MicroRNAs, SNPs and cancer

Angela V. Vitale,1 Huiping Tan,1,2 Peng Jin1

1Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA; 2Division of Histology and Embryology, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, People’s Republic of China

Abstract

MiRNAs are probable regulators of cell events such as differentiation, propagation and apoptosis. These cellular phenomena are also associated with benign and malignant tumor cells, therefore, it is presumed that miRNAs act as natural oncogenes or tumor suppressor genes. Whether a particular miRNA serves as either could almost be moot when the additional problems of SNPs enter the fray. A miRNA involved with SNPs (miR-SNPs) on any regulatory level, whether naturally cancer-inducing or not, could easily undergo an oncogenic transformation. This work reviews targets of miRNAs and the miRNAs themselves frequently containing SNPs reflecting different risks and markers of cancer with emphasis on familial groups and populations of shared heredity.

Introduction

MiRNAs are short (19 to 25 nucleotides long), evolutionarily conserved RNA structures that can bind to the mRNA of protein coding genes.1−3 Most often miRNAs bind within the 3’ untranslated region (UTR),4,5 though there are cases in which miRNAs can bind the 5’ UTR6 and even within coding regions of mRNA.4,7 MiRNAs, like miRNAs, possess 5’ phosphate groups and 3’ hydroxyl termini. They are often found within introns of coding genes, but some miRNAs have their expression driven by a separate promoter. Nucleotides 7 to 8 bases in length, often referred to as seed sites, are located at the 5’ end of the mature form of miRNA.8 The seed sites are usually defined by evolutionary comparison; they are generally conserved among distantly related species, though there is a weaker evolutionary conservation with the 3’ end of miRNAs as the tail is marginally involved in binding to the target site.9 It has been computationally estimated that up to 30-60% of genes can have their expression altered by the presence or absence of miRNAs that bind to the seed site target located on the mRNA.10,11 Single nucleotide polymorphisms (SNPs) are base pair changes with DNA that occur with a frequency of about 1 in 12,500 base pair or at approximately 99% of the sites in which the same residue is present on both homologues of chromosomes.12 SNPs are by far the most common form of mutation in the human genome. SNPs serve as guides to delineate possible markers for disease causing loci, or the loci themselves in databases such as HapMap. SNPs have typically been used for cancer-association studies in different ways. One involves direct examination of genes known to be involved in the cancer pathway; these studies are not always fruitful as they lack statistical power and are limited to a few genes known to interact with a specific oncogene or tumor-suppressor gene.13 The other uses genome-wide association studies (GWAS) in order to examine cancer association within a large population or SNPs can work with and within miRNAs to influence translational control of miRNAs.14,15 SNPs are able in some cases to generate or abolish miRNA binding sites.16,17 SNPs have also been credited as activating miRNAs to become oncogenes or tumor-suppressors.14,15 These SNP base pair changes, whether within the target site of the miRNA or the miRNA itself have been associated with many cancers, both in vitro and in vivo.18,22 It is not likely coincidental that about half of all miRNAs are located at fragile sites as well as sites known to be involved in cancer.23 This review largely covers the interaction of miRNAs with their target sites, but it should be noted that miRNA containing polymorphic SNPs can affect transcription of the primary transcript, and additionally, how the precursor miRNA interacts with downstream miRNA processing proteins. After screening more than a hundred tumor tissues representative of 20 cancers, the expression of one miRNA, let-7e, was significantly downregulated in vivo when a SNP transforming an A to a G (A>G) 17bp downstream of the miRNA was examined.24 Though this was not a bioinformatics study it demonstrated that SNPs within the pri- or pre-regions of miRNA could affect miRNA processing. Exact knowledge on the manufacture of the atypical expression remains elusive.24 Currently, the in silico prediction of miRNA interaction with a purported target site does not always agree with in vivo studies, though these predictions do lead to further avenues of exploration via in vitro studies and effective case control studies.23 The association of population based SNPs with cancers, however, is a somewhat contested issue. It has been suggested that many of the population sizes used to measure the connotation with SNPs and cancer are not large enough to make some of the claims of association and as such more careful case control studies are needed.25 Also imperative is the need to link bioinformatics, in vitro examination, in vivo research and large case control studies.27

