs and esophageal cancer

Santosh Kumar Patnaik1, Reema Mallick2, Sai Yendamuri1,3

1Department of Thoracic Surgery, Roswell Park Cancer Institute, Buffalo, New York, USA; 2Northeastern Ohio Universities College of Medicine, Rootstown, Ohio, USA; 3Department of Surgery, State University of New York at Buffalo, Buffalo, New York, USA

ABSTRACT

Cancer of the esophagus is a highly aggressive disease associated with an overall poor prognosis. There is an insistent need for improving our understanding of the molecular basis of this disease. The recent emergence of observations on the role of microRNAs in cancer and their potential as biomarkers has prompted many investigations to examine their relevance to esophageal cancer. This article provides an introduction to microRNA biology and the techniques involved in studying them, and summates what is now known about their role and utility in regard to neoplastic esophageal diseases.

KEY WORDS

Barrett’s esophagus, epigenetic mechanism, esophageal cancer, gene expression, microRNA

Introduction

Esophageal cancer is the eighth most incident and the sixth most fatal cancer worldwide (1). Histologically, in approximately 95% of cases, the disease occurs as esophageal squamous cell carcinoma (ESCC) or esophageal adenocarcinoma (EAC), with the latter being about 50% more prevalent in the United States but at least 20-times less common in Asian countries (1, 2). The two forms are associated with different etiology, epidemiology, clinical course, and responsiveness to treatment. Among Caucasian men in the United States, while the annual incidence rate of ESCC has been slowly declining, that of EAC increased four- to six-fold to 3.2-4.0/100,000 during the last two decades of the 20th century, presumably due to increased prevalence of gastro-esophageal reflux disease resulting from obesity (3). The overall outcome of patients with esophageal cancer is dismal, with a five-year survival rate of 8%-30% following treatment with only surgery, and with no clearly significant benefit from adjuvant therapies (4). The disease also poses challenges in early diagnosis, staging, prognosis, and selection and delivery of the optimal therapy. Another challenge is that of screening and follow-up of patients with gastro-esophageal reflux disease with or without Barrett’s esophagus (BE), a small proportion of whom develop EAC, and for whom the only reliable method of follow-up requires endoscopy. There is clearly a need for better understanding of the molecular characteristics of the disease for the development of clinically useful biomarkers and treatment modalities. Biomarkers based on changes in genomic DNA, and in expression of specific mRNA or protein molecules, or of metabolites have been explored for many years now (5). MicroRNAs have emerged as a new class of biomolecules with important roles in cellular functions in both health and disease, and have considerable potential as biomarkers. These epigenetic players are ultrashort non-protein-coding RNA molecules which regulate the expression of a large number of proteins by interacting with the coding mRNAs. This review introduces their biology and common approaches to their study, and summizes the current state of knowledge on their involvement in esophageal cancer.

Biogenesis and mechanism of action of microRNAs

The synthesis, processing and action of microRNAs is simplistically depicted in figure 1. Functionally active microRNAs, or mature microRNAs, are 18-22 nucleotide-long, single-stranded RNA molecules with 5’ phosphate
and 3' hydroxyl groups. A nascent mature microRNA, however, arises in pair as a double-stranded RNA molecule known as a microRNA/microRNA-star (\textsuperscript{*}) duplex from a single precursor RNA (pre-microRNA). Pre-microRNAs are ~60-80 nucleotide-long with a hairpin-like stem-loop secondary structure. Endoribonuclease activity of a cytoplasmic RNase III enzyme, Drosha, causes the release of the microRNA/microRNA* duplex-bearing stems from the stem-loop structures of pre-microRNAs. Pre-microRNAs themselves are generated in the nucleus by the action of another RNase III endoribonuclease, Drosha, on much longer, primary RNA molecules (pri-microRNAs) that are transcribed by RNA polymerases II and III from microRNA-encoding genes (6, 7). Two nucleotide-long 3' overhangs on pre-microRNAs are recognized by the Exportin 5 transporter protein which shuttles them into the cytoplasm (8). Many other proteins are involved in this pathway for microRNA genesis. They include the Ran guanosine triphosphatase, which participates in the nuclear export of pre-microRNAs, and the double-stranded RNA-binding proteins DGCR8 (DiGeorge critical region 8) and TRBP (transactivating response RNA binding protein), which work alongside Drosha and Dicer, respectively. Though most microRNAs arise in this framework, exceptions have been observed. For instance, maturation of microRNA \textit{miR-451} does not require the Dicer-mediated cleavage (9), and the precursor of microRNA \textit{miR-1234} is actually an intron (a 'mirtron') that is spliced out of the mRNA of a protein-coding gene (10). The sequences of mature microRNAs can get modified through 3' uridylation or adenylation, or nucleotide substitution, with possible effects on their turnover as well as function (11).

