Published in final edited form as:
Nat Rev Cancer. 2010 June ; 10(6): 389–402. doi:10.1038/nrc2867.

Genetic variation in microRNA networks: the implications for cancer research

Bríd M. Ryan[*,‡], **Ana I. Robles**[‡], and **Curtis C. Harris**[‡]

[*]Cancer Prevention Fellowship Program, Center for Cancer Training, National Cancer Institute, National Institutes of Health, 37 Convent Drive, Building 37, Room 3068A, Bethesda, MD 20892-4258, USA

[‡]Laboratory of Human Carcinogenesis, Center for Cancer Research, National Cancer Institute, National Institutes of Health, 37 Convent Drive, Building 37, Room 3068A, Bethesda, MD 20892-4258, USA

Abstract

Many studies have highlighted the role that microRNAs have in physiological processes and how their deregulation can lead to cancer. More recently it has been proposed that the presence of single nucleotide polymorphisms in microRNA genes, their processing machinery and target binding sites affects cancer risk, treatment efficacy and patient prognosis. In reviewing this new field of cancer biology, we describe the methodological approaches of these studies and make recommendations for which strategies will be most informative in the future.

Human populations are estimated to be 99% identical at the level of the genetic code; thus, human diversity (other than epigenetics) arises from the remaining 1% of variation[1], most of which is due to single nucleotide polymorphisms (SNPs). These are a non-repetitive form of sequence variation that was first identified in 1978 in the β-globin gene cluster2. To date, approximately 10 million SNPs have been identified in the human genome, occurring on average every 100 to 300 base pairs (International HapMap Project website; see Further information)[1,3]. Although most SNPs are silent, epidemiological studies have established a link between variations in gene sequence, environmental interaction and cancer risk. By identifying genetic markers of susceptibility and characterizing gene–environment

Correspondence to C.C.H. Curtis_Harris@nih.gov .

Competing interests statement
The authors declare no competing financial interests.

DATABASES
Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene.
BMPR1B | BRCA1 | BRCA2 | CD86 | DHFR | DICER1 | DNMT3B | ESR1 | GEMIN3 | GEMIN4 | IL1A | KRAS | let-7a-3 | pre-mir-27a | pre-mir-196a-2 | SETD8 | TRBP
MiRbase: http://www.mirbase.org/
Let-7b | let-7e | miR-10a | miR-16-1 | miR-21 | mir-26a-1 | mir-27a | miR-122 | mir-124-1 | miR-182* | mir-149 | miR-189 | mir-196a-2 | miR-378 | miR-423 | miR-453 | miR-575 | miR-582 | mir-BART22 | mmu-miR-10a | mmu-miR-27a | mmu-miR-27c | mmu-miR-155 | mmu-miR-222
UniProtKB: http://www.uniprot.org
RAN | XPO5

SUPPLEMENTARY INFORMATION
See online article: S1 (table) | S2 (box)
ALL LINKS ARE ACTIVE IN THE ONLINE PDF

interactions, it might be possible to reduce cancer mortality through early diagnosis and personalized therapy.

As our knowledge of the topology of the genome has evolved, a new class of non-coding RNAs has emerged called microRNAs (miRNAs). The latest release of the miRBase database has catalogued 721 human miRNAs. Smaller than protein-coding genes, miRNAs can regulate the translation of hundreds of genes through sequence-specific binding to mRNA4, and depending on the degree of sequence complimentary will result in the inhibition of translation and/or degradation of target mRNAs4,5. Interestingly, a recent report shows that miR-369-3p can upregulate the expression of its target, tumour necrosis factor-α (TNFα)6.

Our knowledge and understanding of miRNA biogenesis has evolved in recent years, and is thoroughly described elsewhere[4,7] (FIG. 1). Briefly, mature miRNAs are short RNA molecules of between 19 and 22 nucleotides in length. Nucleotides 2–7 of the mature miRNA sequence create the 'seed region' (REFS 8–11) that primarily specifies the specific mRNA that the miRNA will bind. The degree of specificity conferred by the seed region is comparable to that of the DNA sites recognized by transcription factors12. Although the binding between the seed region is (mostly) in perfect Watson–Crick complementarity, flanking regions do not have to bind with equal precision. In an additional analogy to transcription factors, it is now apparent that base pairing outside the seed region provides a further layer of specificity, just as chromatin structure limits the potential for transcription factor binding. As multiple transcription factors work cooperatively to ignite gene expression, so too can multiple miRNAs bind to cognate sites in the 3′ untranslated region (UTR) of target mRNAs. Indeed, the complexity of translation can be further extended through heterotypic miRNA–mRNA interactions, as genes can harbour binding sites for several miRNAs12–14.

To date, miRNAs have been linked to the aetiology, progression and prognosis of cancer[15], and miRNA expression profiles can uniquely identify cancer types[16,17]. The gain or loss of specific miRNAs can function as an oncogene or tumour suppressor[18,19], the archetypical examples of this being miR-21 and Let-7, respectively. It should also be noted that some miRNAs can have dual oncogenic and tumour suppressive roles in cancer depending on the cell type and pattern of gene expression[20]. In addition, approximately 50% of all annotated human miRNA genes are located in fragile sites or areas of the genome that are associated with cancer[21–23].

Their functional association with cancer, small gene size and potential to simultaneously affect a multitude of genes makes them unique candidate loci for conferring cancer susceptibility, as a small genetic change in an miRNA sequence can theoretically lead to widespread phenotypic effects[24,25].