Canonical pathway of mammalian miRNA biogenesis

The diminutive miRNA strands are first synthesized as longer structures encoded for by RNA polymerase II (RNA Pol II)26 in the typical pathway for miRNA biogenesis. These primary miRNA structures, termed pri-miRNA, can be up to 3 kilobases long. Pri-miRNAs have a 5’ cap, a 3’ polyadenylated tail27 and have significant secondary structure.21 Two RNA polymerase III proteins, Drosha and Dicer, are responsible for cleavage of the pri-miRNA into its subsequent form.22 Before leaving the nucleus the pri-miRNA structures are cleaved by a complex containing Drosophila and another associated protein, DGC41, DGC42 is a double stranded RNA binding protein that plays a critical role within the microprocessor complex. DGC43 simultaneously binds the pri-miRNA at a single stranded and double stranded form. DGC43 recognizes these regions while Drosa is then free to excise a 60 to 70 base pair stem loop structure (pre-miRNA) from the preceding form of miRNA by cleaving single stranded RNA tails near the major stem loop structure.24 Gregory et al.25 found that almost 20 other proteins associate with the microprocessor complex, though they also showed that the Drosa/DGC43 complex is necessary and sufficient for correct cleavage of pri-miRNA into pre-miRNA. Several proteins, including the DEAD box helicases p68 and p72, and hNRPU1-like were found to somewhat
lower the amount of pre-miRNA cut correctly, suggesting there is a larger role for them in the microprocessor machinery. In order to escape the nucleus, correctly spliced pre-miRNA binds with Exportin-5 with assistance by Ran-GTP. Incorrectly spliced pre-miRNA has a lower efficiency of transfer to the cytoplasm.16

Once successfully passed into the cytoplasm pre-miRNA is processed into an RNA duplex.37 The strands of the miRNA duplex are cleaved by a second RNase III enzyme, Dicer,16 which works alongside TAR-RNA-binding-protein (TRBP) to remove the terminal stem-loop structure.19 This cleavage releases two strands of miRNA. The most thermodynamically stable strand, or guide strand, will become the mature miRNA and complex with the Argonaute-2 (AGO2) containing RNA-inducing silencing complex (RISC) and the less stable secondary strand (denoted mir*) is degraded.16 The miRNA target is found by the complimentary mature miRNA via RISC. The mature miRNA and miRNA contain limited base-pairings along the target site. This imperfection could thus allow a single miRNA to potentially interact with hundreds of miRNAs.11 The mRNA target is then translationally repressed and often slated for mRNA degradation.9

**MiRNAs and their relationship to cancer**

Many miRNAs have been associated with certain cancer phenotypes. The first known reporting of miRNAs and their association with some cancers was shown in Calin et al.41 This study showed a deletion of miR-15a and miR-17-92 in chronic lymphocytic leukemia (CLL). This group further demonstrated that a mutation in the pri-miR-16-1 results in down-regulation of the miRNA.41 Other studies have also shown linkage between specific and non-specific cancers.21 For instance, the miR-19-92 cluster is frequently found rearranged within lymphomas23 and the miR-17-92 cluster is found to be highly expressed in a variety of tumors42 and is associated with the binding of c-myc to E-boxes for activation of transcription.44 In vivo and in vitro studies confirm miR-19a targets transcription factor V-maf musculoaponeurotic fibrosarcoma oncogene homolog B (MAFB) and that depletion of miR-19a upregulates HOXA1 expression. It was also shown that miR-19a directly targets the 3’UTR of HOXA1 RNA.46

Conversely, leukemic megakaryocytes show upregulation of miR-101, miR-126, miR-99a, miR-135, miR-20. Additional works have pointed to miRNA differential expression leading to context dependent effects in some cancers. Expression signatures of cancer gene targets within solid tumors are also beginning to be explored45 and recently solid tumors were used for deep sequencing and discovery of new miRNA SNP regions.64 However it is unknown whether these novel sequences will shed light on SNP regions that are differentially expressed across cancers within the same familiar clades.