Mature microRNAs actuate their function through the multi-protein RNA-induced silencing complex (RISC) that is also responsible for the phenomenon of RNA interference caused by small interfering RNAs (siRNAs). MicroRNAs are loaded as microRNA/microRNA* duplexes on RISC complexes where they are unwound into two single-stranded, mature microRNAs (figure 1). One of the strands becomes the 'guide' strand and is retained, whereas the other, the 'passenger' strand, is degraded. The selection of the guide strand is not random and is biased by lowered thermodynamic stability at the 5' end and other sequence-specific features of the strands (12, 13). The Argonaute family of proteins (Ago 1-4 in humans), key components of the RISC complex, participate in this strand-selection process. RISC complexes are guided to target mRNA molecules by the mature microRNA that is retained as the guide strand to degrade them or to inhibit their translation through mechanisms such as endonucleolytic cleavage and premature dissociation of ribosomes (14). It should be noted that mature microRNAs can be detected within the nucleus as well (15), and their specific roles include directly, and either positively or negatively affecting gene transcription have been documented (16, 17).

The targeting of mRNAs by microRNAs requires only partial sequence complementarity between the microRNA and the apposite microRNA-target site in the mRNA, which can be in either the coding or the untranslated region of the mRNA. A mature microRNA can thus target hundreds of different mRNAs, and the same mRNA can be targeted by scores of different microRNAs. A majority of microRNA-target sites show perfect sequence complementarity with the 'seed' sequence (nucleotide positions 2-7) of the mature microRNAs targeting them (18). Imperfect complementarity for the seed sequence can, however, be compensated by enhanced base-pairing at the 3' end of the microRNA (19). Target sites lacking both perfect seed pairing and 3' compensatory pairing but depending on Watson-Crick base-pairing with the central 11-12 nucleotides of microRNAs have also been identified (20). Bioinformatic algorithms such as miRanda and PicTar that consider such factors to predict mRNA targets of individual microRNAs exist, though their accuracies are not high (21). Biochemical techniques relying on co-immunoprecipitation of target RNA with proteins associated with the RISC complex have been developed to identify microRNA-targeted mRNAs (22, 23). Experimental verification of individual microRNA targets typically involves correlating changes in mRNA and protein levels with changes in the level of the targeting microRNA. Reporter mRNAs, such as those encoding for fluorescent or luminescent proteins, engineered to bear microRNA-target sites are also often used in such studies. Biologically, the degree to which mRNA transcripts are targeted by microRNAs, and the effect of such targeting depends on factors like the amounts of the microRNAs and the mRNAs (24), and the number of microRNA-target sites on the mRNAs. As has been shown for the \textit{PTEN} and \textit{KRAS} genes (25), the targeting by microRNAs can also be diluted by mRNAs which bear target-sites for the same microRNAs and thus act as decoys. Sequence polymorphisms or mutations in microRNAs or microRNA-target sites can enhance or diminish mRNA targeting (26). Epidemiological studies have shown correlations between such polymorphisms and the nature of various diseases in human populations (27, 28).