The initial demonstration that miRNA-related SNPs can affect phenotype was elegantly depicted by Abelson *et al*.[26] who found that a mutation in the miR-189 binding site of *SLITRK1* was associated with Tourette's syndrome. Since then, several studies have used systematic sequencing or *in silico* approaches to identify SNPs in miRNA-related genes, catalogues of which have been created and made public[27–29]. These reports provide fertile

FURTHER INFORMATION
Curtis C. Harris' homepage: http://www3.cancer.gov/intra/LHC/LHCPAGE.htm
dbSNP: http://www.ncbi.nlm.nih.gov/projects/SNP/
International HapMap Project: http://hapmap.ncbi.nlm.nih.gov/
Patrocles database: http://www.patrocles.org/
PolymiRTS database: http://compbio.uthsc.edu/miRSNP/
RNAfold software: http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi
Targetscan database: http://www.targetscan.org/

ground for follow-up case–control studies to determine the association between these genetic markers and cancer risk.

At a glance

- Single nucleotide polymorphisms (SNPs) in microRNA (miRNA) genes (miR-SNPs) can be predicted to affect function by modulating the transcription of the primary transcript, pri-miRNA and pre-miRNA processing and maturation, or miRNA–mRNA interactions. Functional support for each of these mechanisms has been found for several individual miR-SNPs.
- SNPs in mature miRNAs and miRNA binding sites function analogously to modulate the miRNA–mRNA interaction and create or destroy miRNA binding sites.
- Several elements have to converge for an miRNA binding site SNP to be considered functional: the SNP must have a proven association with cancer, both the miRNA and its predicted target must be expressed in the tissue, and the allelic changes must result in differential binding of the miRNA and affect expression of the target gene.
- Computational prediction of miRNA binding sites and up-to-date coverage of SNPs is an essential part of these studies. Programs such as Patrocles and PolymiRTS intercalate and cross-reference these data with dbSNP information, and as such are invaluable in aiding the search for polymorphic miRNA binding sites.
- Case–control studies have provided evidence for an association of miR-SNPs and SNPs in miRNA-binding sites and cancer risk. These studies differ in the degree of functional support for the predicted interaction and mechanistic insight, as well as validation status.
- Although still lacking biological validation, SNPs in the miRNA processing machinery are likely to affect the miRNAome as a whole, perhaps leading to overall suppression of miRNA output. Despite several reported associations, none of the studies of SNPs in miRNA processing machinery has been independently validated, nor has the biological mechanisms of how they affect miRNA maturation and cancer been delineated.
- IsomiRs are miRNA structural variants that may arise from variable cleavage sites for DROSHA and DICER1 in the hairpin. A few isomiRs have been implicated in cancer, but associations with cancer risk have not been established.
- Both the regulatory and coding regions of genes can harbour miRNA binding sites, but research in this area remains scant. Sensitive alleles identified in epidemiological studies, but with obscure functional roles, should perhaps be tested under miRNA prediction algorithms that are not limited to the 3′ untranslated region of genes, particularly if evidence indicates that altered expression of that gene can be associated with the phenotypes.

SNPs in miRNA genes and cancer

Sequencing has shown that SNPs in miRNA coding genes, and specifically in miRNA seed regions, are rare[27,30] and would at first glance seem to be of limited functional importance. On close inspection, however, it is apparent that such rare occurrence is the result of well-documented evolutionary selective pressures[27,31] (BOX 1).

SNPs in miRNA genes are thought to affect function in one of three ways: first, through the transcription of the primary transcript; second, through pri-miRNA and pre-miRNA processing; and third, through effects on miRNA–mRNA interactions (FIGS 2,3). See Supplementary information S1 (table) for a list of SNPs in all currently known pre-miRNAs and mature miRNAs, which has been created using build 130 of dbSNP (see Further information). Most studies that have followed a biologically based candidate gene approach to search for SNPs in miRNAs that might confer cancer susceptibility rely on knowledge of a functional link between a particular miRNA and gene target.

The first evidence that point mutations in miRNA genes can have a functional effect and confer cancer susceptibility comes from a seminal study by Carlo Croce's group[32] in which a germline mutation in *pri-mir-16-1* was found in a kindred with familial chronic lymphocytic leukaemia (CLL) and resulted in low levels of miR-16-1 expression (FIG. 2). Moreover, the mutation was subsequently discovered in New Zealand black mice that naturally develop CLL-like disease[33]. However, this mutation was not detected in a large panel of other tumour types, thereby showing specificity for cancers of a particular origin[34]. Notably, the oncogenic *mir-17–92* cluster contains two miRNAs that harbour three SNPs (see Supplementary information S1 (table)), but the relevance to inherited cancer risk and carcinogenesis remains to be investigated.

To date, the literature suggests that the functional consequences of SNPs in pri-miRNAs relate to processing and levels of the mature miRNA. For example, SNPs in the pri regions of let-7e and *mir-16* lead to decreased mature miRNA levels[29,32]. Indeed, several studies show an association between pri-mir SNPs and cancer risk (TABLE 1). Specifically, rs7372209 in mir-26a-1 was associated with a 64% decreased risk of bladder cancer in females[35], and a twofold increased risk of premalignant oral lesions[36]. The rs531564 SNP in mir-124-1 is associated with an increased risk of bladder cancer[35] and oesophageal cancer in males[37], and the pri-miR SNP rs213210 in *mir-219* increased the risk of oesophageal cancer[37]. Another pri-miR SNP, rs2292932 in mir-149, has been tested in several cancers but has not been associated with cancer risk[38,39]. This suggests that the molecular mechanisms underlying the genetic associations of mir-SNPs with cancer are complex and vary by cancer site.