**Breast Cancer**

A SNP in the precursor form of miR-146a could be a target for predicting age of onset for both ovarian and breast cancer47 though there is some doubt about the case control methodologies.48 A SNP in the gene antecedent (rs2910164) changing a G-U pair to a C-U pair in the stem region was recently associated with age of onset of breast cancer (BC) and ovarian cancer (OC) in unrelated groups.49 In vitro analysis demonstrated that the rare SNP variant binds the 3’UTR of BRCA1 more commonly than the more common allele. This study suggests that the miR-146a mutant precursor may be concomitant with ovarian cancer and breast cancer.50 A later study showed that the miR-146a pre-miRNA rs2911064 C>G allele was in Hardy-Weinberg equilibrium with the rest of the comparative population with in a case control study among Chinese women.49 Other studies point toward bioinformatics methodologies that could shed light on both miRNAs and their target sites with a role in cancer.46-51 Recently, in a case control study involving unrelated Chinese women of Han ethnicity, two out of four pre-miRNAs studied were shown to have significance with increased risk of BC. The Hu et al. study indicated that hsa-mir-1889a3 rs11614913: T>C and hsa-mir-499 rs3746644: A>G were distributed more heavily in women of like descent.49 The research also points out two genes, LSP1 and TOX3, according to GWAS studies, are associated with hsa-mir-19a2-3p and hsa-mir-196a2-3p as newly identified BC susceptibility markers.52

In a smaller case study, the estrogen receptor 1 (ESR1) protein product has been shown to affect BC risk in women; based on a study predicting polymorphic SNPs effect on gene expression an ESR1 miRNA binding site was examined for association with BC onset.53 The populations amassed for the study included familial BC cases and isolated cases of early BC. A minor allele of ESR1 (ESR1 rs2747648T>C) within a predicted miR-453 binding site was negatively correlated with premenopausal women and the onset of BC.54 The allele (T>C) has a protective effect against BC, even more so in cases of familial BC and when the C allele was present in the homozygous condition.55

**Lung cancer**

Lung cancer is the third largest cause of cancer-related deaths among men and women in the United States.56 Let-7 has been implicated as an oncogene in many human cancers. Let-7 is a direct downstream target of the RAS gene family and a recent report by Chin et al. examines the connection between Let-7 and KRAS, a gene within the RAS superfamily.57 LCS6 is a newly found SNP in the 3’ UTR of KRAS, a target of Let-7. Expression of KRAS,
and a consequent lowering of expression of Let-7, was found to be significantly associated with non-small lung cancer.62 Shortly after LCS6 was shown to be tied to lung cancer involving low-dose smokers, a reexamination of the data cast doubt on the association.63 Though involvement of KRAS as a lung cancer oncogene and lowered Let-7 expression is not in doubt, the LCS6 SNP is not found to be involved with greater risk of lung cancer.16 The second study, involving the same populations as the Chinn et al. paper uses a slightly different analysis of the data and the authors discuss that the use of a smaller subset of cases coupled with a very low association of LCS6 with lung tumors may play a role in the most current study negating the former study.16

Other extremely important genes that require deeper investigation are those directly responsible for of miRNA processing. It cannot be ignored that any polymorphisms including SNPs within miRNA biosynthesis genes can have a direct effect on an individual's cancer susceptibility. In a two stage study using a Sequenome mass spectrum-based genotyping assay for stage 1, 11 miRNA were examined for lung cancer association.65 The intronic AGO1 rs368332A>G was found to be a good candidate for further study using a larger population for analysis in a case-control study investigating a Korean population with a little over both 500 cancer patients and control healthy patients. Interestingly, individuals with at least one AA allele at rs636832 have a higher risk of lung cancer while those with an AG or GG alleles have a protective effect against lung cancer.66 A possible link between cancer risk of smokers and non-smokers was examined with regards to AGO1 rs636832. The AA allele was found to be initially correlated to lung cancer prediction in heavy smokers, but in a multivariate logistic regression this correlation was not found.67 Further study on this AGO1 SNP in larger case-control studies and in different ethnic groups is needed to elucidate its relationship with lung cancer.