\textbf{Nomenclature, isolation, and detection of microRNAs}

As of July 2010, 940 mature microRNAs in humans, and more than 14,000 mature microRNAs in 132 other organisms, including viruses, protozoa, flies, and plants, had been identified. Novel microRNAs are generally discovered through high-throughput, direct sequencing of
Figure 1. Synthesis and action of microRNAs. Mature microRNAs are generated as a pair of partially complementary RNA molecules, microRNA/microRNA* duplexes, from precursor molecules called pre-microRNAs by the endoribonuclease activity of Dicer in the cytoplasm. Pre-microRNAs are single-stranded RNA molecules with double-stranded secondary structures containing the mature microRNA sequences. They are produced in the nucleus by the action of another ribonuclease, Drosha, on primary microRNA transcripts, and are exported to the cytoplasm by the Exportin 5 transporter. The targeting of protein-encoding mRNA transcripts by microRNAs requires their loading on to multi-protein RISC complexes. After RISC loading, one of the two RNA strands of the microRNA/microRNA* duplex is degraded, and the remaining one guides RISC for mRNA targeting. A single microRNA can target different mRNAs, and a single mRNA can bear one or more targeting sites for multiple microRNAs.
RNA molecules isolated from biological specimens, through cloning them for sequencing, or through bioinformatic prediction following analyses of genomic sequences (29, 30). The miRBase microRNA registry is responsible for assigning names to microRNAs (31). MicroRNA names have a numerical component, and a prefix indicating the taxonomic species of origin is often added to them. Thus, have a numerical component, and a prefix indicating the assigning names to microRNAs (31). MicroRNA names are also often qualified by appending '3p' or '5p' to indicate the strand of the hairpin stem from which they arise. Sometimes an asterisk (*) is appended to indicate that that microRNA usually becomes the passenger strand and is thus less abundant than the sister mature microRNA. Thus, microRNAs with identical or very similar sequences, and therefore considered members of the same microRNA family, but which are products of different genes are distinguished by suffixes. Examples are miR-16-1 and miR-16-2, and miR-200a and miR-200b. Because the same pre-microRNA can generate two different mature microRNAs, microRNA names are also often qualified by appending '3p' or '5p' to indicate the strand of the hairpin stem from which they arise. Sometimes an asterisk (*) is appended to indicate that that microRNA usually becomes the passenger strand and is thus less abundant than the sister mature microRNA. Thus, miR-200b-3p may also be referred to miR-200b*. Because of historical reasons, most family members of the orthologs of the let-7 microRNA of the nematode worm Caenorhabditis elegans, such as hsa-let-7b and mmu-let-7e, do not contain 'miR' in their designations.

In mammals, microRNA loci are present on all but the Y chromosome (e.g., 32). A significant number of microRNA loci occur in clusters, with consistent expression observable among the members for a majority of the clusters (33). Though microRNAs are ubiquitously expressed in cells, the amounts of individual microRNAs can be cell-type-specific, and can vary temporally or as per physiological or disease state. Some microRNAs, like let-7a, miR-16 and miR-21, appear to be extremely abundant in most mammalian cells, while some like miR-302a and miR-122 have expressions that are highly restricted to specific tissues (e.g., 29, 34). In terms of weight, microRNAs are believed to constitute less than 0.05% of cellular RNA, but because of their small size, they form a sizable molar fraction. Extracellular microRNAs, most of which appear to be secreted within microvesicles from cells (exosomes, figure 1), are found in bodily fluids such as urine, milk, serum and sputum. The microRNAs are protected from the strong ribonuclease activity present in such fluids because of their encapsulation within the vesicles (e.g., 35, 36) and possibly because of protection by specific proteins that bind them (37). Total RNA extraction methods, such as those using organic solvents or spin-columns with RNA-binding matrices, are used for the extraction of microRNAs. Techniques to enrich the microRNA-containing small RNA fraction of total RNA preparations are also available. Perhaps because of their small size, microRNAs appear to be preserved very well in formalin-fixed and paraffin-embedded (FFPE) tissues (e.g., 38) as well as in degraded total RNA preparations (39). Extracellular microRNAs have been found to be preserved well in desiccated bodily fluids even without refrigeration (40, 41).

RNA quantification techniques like Northern blotting, reverse transcription-PCR (RT-PCR), in situ nucleic acid hybridization, and microarrays are used for detecting microRNAs. Novel methods that rely on principles such as surface-enhanced Raman spectroscopy (42) and nanomechanical sensing (43) have also been developed. The sensitivity, specificity and cost associated with the different microRNA detection technologies vary, though many of them offer unique advantages (44). For instance, in situ hybridization provides additional information on the spatial distribution of microRNAs, and Northern blots can be used to simultaneously quantify pre-microRNA levels.