Case–control studies have also provided evidence for an association of pre-mir SNPs and cancer risk. Rs11614913 in pre-mir-196a-2 contributes to the risk of developing breast[40], lung[39] and gastric cancers[41] in the Chinese population (TABLE 1). In each case the rs11614913 variant homozygote CC was associated with increased cancer risk. Risk of developing oesophageal cancer in Caucasian males and never-smokers was significantly associated with the rs11614913 variant homozygote TT, the minor allele in this population[37]. Rs11614913 is located in the 3′ passenger (3p) strand mature sequence of mir-196a-2, thereby possibly affecting both maturation and the repertoire of target mRNAs with which it interacts. Indeed, previous studies have shown that sequence variations in mature and precursor miRNA sequences affect miRNA biogenesis[28,42], and levels of mature miR-196a-2 were lower in CC carriers than in TT carriers[39]. Notably, this SNP has also been associated with poor survival in patients with lung cancer[43]. Indeed, this was the first demonstration that miRNA-related SNPs could be related to cancer prognosis.

An SNP in the terminal loop of pre-mir-27a, rs895819, confers a reduced risk of developing breast cancer in families with a history of non-BRCA-related disease[44] and in families with mutant BRCA2 (REF. 45). It has been shown that artificial mutations in the terminal loop of miRNAs such as *mir-21* and *mir-30* can block miRNA maturation[46], so it is conceivable that the variant allele of rs895819 might impair the maturation of oncogenic mir-27a, thus explaining the protective effect of the SNP. However, no alterations of free energy or conformation of the miR-27a–mRNA duplex were predicted *in silico*, thereby leading to the

assumption that the SNP would not affect *mir-27a* maturation or targeting[45]. Nevertheless, SNPs in pre-miRNAs may affect expression even in the absence of apparent effects on its secondary structure. Such is the case for an SNP in *let-7e* (rs41275792), which leads to reduced levels of the mature miRNA *in vivo* even though its secondary structure is not predicted to change[29]. The location of the rs895819 SNP in the centre of the terminal loop of *mir-27a* is likely to decrease the size of the loop and affect the binding of DROSHA, thus decreasing miRNA maturation29,46. Alternatively, the SNP might influence the binding affinity of several DROSHA inhibitors, such as Lin28 (REFS 47,48).

Another pre-miRNA SNP, rs6505162 in *pre-mir-423*, is associated with an increased risk of bladder cancer[35] and ovarian cancer in carriers of mutant *BRCA2* (REF. 45), and decreased risk of oesophageal cancer in Caucasians37. There is no clear explanation for the opposing effects of this SNP in different cancer types within the same population. Modulations of mature levels of miR-423 have not yet been functionally linked to this SNP, and the RNAFold software (see Further information) does not predict a change in secondary structure[45]. An A/C SNP in *mir-30c-2* was predicted to cause the greatest change in target gene identification, and as such was postulated to affect cancer risk[31], although results from a study in hepatocellular carcinoma (HCC) did not support this hypothesis[49].

A unique example of a functional miRNA SNP is rs2910164, which is located in the 3p strand of *mir-146a* (FIG. 3). This polymorphism involves a mispairing in the hairpin of the precursor, which leads to altered processing, lower expression of the mature sequence and predisposition to papillary thyroid carcinoma50. Intriguingly, individuals with a heterozygous genotype have a greater risk of developing papillary thyroid carcinoma than homozygous individuals[50]. Heterozygosity as a genetic risk factor is rare, and as postulated by the authors may be a form of genetic epistasis in which the phenotype of the heterozygote differs from the sum of the phenotypes of both alleles. Indeed, in a follow-up study, the same authors found that the SNP fell within the seed of the 3p strand, and would give rise to three mature miRNAs (one from the leading strand and two from the 3p strand) instead of the expected two observed in homozygous individuals[51]. The resulting complexity is underscored by the finding that each mature miRNA will target a specific repertoire of mRNAs51. Interestingly, there is evidence of homozygous to heterozygous somatic mutations at rs2910164 in several patients with papillary thyroid tumours50. These seminal studies elegantly decipher how a small genetic change can influence the gene expression profile and show that somatic mutations in miRNAs can be an oncogenic event. In separate studies, rs2910164 was associated with an increased risk of developing prostate cancer52 and HCC[53] in males through reduced expression levels of miR-146a[53]. Although not directly associated with breast cancer risk, rs2910164 was associated with a younger age of breast cancer diagnosis in familial breast cancer after adjustment for *BRCA1* and *BRCA2* mutation status[54]. The variant allele of rs2910164 leads to increased levels of mature miR-146a and binds with greater affinity to *BRCA1*. Predisposition may therefore develop through the downregulation of BRCA1 (REF. 54). Alternatively, rs2910146 could disrupt the well-documented role of miR-146a as a mediator of the pro-apoptotic transcription factor nuclear factor-κB (NF-κB)55,56. In support of this possibility, genes involved in the regulation of apoptosis were differentially transcribed in heterozygote rs2910146 carriers51.