Hepatocellular carcinoma

Hepatocellular carcinoma (HCC) is responsible for the majority of liver cancers44 with the prevalence occurring in China.65 Worldwide, HCC is the fifth most widespread cancer and is responsible for a third of cancer deaths.66 Mir-146a rs2910164 SNP GG genotype was coupled to hepatocellular carcinoma (HCC) in males.44 This is in contrast to another SNP change in the mir-146a, a G>U pair to a C>U pair.46 This case control study used male and female unrelated Han Chinese participants, 479 HCC patients and 504 controls without occurrence of HCC.44 Hepatitis B Virus (HBV) was found in a large cohort of the HCC patients (88.9%) suggesting that HBV plays a role in cancer onset of HCC and may have some interaction with the mir-146a variant allele. Production of mature mir-146a was also studies by transiently transfecting 293T cells with either the GG or CC allele. The cells transfected with the GG allele produced more mature mir-146a. Also discovered was the ability of the GG allele to promote colony formation and proliferation in transfected NIH3T3 cells.44

Two recent studies investigated origins of HCC and found that mir-196a2, coupled with cirrhosis of the liver, has diagnostic implications for HCC.67,68 Using a Han Chinese population, 310 HCC patients with cirrhosis and 222 individuals with cirrhosis but without HCC, were examined for an rs1614913 miR-196a2 polymorphism in the first study.67 Patients with HCC and cirrhosis had a higher level of the rs1614913 CC genotype. Various stages of tumor tissue were collected from 59 HCC patients and the expression levels of miR-196a were examined. No significant differences between rs1614913 phenotypes were seen among the different grades and stages of tumors, though a slight association with the T allele was shown with tumor progression.67 Patients with a CC or CT genotype overall had a higher preponderance of HCC. Because mirR-196a2 has been previously shown to effect expression of mature mir-196a it was thought that levels of mirR-196a2 might increase levels of mirR-196a. Indeed mirR-196a expression levels were increased in patients with a CC or CT genotype in mirR-196a2 suggesting these miRNA polymorphisms play a role in HCC onset in patients also displaying cirrhosis.57 In the second study 560 patients examined had HCC and 391 individuals without HCC were used as the control population.68 As in the previous study males with mirR-196a2 rs11614913 CC genotype had higher levels of HCC diagnosis. This case-control study also examined mirR-196a2 expression in different tumor stages and concluded that the miR-196a2 rs11614913 C allele was indicative of patients with certain types of tumors but not in patients with large tumor, advanced-stage tumor or lymphatic metastasis thus suggesting differential gene regulation is playing a role in HCC stages.58

Other cancers

Other mir-SNPs, through case control studies, have been found to amplify or diminish risk of other cancers. For instance, cervical cancer is the second most globally reported cancer for which Human Papilloma Virus (HPV) is responsible.69 Han Chinese women were used in a case control study recently and a mir-SNP was examined in the LAMB3 pathway.70 HPV through genes E6 and E7 blocks the expression of mir-218. LAMB3 expresses laminin-5 and this protein is greatly reduced in the absence of mir-218. The lack of laminin-5 then further stimulates the HPV. The study yielded an interesting result, SNP rs11134527 within pri-mir-218 has a variant linked to increased cervical cancer.70 It is even postulated that the variant may play a role in increased risk of HPV infection. LAMB3 was shown to be a direct target of mir-218 through this work. The study could, by the author’s admission, be expanded as the numbers involved in the case study were rather low and the controls were women who has self-reported to have no cancers, but may have been harboring other unknown cancers.70

A non-synonymous substitution in GEMIN3 has been coupled with increased bladder cancer risk.71 This mutation is somewhat similar to the AGO1 SNP as there is an increased risk of lung cancer which examines an miRNA biosynthesis gene's relationship with cancer.43 GEMIN3 codes for a core protein of a larger complex that plays a role in pre-miRNA splicing;72 the protein is also in a 15S ribonucleoprotein complex containing eIF2C, another protein that is of great consequence with regard to miRNA processing.73 The population was a large and homogenous, composed of Caucasian patients diagnosed with bladder cancer and a control group.73 This analysis once again points to the importance of further study of miRNA processing genes that may alter expression of a myriad of miRNAs potentially involved in tumorgenesis.