Our knowledge of the functions and mRNA targets of specific microRNAs is currently limited, and studies of microRNA functions often start by first identifying microRNAs whose levels are significantly affected in a disease state. Unlike for microRNAs, there is a significant body of information associating mRNA expression profiles with esophageal cancer (45). At least some of the biological functions of many genes are known, and compared to microRNA profiling, mRNA profiling can more readily delineate the immediate pathways involved in biological processes. However, unlike the latter, microRNA expression studies do not require fresh or frozen specimens and can use cell-free bodily fluids. Further, probably because microRNAs are 20-30-times less in number than mRNAs, their profiles might be more robustly analyzable, yielding more accurate classifiers (46).

Alterations in microRNA levels, and its engineering

Changes in levels of specific microRNAs in tissues have been associated with diseases such as cancers (47) and diabetes (48), and with particular physiological conditions such as pregnancy (49) and muscle hypertrophy (50). Profiles of microRNAs in bodily fluids such as serum, saliva and urine too have been correlated with conditions such as myocardial injury, Sjögren’s syndrome and urinary bladder cancer, respectively (51-53). In general, the exact causes underlying such alterations are not known for most cases, though the molecular bases are known for many. Deletions of the genes for miR-15 and miR-16 have been shown to cause down-regulation of levels of those microRNAs in chronic lymphocytic leukemia (54). In many cases of mixed lineage leukemia-rearranged acute leukemias, DNA copy number amplification is known to cause overexpression...
of microRNAs of the \textit{miR-17-92} cluster (55). The reduced amount of microRNA \textit{let-7} that is seen in many tumors is believed to be because of overexpression of Lin28, an RNA-binding protein that causes polyuridylation and degradation of the \textit{let-7} pre-microRNA (56). Global reduction in microRNA levels in cancer cells have also been noted (46). This has been attributed to causes such as mutations in the Dicer-encoding \textit{DICER} gene in familial pleuropulmonary blastoma (57), targeting of transcripts for Dicer itself by microRNAs \textit{miR-103} and \textit{miR-107} in metastatic breast cancer (58), and mutations in the gene encoding for TRBP protein in many cases of carcinomas (59). A global increase in microRNA levels too has been found. In high-risk myelomas, this is believed to be caused by an overexpression of the gene encoding for the Ago 2 protein (60).

In vitro studies using cell-culture models have unveiled many pathways responsible for physiological changes in levels of specific microRNAs. For example, during induction of the contractile phenotype in smooth muscle of the human vasculature, signal transduction through the transforming growth factor β (TGFβ) and bone morphogenetic protein (BMP) family of growth factors causes a rapid increase in levels of \textit{miR-21} (61). In human breast cancer cells, activation of the estrogen receptor α (ERα) results in reduced levels of many microRNAs, such as \textit{miR-16} and \textit{miR-145}, by suppressing their maturation (62). Binding of hypoxia-induced factor 1α (HIF1α) to a hypoxia-responsive element in the promoter of the \textit{miR-210} gene is responsible for the overexpression of \textit{miR-210} in hypoxic cells (63).

Levels of specific microRNAs can be engineered both in vivo and in vitro to study their biology as well as potential as therapeutic targets. Transgenic techniques for gene knock-out or conditional expression have been used for causing aberrant or conditional up-regulation or down-regulation of microRNAs in animals such as mice and in cultured cells (e.g., 64, 65). Overexpression can also be achieved through traditional molecular biology methods such as transfection of plasmid DNA bearing microRNA genes or of precursor microRNA molecules, and transduction by engineered lentiviruses. Antisense nucleic acid molecules are commonly used to cause a knockdown of microRNA levels (66). Functional knockdown of a microRNA in vivo or in vitro without an actual reduction in levels of the microRNA has also been accomplished using lentiviruses that express decoy RNA with microRNA-target sites (67). Many studies on the therapeutic potential of such microRNA engineering have shown promising results. For instance, intratumoral as well as systemic delivery of synthetic \textit{let-7} microRNA, whose level is downregulated in lung cancer, was found to cause tumor regression in a mouse model of lung cancer (68), and disease progression in a mouse model of hepatocellular carcinoma was found to be halted by systemic delivery of adeno-associated viruses engineered to express \textit{miR-26a} (69).