Although studies have started to reveal the nature of the association between miRNA SNPs and cancer risk, several considerations remain: most of the studies used a candidate gene approach; of those that used a systematic approach, their lists are outdated owing to enhanced screening techniques that have identified new miRNA genes and updated builds of genome-wide SNP repositories. According to the most recent updates of miRNAs, 202 pre-miR genes have 283 SNPs (Supplementary information S1 (table)). It is likely that as new builds of dbSNP are formulated more SNPs with potential relevance to cancer risk will be identified. In addition,

the minor allele frequencies of many of the mir-SNPs already identified have not been determined. Therefore, population studies should be conducted that will ascertain whether or not these SNPs are polymorphic and if so in what populations. This is an important consideration, as data are emerging to suggest that some mir-SNPs have evolved to a high level of variance in distinct populations. For example, the variant alleles of several SNPs occur only in populations of African descent57. These epidemiological associations need to be validated in independent populations and functionally tested[58]. Furthermore, the inclusion of mir-SNPs in future genome-wide association studies (GWAs) will help to unveil low-penetrance susceptibility mutations. Most mir-SNPs are not included in current GWAS designs, and as such there is a paucity of information in this regard. The identification of tag sNPs for miRNA-related SNPs will also be a useful endeavour. Clarification of the extent of the pri region of miRNA genes is also needed to more accurately assess miRNA-related genetic variation. Many studies currently limit their analysis of mir-SNPs to the pre and mature regions, as these are clearly defined.

Box 1

Evolutionary selective pressures on mir-SNPs

From the seminal findings by Ambros[150], Ruvkun[151] and colleagues that *lin-4* (REF. 150) and the 3′ untranslated region of its target, lin-14 (REF 151), had blocks of conservation across *Caenorhabditis elegans* species, cross-species sequence conservation has become part of the standard used in prediction of both microRNA (miRNA) genes and their target sites. The widespread understanding that has stemmed from these and subsequent studies is that the functional importance of these sequences leads to selective pressure that limits the frequency of deleterious alleles. This is in fact the case, as single nucleotide polymorphism (SNP) density in miRNA genes and miRNA target sites has been found to be lower in human miRNA loci[27], as well as predicted miRNA binding sites, particularly seed sites, than their flanking regions30. SNPs in miRNA seed regions were predicted to occur at a frequency of 1%27. Conversely, given the constraint on sequence conservation as a means of prediction, it is not entirely surprising to find that SNP density is, in fact, lower in these regions. It is satisfying that evidence for negative selection has also been observed in human-specific miRNAs; that is, those that do not show cross-species conservation13. Although globally rare, mir-SNPs may be positively propagated, and so contribute to gene expression differences that become established in species. Such is the indication from miRNAs that are conserved in primates only[152] or that are human-specific[57], and there is evidence of local positive selection that may have contributed to expression diversity and adaptation, as well as higher-order thinking in humans. mir-SNPs may be another genetic modification that explains phenotypic differences in humans compared with other primates in the absence of high genotypic diversity. From the perspective of the molecular epidemiology of cancer, because mir-SNPs are so rare, their likelihood to be disruptive is higher.

SNPs in miRNA binding sites

The power of SNPs to affect phenotype is highlighted by the myriad of disorders and traits discovered in association with them[59–71]. In an analogous manner to seed region SNPs, an SNP in the 3′ UTR of a gene may create, as well as destroy, an miRNA binding site (FIG. 3). Disruption of miRNA-dependent regulation by SNPs in the miRNA binding site of target mRNAs is a bona fide mechanism for altered gene expression in cancer. Let-7 binds to the 3′ UTR of KRAS and regulates its expression[72], and both *let-7* and *KRAS* are implicated in lung carcinogenesis[73]. Chin *et al*.[74] sequenced the 3′ UTR of *KRAS* and identified ten Let-7 complementarity sites (LCS). A new SNP in LCS6 (now designated rs61764370) was present

in 20% of lung cancer cases and 5% of the control population (TABLE 2). Correspondingly, the SNP was associated with a 2.3-fold increased risk of developing lung cancer in moderate smokers. Having validated their findings in a second large cohort, the authors demonstrated that the SNP led to increased luciferase activity and decreased levels of Let-7 family members, especially Let-7b [75], suggesting the possibility of a negative feedback loop, similar to that between Let-7 and Lin28 (REFS 76,77). It is possible that the presence of the SNP along with mutant *KRAS* could lead to an amplified oncogenic hit, and might identify a group of patients who are at a particularly high risk of developing lung cancer[75]. Subsequently, this SNP has been associated with reduced survival in patients with oral cancer[78].

As illustrated in this example, several elements have to converge for an miRNA binding site SNP to be considered functional: the SNP must have a proven association with cancer, both the miRNA and its predicted target must be expressed in the tissue, and the allelic changes must result in differential binding of the miRNA and affect expression of the target gene. Choosing candidate genes to analyse in specific cancer types has been one favoured approach to find such interactions. In this manner, rs17281995 in the 3′ UTR of CD86 was predicted to disrupt the binding sites for five miRNAs and was found to be associated with an increased risk of colorectal cancer[79]. miR-582, which is expressed in normal colon tissue, bound less tightly to the variant allele of *CD86*, thus increasing its expression level. CD86 functions as a co-stimulatory molecule and increases the production of the pro-inflammatory cytokine interleukin-4 (IL-4)[80], which might explain the contribution of this SNP to colorectal cancer risk.