Also targeting Mir-SNPs for targeted exploration of linkage in an effort to aid in identification, early-stage head and neck cancer patients with high and low risk secondary primary tumor (STP) and high- and low-risk cancer reoccurrence.74 The population contained only 150 patients and 300 controls matched by age, gender and ethnicity.74 Though 18 mir-SNPs were found to be associated with STP and/or reoccurrence, one mir-SNP in particular, rs5747238, is located in a miRNA binding site within SMC1B.75 The 18 mi-SNPs were examined and found to be tied to STP/reoccurrence in a dose dependent fashion.75 Almost half of the SNPs were located in RNASEN (DROSHA).74 Though mutations in the RNASEN are likely to equally affect the processing of all pri-miRNA equally, it is postulated that since miRNAs are expressed differentially in tissue the RNASEN would then differentially affect tumorgenesis.75 SMC1B is suggested to play a role in chromosome structure during meiosis and mitosis.75 Polymorphism in microRNA Target Site (PolymiRTS) is a database of DNA changes in presumed microRNA target sites.76 PolymiRTS found that a SNP within SMC1B is likely to create miRNA binding sites for mir-609 and mir-124a.77 This SNP is thought to lower expression of SMC1B.
leading to increased carcinogenic potential via further genome volatility.75

Responsible for a third of deaths among genitourinary malignant cancers, renal cell carcinoma (RCC) causes 40% mortality among these patients.77 Though surgery is still the best therapy for RCC, recurrence will arise in 20-40% of patients.78 Because this cancer has such a high rate of mortality it is imperative that biological markers predictive of clinical outcome, including miR-SNPs be recorded. Seven miR-SNPs were found linked with cancer survival, however five miR-SNPs were associated with an additional RCC episode.79 GEMIN4, a protein functionally coupled with GEMIN3, has two SNPs, rs78137T>C and rs91025G>C associated with almost 1.75% risk of mortality.79,80 SNP rs3744741C>T was found concomitant with decreased risk of mortality.79,80 MiR-146a, miR-196a-2, miR-423, miR-608 and miR-601 are also concomitant with RCC recurrence.79 All of the SNPs found in the miRNAs are located in the pre-miRNA form.79

There is a scenario in which the heterozygosity of the SNP within a miRNA area can lead to greater likelihood of disease. SNP rs2910164 located in the 3 prime strand of miR-146a can lead to an increased risk of papillary thyroid carcinoma especially when present as a heterozygote.81 It is suggested that the heterozygosity somehow leads to a gestalt phenomena wherein the sum of the parts is less than equal to the whole via an epistatic effect between the two alleles.81 The same group then showed that the SNP produces three miRNAs unlike the normal two produced in homozygous affected individuals. Unlike the phenotype normally seen with heterozygous individuals, two of the mature miRNAs are produced from the 3 prime end of miR-146a and a third is produced from the leading strand.82 These three mature miRNAs have the ability to bind various miRNAs thus interrupting the normal miR-146a interaction with a predicted variety of miRNAs. It is thought that DNA-damage response pathways acting on cell death signals are invoked within the SNP heterozygote.82 This group demonstrates the genetic complexity of miRNA interactions with target site, and the importance of somatic mutation with regards to an oncogenic phenotype.82

As related with BC risk and HCC, miR-146a rs2910164 is found once again to be associated with a distinct cancer, this time esophageal squamous cell carcinoma (ESCC) within a case control study among the Chinese Han.83 In this case the GG genotype was attached to the ESCC state83 rather than a change from the G>U pair to a C>U.87

Like HCC patients, a North Indian population was shown to have a significant risk of developing prostate cancer with a polymorphism in miR-196a2 (rs11614913).84 In this candidate gene study, miR-499 (rs4984364) also demonstrated significant association with prostate cancer in this case-controlled study which examined 159 prostate cancer patients and 230 controls.84 In this case the heterozygote allele CT in miR-196a was associated with disease outcome and interestingly a heterozygote CT genotype in miR-499 also showed linkage to prostate cancer. MiR-196a and miR-499 may work in concert to produce influence the onset of prostate cancer and my serve as prognostic and diagnostic indicators.84