**MicroRNAs and esophageal carcinoma**

Guo and colleagues were the first to report microRNA expression profiles in esophageal cancer, in 2008. Microarrays were used to profile 435 microRNAs in RNA extracted from fresh-frozen specimens of 31 pairs of ESCC and corresponding adjacent normal esophageal tissues (70). One-hundred-ninety-one microRNAs were considered detectable, and their expression profiles could be used to discern cancerous from normal tissue with >90% accuracy. MicroRNAs \textit{miR-25}, \textit{miR-424} and \textit{miR-151} showed upregulation, and \textit{miR-100}, \textit{miR-99a}, \textit{miR-29c}, and \textit{miR-140} showed reduction in cancerous tissue. Higher expression of \textit{miR-103} and \textit{miR-107}, known to affect metastatic potential of cancers by downregulating Dicer levels (58), was associated with poor prognosis. In a study that was published in the same year, Feber, et al., used RNA from fresh-frozen tissue samples from ten cases each of EAC and ESCC, and five cases of BE, to assay the expression of 328 human microRNAs (71). Compared to normal esophagus, \textit{miR-203} and \textit{miR-205} were expressed two-to-ten-fold less in all three diseases, whereas \textit{miR-21} levels were three-to-five-fold higher. Reduced levels of \textit{miR-203} and \textit{miR-205} were also observed in columnar epithelium compared to normal squamous epithelium in a study that examined 377 microRNAs in 16 individuals using microarrays (72). Levels of \textit{miR-205} were also found to be lower in BE mucosa compared to normal adjacent epithelium as well as to neosquamous epithelium generated following ablation of Barrett’s epithelium with Argon plasma coagulation in a study involving nine patients (73). MicroRNA \textit{miR-21} was also identified as overexpressed in a study that used RT-PCR to examine 20 cases of ESCC and seven ESCC cell-lines, and in two other studies, and it has been shown to be an oncogene that promoted cell transformation by targeting transcripts for the Programmed cell death 4 (PDCD4) protein (74, 75-76). Though some microRNAs, such as \textit{miR-21}, \textit{miR-100}, \textit{miR-203} and \textit{miR-205}, were identified as being affected in esophageal carcinoma in more than one of the aforementioned studies, many, like \textit{miR-143}, \textit{miR-145} and \textit{miR-215}, whose levels are increased in EAC as well as BE (74), were not. Characteristics of patient populations and RNA quantification technologies, and differences in sample-sizes and data analyses are believed to be responsible for this, a theme that occurs recurrently in such biomarker discovery work.

In a large study with a sample-size of 170 (100 EAC and 70 ESCC cases), in which 329 microRNAs were quantified using
microarrays, differences in microRNA expression between the two histological types were clearly identified. Specifically, miR-194 and miR-375 were found to be expressed 5-6-times more in EAC compared to ESCC (74). In EAC patients with Barrett’s, but not in those without, low expression of miR-375 was associated with worse prognosis (hazard ratio [HR]=0.3, 95% confidence interval [CI]=0.2-0.7). Among ESCC patients, increased miR-146b, miR-155 and miR-188, and decreased miR-21 were associated with poor prognosis, with HR values ranging from 2 to 4. MicroRNA expression differences between BE and EAC were also been examined by RT-PCR in a cohort of 32 cases, and expression of miR-143, miR-145 and miR-215 was higher in the former (72). In a similar study involving 50 and 25 cases of BE and EAC, respectively, expression of miR-143 and miR-145, but not of miR-215, was higher in BE than in EAC (77). In the same study, using microarray-based assays for some of the cases, alterations in levels of microRNAs between diseased and adjacent normal tissue were seen for 0, 32 and 39 of 470 quantified microRNAs in BE with low-grade dysplasia (n=5), BE with high-grade dysplasia (n=5), and EAC (n=6), with 14 and ten up- and down-regulated similarly in the last two diseases.