Among the 120,000 known SNPs that occur in 3′ UTRs, ~17% destroy putative conserved or non-conserved miRNA binding sites. Furthermore, 8.6% create new predicted target sites according to the Patrocles database[81] (see Further information; Supplementary Information S2 (box)). Yu *et al*.[82] searched the 3′ UTRs of all genes in the genome and cross-referenced this with dbSNP 126 and the Targetscan database (see Further information) for possible overlap. Interestingly, they observed that 12 of these SNPs were differentially expressed in cancer EST databases[82]. Subsequently, one of these SNPs, rs16917496 in SETD8, was associated with an increased risk of breast cancer[83]. Landi *et al*.[84] have now catalogued 79 SNPs in the 3′ UTRs of 129 colorectal cancer-associated genes. One of these, an insertion/deletion (indel) polymorphism (rs3783553) in the 3′ UTR of IL1A led to a 38% decrease in the risk of developing HCC. The TTCA insertion allele for rs3783553 disrupts a binding site for miR-122 and miR-378, thereby increasing transcription of *IL1A in vitro* and *in vivo* [85].

Saetrom *et al*.[86] mapped HapMap SNPs to putative miRNA recognition sites in genes deregulated in oestrogen receptor-stratified breast tumours and used local linkage disequilibrium patterns to identify high-ranking SNPs in the Cancer Genetic Markers of Susceptibility (CGeMS) breast cancer GWAS. Two SNPs, rs1970801 and rs1109745, were in strong linkage disequilibrium with rs1434536, an SNP in an miR-125b target site in the 3′ UTR of bone morphogenetic protein receptor type 1B (BMPR1B). Subsequently, rs1434536 was validated and miR-125b was shown to differentially regulate the C and T alleles. These results suggest that allele-specific regulation of *BMPR1B* by miR-125b explains the observed disease risk[86].

Drawing on the list of putative polymorphic miRNA binding sites in the genome provided by Chen *et al*.[87], 11 possible candidate SNPs were selected for their potential relevance to breast cancer[88]. Subsequently, rs2747648, which resides in a predicted binding site for three miRNAs in the oestrogen receptor-α (ESR1) gene, was associated with a 27% reduction in breast cancer risk in premenopausal women. When the C allele is present, miR-453 binds with greater affinity to *ESR1*, thus leading to decreased levels of eRα. Postmenopausal women already have reduced levels of endogenous oestrogen, perhaps explaining why this SNP is relevant only in premeno-

pausal women. Indeed, the authors pose the interesting question of whether or not carriers of the ancestral T allele would respond better to endocrine therapy, given that they will naturally express increased levels of the receptor[88]. Moreover, there is evidence that the variant allele of another SNP in *ESR1*, rs93410170, enhances the binding between miR-206 and the 3′ UTR, thereby decreasing eRα levels[89]. Subsequently, the authors postulated that the lower incidences of breast cancer in Hispanic and european populations could be partially associated with the increased prevalence of the variant T allele.

The idea that an SNP in an miRNA binding site could specifically affect pharmacokinetics is illustrated in a study that described the effect of a C–T SNP near the miR-24 binding site in the 3′ UTR of human dihydrofolate reductase (DHFR) on its translation. The variant allele interferes with miR-24 binding to the 3′ UTR and leads to a twofold increase in the mRNA half-life, DHFR overexpression and methotrexate resistance90. Moreover, a recent follow-up study found that DHFR imparts a selective growth advantage and neoplastic transformation in immortalized cells91. However, an important caveat of this study, and one that introduces an additional level of complexity when studying miRNA-related SNPs, is that the SNP that affected drug resistance was not located in the miRNA binding site. Rather, it was located further downstream. The relationship between miRNAs and pharmacogenomics was recently discussed92. The concept of 'integrative epidemiology' (REF. 93) proposes several models that depict how SNPs may overlap in their relationships to therapeutic response, susceptibility and prognosis94. An intriguing observation reminds us that multiple SNPs in the miRNA network may interact in a similar manner in carcinogenesis. *DHFR* is one of the predicted targets of miR-196a-2, the SNP of which, rs11614913, is associated with cancer risk43. Thymidylate synthase, the target of the cancer drug 5-fluorouracil, is another predicted target of miR-196a-2 (REF. 43). Indeed, recent studies have found that genetic variations in the 3′ UTR of thymidylate synthase are associated with resistance95,96. Whether rs11614913 affects expression of these genes and therapeutic efficacy is an intriguing question.

A recent paper provides new insight into mir-SNPs and their relevance to cancer. New miRNAs have been found in epstein–Barr virus, which is a causal factor for several cancers[97]. One of these miRNAs, mir-BART22, harbours a single genetic variant that increases its mature levels and downregulates its target, latent membrane protein 2A (LMP2A), a potent immunogenic viral antigen. As a consequence, it is possible that miR-BART22 could accelerate carcinogenesis through the downregulation of LMP2A and the evasion of the host immune response[98]. Whether similar interactions contribute to other viral-related cancers remains to be determined.

Although much less explored, miRNAs can also bind to target sites in the 5′ UTR and open reading frames[99]. Such binding sites can occasionally be interrupted by introns and therefore require splicing to bind with their complementary miRNAs[100]. Several lines of evidence support miRNA binding to the 5′ UTR[101–103] and coding sequences[9,11,104]. For example, miR-10a binds to the 5′ UTR of ribosomal proteins to increase their translation[105], and miR-148 regulates DNMT3B expression through a conserved site in the protein coding sequence. Interestingly, the target site is absent in the *DNMT3B3* splice variant. Therefore, the expression of miR-148 changes the relative abundance of *DNMT3B* splice variants[106]. Let-7 directly targets DICER1 in its coding sequence, thus establishing a mechanism for a miRNA–DICeR1 negative feedback loop[107]. This work suggests that the search for miRNA-related susceptibility loci should be expanded to include both the 5′ UTR and coding regions. Sensitive alleles identified in epidemiological studies, but with obscure functional roles, should perhaps be tested under miRNA prediction algorithms that are not limited to the 3′ UTR of genes, particularly if evidence indicates that altered expression of that gene can be associated with specific phenotypes.