**MiRNAs: Clinical outcome predictors?**

Discussed above are some of the miRNA binding sites and miRNA SNPs that putatively affect the outcome of BC. In silico scenarios suggests that miR-453 binds more strongly to an ESRI SNP and may thus effectively lower the amount of estrogen produced.85 Interestingly the majority of BCs do express estrogen receptors, though the cancers that do not are more difficult to treat.86,87 The facts that these estrogen receptor positive cancers, though more easily treated with certain drugs, do become easily resistant. It is suggested that hormone replacement therapy would be a valuable variable to study in addition to familial cancer patients and sudden onset patients.85 It would be invaluable to know if those patients with the ESRI SNP allowing more dynamic binding of miR-453 would positively respond to endocrine treatment, thus leading to an individualized plan of treatment.83 The Kontrovich et al, study addresses their shortcomings.88 More affected and unaffected individuals need to be included in the study, especially in light of the preliminary BRCA2 data suggesting that this particular population contains a mutant SNP possibly affecting regulatory actions leading to BC. Again, with more case control studies focusing on larger and more diverse populations (this initial study only focused on an Ashkenazi population displaying little heterozygosity at the SNP mutation sites), the possibility arises that personal treatment plans could be designed for patients.84 Yet another study involves integrins which dictate cell adhesion to the extracellular matrix.85 In a large case control study 746 Swedish patients with current or former instances of BC were examined along with 1493 individuals without BC as controls. Probable target sites of miRNA SNPs were examined in integrin genes and a strong association between the ITGB4 rs743554 A allele and aggressive tumor formation was discovered.86 This allele could be a strong predictive indicator of BC risk.

It has also been shown that several cancer cell lines will alter protein expression based on differential regulation of miR-638 and miR-628-5p and was concluded that small differences in protein expression caused by the interaction of certain regulatory genes SNPs and miRNA will influence the onset of certain cancers.85 In one case control study miR-21 was discovered to be involved with lung cancer and proved a chemotherapy response marker.86 MiRNA related SNPs have also been associated with colon cancer in patients treated with 5-flourouracil and irinotecan.87 These SNPs were associated with various genes, including rs1834306, within pri-miR-100 and rs7372209 also located in a pri-miRNA, pri-miR-26a.86 Investigation of SNPs directly involved with deleterious effects of cancer drugs would greatly facilitate basic research studies for cancer. Further studies involving GWAS and large case control studies could certainly go a long way to advancing these research studies into something that could be used to tailor-make a treatment plan for cancer.

A significant movement advocates personalized medicine.85,86 This group feels that much can be done to assist individuals with a particular SNP (or greater than one SNP) that may leave them more likely for cancer onset. Foremost amongst this concern is the large number of adverse drug reactions among cancer patients. Dihydrofolate reductase (DHFR), when overexpressed, leads to methotrexate resistance, a drug primarily used to treat cancer but is also used to treat such conditions as psoriatic arthritis.88 SNP 892C>T mutation near the 3’UTR of DHFR hinders miR-24 from binding its target site located within the 3’UTR. This causes upregulation of DHFR along with its consequent drug resistance leads to recurrence or even inability to fight cancer. Acknowledging that much more can be done to identify and validate particular miR-SNPs and their association to a diseased state a recent paper points to the new laws enacted by the United States to both encourage parents of individuals or adult individuals to have their genome sequenced in its entirety.82,83 By having a trove of genetic information some sense may be made of miR-polymorphisms and how to effectively diagnose and treat individuals possessing known SNPs showing incomplete penetrance within a population or limited expressivity within an individual.

Excitingly, a new study shows that miR-SNPs have been associated with prostate cancer in men can be used to effectively predict how effective of androgen-deprivation therapy (ADT).84 15 total SNPs spread between three prognoses (disease progression, prostate cancer specific mortality and all-cause mortality) were found.84 These SNPs were found within miRNAs and miRNA binding sites. Combinatorial analysis between SNPs show that during ADT, patients having a larger number of adverse genotypes have a more rapid
time to progression and poorer prostate cancer-specific survival rates. In effect miR-SNPs could act as prognostic markers in patients.