The ability to predict a cancer patient’s response to chemotherapy or radiotherapy is a major goal of current translational research. Such predictability can be particularly applicable and relevant in esophageal cancer because of the ease with which pre-treatment cancer tissue can be sampled by endoscopy, and the current norm of administering chemo- or radiotherapy before surgery, in spite of limited pathologic response to it. MicroRNA profiling of the NCI-60 cell-lines has demonstrated associations between microRNA expression and sensitivity to chemotherapeutic drugs, suggesting that microRNAs might be usable as predictors, and possibly even modulators, of chemosensitivity (e.g., 78, 79). Recently, Hong, et al, showed that miR-29b, high levels of which were associated with poor prognosis in ESCC, targets transcripts of the MDRI drug-resistance gene and affects sensitivity of many esophageal cancer cell-lines to a variety of anti-cancer drugs (80). Targeting of MDRI by another microRNA, miR-27a, to alter esophageal cancer cell-line chemosensitivity has also been observed (81).

A few studies have examined the association of esophageal cancer with other molecular determinants of microRNA biology, besides microRNA levels per se. In a study involving 71 cases of esophageal cancer, post-operative survival was negatively associated with increased levels of RNASEN mRNA, while levels of transcripts for Dicer and DGCR8 had no correlation (82). The HR was 4.6 (95% CI=1.5-13.8). Further, RNASEN knockdown reduced proliferation of esophageal cancer cell-lines in vitro. The RNASEN protein interacts with DGCR8 and affects pre-microRNA processing (83). Genetic variations in a number of microRNA-related genes were identified as associated with susceptibility to the disease in a study of 346 Caucasian patients in whom 41 variations in 26 genes, including those encoding Dicer, DGCR8 and Ago 1, were examined (84). Certain polymorphisms in the genes for miR-196a-2 and miR-631 were associated with an increased risk for the disease (odds ratio [OR] of 1.7 in both cases), whereas a particular polymorphism in the gene for miR-423 was associated with a reduced risk (OR=0.6). Polymorphisms in the gene for miR-196a-2 have also been linked with risks for cancers of the liver, lung, breast, stomach, and head and neck (27, 28, 85-87). In a cohort of 11 patients, miR-196a was found to mark the progression of BE to low-grade dysplasia, high-grade dysplasia, and EAC, with rising levels (88). Some of these findings on miR-196a might be explained through its targeting of the transcript for Annexin A1, an anti-proliferative and apoptosis-mediating protein (88). The microRNA has also been shown to target transcripts for the S100A9 protein, also referred to as MRP14 (migration inhibitory factor-related protein 14), reduction of whose product has been associated with poorly differentiated ESCC (89). In a study of 444 sporadic ESCC cases among the Chinese Han, a single nucleotide polymorphism in the gene for miR-146a was found to be associated with an increased risk for the disease (OR=2.4, 95% CI=1.4-4.2), with risk being higher for smokers (OR=3.2, 95% CI=1.7-4.5) (90). A separate polymorphism was associated significantly with higher clinical tumor-node-metastasis (TNM) staging (OR=1.6, 95% CI=1.2-2.2).

In vitro studies using esophageal cancer cell-lines have helped identify roles for certain microRNAs in the biology of esophageal carcinoma. For example, miR-373 has been shown to target transcripts for LATS2 (large tumor suppressor homolog 2) protein, whose gene-locus, a locus for which loss of heterozygosity has been reported for esophageal cancer, was associated with poorly differentiated ESCC (91). MicroRNA miR-10b was found to cause increased invasiveness and motility of cells by targeting transcripts for KLF4 (Kruppel-like factor 4) protein (92). Elevated expression of the microRNAs in esophageal cancer tissues was shown in both studies. Similarly, miR-145, miR-133a and miR-133b, all of which are downregulated in ESCC, have been shown to target transcripts for FSCN1 (actin-binding protein, Fascin homolog 1) that is associated with esophageal squamous cell carcinogenesis (93).

Conclusion

The study of the role of microRNAs in esophageal cancer appears to be emerging from infancy, and one can anticipate
more extensive examinations in this area in the near future. Many of them will help elucidate biology of the disease, especially when considered in concert with mRNA and protein expression studies. Some may have an immediate translational value through the development of microRNA biomarkers to improve disease screening and management.

Acknowledgements

We acknowledge the support provided by the American Association for Thoracic Surgery in the form of a summer internship award to R.M.

References