Although the data on SNPs in miRNA binding sites and cancer risk are exciting, several limitations and caveats remain that may affect the field as it moves forwards. One of the major pitfalls in these studies is the ambiguity of computationally predicted miRNA binding sites. Programs such as Patrocles and PolymiRTS (see Further information; Supplementary information S2 (box)) intercalate and cross-reference this data with dbSNP information, and as such are invaluable in the search for polymorphic miRNA binding sites. However, miRNAs bind to their targets in a manner that is either 5′ dominant (perfect base pairing at 7–8 nucleotides in the 5′ end) or 3′ compensatory (imperfect 5′ binding, therefore the 3′ end has a stronger degree of complementarity)[108]. Algorithms that do not take this into account might miss SNPs that could be important for 3′ compensatory-based matching. Furthermore, the databases need to keep pace with the discovery of new SNPs and miRNA genes to maintain their relevance for researchers. Therefore, it is noteworthy that several SNPs[50,51,57,74], including the *KRAS* LCS6 SNP, were identified through direct sequencing. A major lesson from GWAS is that variants in regulatory regions are much more likely to cause disease than nonsynonymous coding SNPs, and as such miRNA binding regions should be considered in future GWAS approaches[109].

SNPs in miRNA-processing machinery

The global repression of miRNA maturation promotes cellular transformation and tumorigenesis[15], thus SNPs that affect the proteins involved in miRNA biogenesis may have deleterious effects on the miRNAome (FIG. 4). Furthermore, recent data suggest that SNPs in the biogenesis machinery are also linked to cancer risk, drug response and prognosis. Low levels of DROSHA, for example, are associated with poor cancer survival[110]. However, SNPs in *DROSHA* and *DGCR8* did not predispose to cancer susceptibility[35–37,111] (TABLE 3).

By contrast, the nuclear export proteins XPO5 and RAN have been associated with cancer risk[37]. XPO5 is responsible for miRNA nuclear export, and knocking down its expression leads to reduced miRNA levels[112]. XPO5 is downregulated in bronchioloalveolar carcinoma and stage 1 lung cancer[113] but upregulated in high-grade prostate cancer[114]. The SNPs in *XPO5* and *RAN* occur in their 3′ UTRs, suggesting that they might affect mRNA stability. Our analysis of the *RAN* SNP, rs14035, in the PolymiRTS database suggests that the ancestral allele lies in a binding site for miR-575, which is disrupted by the derived allele that in addition creates a binding site for miR-182*. Although these are *in silico* results, they raise the possibility that in addition to affecting cancer risk through the disruption of miRNA nuclear export, a more intricate pathway may be involved that includes miRNA regulation.

DICER1 and transactivation-responsive RNA-binding protein (TRBP) mediate pre-miRNA processing. A recent study indicated that DICeR1 functions as a haploinsufficient tumour suppressor in cancer[115]. Indeed, lower levels of *DICER1* mRNA have been associated with decreased cancer survival[110]. Interestingly, rs3742330 in the 3′ UTR of *DICER1* was associated with an increased risk of premalignant oral lesions in individuals with leukoplakia and/or erythroplakia[36]. Melo and colleagues[116] identified two frameshift mutations in TRBP that introduce premature stop codons, resulting in reduced TRBP expression. One function of TRBP is regulating DICER stability, thus these mutations resulted in reduced DICER expression and lower miRNA production and were associated with higher cellular proliferation levels[116].

RISC has a pivotal role in guiding single-stranded mature miRNA sequences to their target mRNA sites. The variant allele of the GEMIN3 nonsynonymous SNP, rs197412, was associated with a reduced risk of premalignant oral lesions, and rs197414, also in *GEMIN3*, was associated with an increased risk of bladder[35] and oesophageal cancer[37]. Therefore, *GEMIN3* variants could alter global miRNA homeostasis and have a major effect on cellular

signalling pathways. Two SNPs in GEMIN4, rs2740348 and rs7813, were associated with a decreased risk of renal cell carcinoma[111] and reduced transformation of Hep3B cells[117].

The tumour suppressor p53 was recently implicated in miRNA processing[118]. Through interaction with p68 and DROSHA, p53 facilitates pri-miRNA to pre-miRNA processing. Given the well-documented relationship between p53 mutations and cancer[119–121], it is possible that there might be p53 mutations or SNPs that affect miRNA processing and so increase or decrease the risk of cancer development. This also raises the possibility that p53-associated conditions, such as Li–Fraumeni syndrome, may relate to a global decrease in miRNA production and function.

On a cautionary note, although these associations are interesting and may have far-reaching implications, none of the studies of SNPs in miRNA processing machinery has been validated in independent studies, nor has the biological mechanisms of how they affect miRNA maturation and cancer been delineated.

Alternative miRNA structural variation

IsomiRs

A new form of miRNA sequence heterogeneity has recently been identified. Before their official annotation as 'isomiRs' by Morin *et al*.[122] in 2008, miRNA variants had been identified in cloning studies but were ambiguously misclassified as putative experimental error[123,124]. In cases in which several closely matching sequences were discovered for a single miRNA, the most commonly detected sequence was chosen as the reference[122,125,126]. So far, three main types of miRNA sequence modification have been described: 3′ deletion/addition, 5′ deletion/addition and internal modifications (BOX 2). Distinct from SNPs, the location of these variations suggests that they generally arise from variable cleavage sites for DROSHA and DICER1 in the hairpin. Their identification was possible owing to the application of next-generation deep 454 sequencing approaches to miRNA discovery.