Limitations of miR-SNP studies

MiR-SNP studies with regards to disease relationship have at least two major caveats, the population size used in the study and the diversity of that population. There are also conflicting reports about miR-SNPs and their ability to affect cancer outcome. A recent report on colorectal cancer (CRC) in Han Chinese using 126 CRC patients and 407 healthy individuals showed that a mir-196a2 polymorphism (rs11614913 T>C) is not associated with CRC. This is in direct conflict with non-epidemiological studies demonstrating that the polymorphism could be involved with CRC onset. Therefore mir-196a2 would not be indicative of a CRC condition. The gene bcl-2, its overexpression regulated by miR-16 +7, has been indicated as tumor suppressor with regards to chronic lymphocytic leukemia (CLL). A recent study using 39 CLL patient demonstrated that miR-16 +7 would be a poor diagnostic marker of CLL.

It should be noted however that these case control studies did not incorporate GWAS studies. GWAS is a more useful tool than traditional methods involving linkage studies and candidate gene analysis. Because many cancers have polycogenic origin it is important that many loci are examined. GWAS has the statistical power to perform this task in an unbiased way.

GWAS has its own limitations however in that the data set produced by many studies is large and complex, leading to confusion about which SNPs may be relevant to disease. GWAS analysis also requires validation in unrelated populations are necessary. To assist in data analysis tools have been constructed that will closely examine the amount of linkage disequilibrium between SNPs. These tools are useful in analyzing miR-SNP data as some studies do not incorporate GWAS, but instead concentrate solely on linkage data and candidate gene analysis.

Conclusions

MiR-SNPs have already been shown on a molecular level to be associated with a plethora of cancers. It is imperative not only to have insight into which cancers in general have differentiating levels of miR-SNPs, but also to have an understanding of which miR-SNPs would best predict cancer onset and outcome. MiR-SNPs, used within GWA studies with a large case-control group reflective of true population heterogeneity may ultimately prove successful in predicting cancer risks.

Interestingly, mir-146a has been shown to be involved in a large range of cancers including BC, OC, papillary thyroid carcinoma, renal cell carcinoma and ESCC. Much work has been done on BC. miRNA mutation as well as mRNA with modified miRNA binding sites are likely to be useful in disease prognosis. A critical point is that heterozygosity of miRNA-SNPs can have an epistatic effect on gene expression. Also crucial, not only do cancers originate from SNPs located on miRNA binding sites of genes directly associated with cancer and SNPs within the miRNAs themselves, cancers are also associated with SNPs within the miRNA machinery genes themselves. Some cancers show a relationship between miRNA binding sites in genes such as GEMIN3 and RNASEN in bladder cancer and early-stage head and neck cancer respectively.

Genetic differentiation among cancers across populations is only starting to be extensively documented. As just one example, renal cell carcinoma would greatly benefit from miR-SNP markers to trace patients likely to need advanced forms of therapy to avoid remission after initial treatment. One cautionary note to building large databases with miR-SNP data from many populations would be maintenance of privacy to the donors. MiR-SNPs will prove a notable trove of data and most likely be very effective as a clinical outcome predictor.

References

48. Li XD, Li ZG, Song XX, Liu CF. A variant in an miR-125a alters the processing of pri-miRNA. Hum Mol Genet 2008;17:1793-801.
64. El-Serag HB, Rudolph KL. Hepatocellular Carcinoma: Epidemiology and Molecular Carcinogenesis. Gastroenterology 2007;132:2357-76.
82. Jazdzewski K, Liyanarachchi S, Swieniak
85. Shpargel KB, Matera AG. Gemin proteins
88. Janzen NK, Kim HL, Figlin RA, Belldegrun
89. Mishra PJ. MicroRNA polymorphisms: a
90. Rossbach M. Small non-coding RNAs as
91. Schepeler T, Reinert JT, Ostenfeld MS, et
92. Sethupathy P, Collins FS. MicroRNA target
104. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661-78.