The specificity of isomiRs as bona fide genetic variants of miRNAs, and not experimental artefacts, is strengthened by the detection of the new miRNA sequences and analogous nucleotide modifications in different genomes[122,125,126]; their detection using both a linker-based miRNA cloning approach[127] and massively parallel sequencing[122]; an observation that the frequencies of the nucleotide modifications were remarkably higher than the estimates attributed to sequencing errors; and the non-random positioning of the nucleotide changes[125,128]. Indeed, an analysis of the most prevalent 3′ additions in human and mouse tissues demonstrated that the nature of the nucleotide change is evolutionarily conserved[125].

Despite the large number of isomiRs detected, their role in post-transcriptional regulation remains to be experimentally determined. It is postulated that the modifications could affect miRNA half-life, subcellular localization and miRNA target specificity. IsomiRs resulting from variation at the 5′ end may be of particular interest, as they have different seed sequences from the reference miRNA, with the subsequent ability to potentially target different transcripts.

Several isomiRs have been implicated in cancer. In a mouse model of leukaemia, several isomiRs of mmu-miR-10a, mmu-miR-155, mmu-miR-27a, mmu-miR-27c, mmu-let-7a and mmu-miR-222 were differentially expressed[125]. Concordant with their reference sequences, most of the isomiRs were downregulated in the tumour cells. One isomiR of mmu-miR-223-5p was downregulated ~2,500-fold in metastases[125]. The isomiR count for members of the Let-7 family is among the highest detected[129]. For example, Let-7a-5p has 78 sequences derived from various combinations of 5′ and 3′ modification: some of these sequences had counts

greater than 4,000 and were therefore highly expressed. The mir-181 family of putative tumour suppressors and *mir-21*, an oncogene, also have a remarkable level of sequence variation. A full description of the detected isomiR sequences is provided by Morin *et al*.[122]. Although these variations have not been interrogated in human cancer, it is plausible that they are relevant to tumorigenesis and it is likely that navigating this mercurial maze may lead to many answers underlying both normal and cancer cell biology.

Epigenetics

miRNA expression can also be affected by epigenetic silencing. Indeed, many miRNAs are found in CpG islands (Supplementary information S1 (table)). epigenetic silencing of several miRNAs is a frequent and early event in breast cancer[130,131], and although the *let-7* family is globally downregulated in lung cancer[73,132] there is evidence of let-7a-3 hypomethylation133; this is perhaps another example of how miRNAs can have bivalent roles in malignancy[20]. A case–control analysis exploring a possible relationship between aberrant epigenetic profiles and cancer risk has yet to be instigated, and so the implications of this form of genetic variation at the population level are unknown. Moreover, it is likely that mir-SNPs in CpG islands might also affect the pattern of miRNA expression and contribute to cancer susceptibility. Consistent with this idea, an SNP occurring in the promoter of an miRNA (whether in a CpG island or not) would also be predicted to affect miRNA levels. Indeed, Sevignani *et al*.[23] found that most of the sequence differences in miRNA genes in tumour-susceptible mice rather than tumour-resistant mice occur in the promoters.

Alternative splicing, alternative cleavage and polyadenylation

Approximately 60% of human genes are thought to undergo alternative splicing[134]. The implications of this form of transcriptional control for miRNAs remain largely unexplored. However, there is evidence that proliferating cells have shorter 3′ UTRs, fewer miRNA binding sites and therefore diminished regulation by miRNAs[135], although the converse is also true[64]. Furthermore, miR-124 augments neural differentiation by targeting *PTBP1* mRNA, a global repressor of alternative pre-mRNA splicing in non-neuronal cells[136]. Sandberg *et al*. [135] found evidence of shorter 3′ UTRs in proliferating T cells mediated by alternative cleavage and polyadenylation. Moreover, it was subsequently shown that globally cancer cells have shorter 3′ UTRs than untransformed cells and so escape regulation by miRNAs — these shorter isoforms can give rise to tenfold more protein. In addition, the expression of the shorter isoform of the proto-oncogene insulin-like growth factor 2 mRNA binding protein 1 (IMP1) led to oncogenic transformation, although the longer form did not, thus showing that the loss of repressive elements in the mRNA sequence through alternative cleavage and polyadenylation promotes an oncogenic phenotype[137,138].

Conclusions and future perspectives

This Review has focused on the genetic variations known to occur in miRNAs, their binding sites and the genes that facilitate their processing. However, as the ‘miRNAome’ evolves, it is likely that new candidate SNPs and forms of genetic variation linked to cancer susceptibility will emerge. Furthermore, given the differential cell of origin for cancers arising from different anatomic sites, and the cell type specificity of miRNA transcriptomes, it is reasonable to assume that the effects of mir-SNPs will be modulated in a cell type-specific manner. The incorporation of miRNA target co-expression and expression Quantitative Trait Locus (eQTL) mapping should aid in deciding whether mir-SNP is functional. Such features are part of the databases Patrocles and PolymiRTS (Supplementary information S2 (box)).

Although candidate gene approaches can certainly ascertain the effect of a single SNP on an individual’s risk of cancer, the cumulative effect of the inheritance of multiple SNPs in miRNA-

related genes might augment risk. Consistent with this idea, an increased risk of oesophageal and bladder cancer was observed in individuals with SNPs in both miRNAs and miRNA processing genes[37]. Furthermore, some of those mir-SNPs might be in linkage disequilibrium and therefore inherited together. The polygenic model of inherited breast cancer purports that unfavourable combinations of polymorphic genetic variants in low-penetrance susceptibility genes contribute to the excess familial breast cancer risk and most of these genes have not yet been discovered[139–141]. As mir-SNPs are rare[27,30], and their minor allele frequencies are globally low, large studies will be needed to draw out their importance.

The complexity of the miRNA network is further intensified by the discovery of miRNA functions that fall outside their classic range. For example, there is evidence of miRNA-mediated increases in protein translation[6], nuclear import of miRNAs with distinctive hexanucleotide terminal motifs[142] and the secretion of miRNAs[143,144]. Furthermore, an alternative miRNA processing pathway has been uncovered in both *Drosophila melanogaster* and *Caenorhabditis elegans* that bypasses DROSHA and instead uses a splicing technique to generate miRNA precursors from short intronic sequences (mirtrons)[145–147]. How SNPs in miRNAs affect these pathways remains to be tested.

The data reviewed here explore the strong link between alterations in miRNA structure and function and inherited cancer risk. Although the pathways that mediate this risk have not been fully elucidated, there is a clear suggestion that cancer risk is mediated by changes in miRNA sequence and maturation. mir-SNPs affect cancer susceptibility, response to treatment and prognosis[75,78,148]. In addition to broadening our understanding of the astounding complexity of how miRNAs function, the study of genetic variation in miRNA networks has expanded our knowledge of the myriad ways in which miRNAs can affect cancer. The 3′ UTRs of genes are involved in multiple levels of regulation[87]. These oftneglected regions are now known to be prime regulators of the transcriptome, and the importance of SNPs in these regions for human traits is exemplified by the range of phenotypes affected by these mutations (for reviews see REFs 59,149). Historically, these regions were not extensively mined for SNP discovery, something that should now be addressed[74]. Moreover, it is clear that miRNA genetic miscellany can affect the diversity of the genome and is related to cancer susceptibility. Several approaches have been used to apply this genetic basis to cancer, and a polygenic, network-based approach should be adopted in the future140. The validation of these findings in multiple cohorts and the testing of their applicability to different ethnic populations is also required. Furthermore, the linkage of population-based studies to functional validation is crucial for both basic science and the advancement of these findings to clinical applications — this step should no longer be overlooked in epidemiological studies if we are to unravel the implications of these networks in human disease. Therefore, despite the initial insights covered in this Review, we believe that a vast anthology of knowledge remains to be discovered.

Box 2

Main forms of isomiR variation

3′ modification

The most prevalent type of modification noted among mature microRNA (miRNA) sequences is single nucleotide 3′ extensions[122,127,153,154]. These modifications produce an isomiR that matches the genome at every position except the terminal nucleotide. A 3′ extension was found in 66% of *mir-326* reads[129]. The nucleotides most commonly added were adenine and uridine, followed by cytosine and guanine[122,153]. Intriguingly, in the study by Kuchenbauer *et al.*[125] 151 miRNAs and miRNA*s had a 3′ variation that not did not match the genome, suggesting the possibility of an as yet unknown new mechanism of miRNA processing. The changes in terminal nucleotide were proposed to be partly due to

deamination of cytosine to uracil by cytidine deaminases (CDARs) or deamination of an adenosine to inosine by adenosine deaminases (ADARs).

Uridylation at the 3′ end of miRNAs has also been reported[76,155,156]. The biological importance of 3′ uridylation is unclear. It could mediate miRNA turnover or facilitate mRNA–miRNA binding in cases in which 3′ compensatory binding is predominant. This signal may also function as a degradation signal, perhaps in a manner that is analogous to protein ubiquitylation[156].

5′ modification

Contrary to the idea that pre-miRNA processing leads to a mature miRNA sequence with a fixed nucleotide at the 5′ end, recent data indicate that isomiRs may also result from variation at the 5′ end. They are of particular interest as they have a different seed sequence from the reference miRNA, and therefore have the potential to bind a different repertoire of targets. Not surprisingly, nucleotide modification to the 5′ end of mature miRNAs seems to be less likely than at the 3′ end. It is currently unclear whether these non-canonical variants associate in RNA-induced silencing complex (RISC). If so, the presence of isomiRs may have implications in future annotation of miRNAs and the development of new target prediction algorithms. Modification at the 5′ end of miRNAs has been noticed in T cells, in which distinct *mir-142* variants seem to regulate different target gene pools[157]. IsomiRs of *mir-142* seem to arise from shifting of the processing sites in the pri-miRNA sequence by DROSHA to generate alternative pre-miRNA variants that can then be independently processed by DICER1 to generate mature miRNAs that might have altered target specificity. The seed-matched target binding sites of the different miR-142 variants seem to be evolutionarily conserved[157] and these distinct miR-142 variants seem to regulate different 3′ UTR targets. It is also noteworthy that miR-142 is the most highly expressed miRNA in naive T cells. There is also evidence for 5′ end processing in *Caenorhabditis elegans*, mammals, viruses and *Drosophila melanogaster* [25,123,157,158]. Therefore, 5′ end modifications are a conserved phenomenon[157]. It is possible that many more miRNAs with 5′ shifts will be found in the future.

Internal modification

Using massively parallel sequencing in mouse ovary, Reid and colleagues[128] found evidence of internal editing of murine *let-7a* in the form of internal insertion, deletions and substitutions. There is a selection against nucleotide alterations in nucleotides 3–7 (the seed) and 10–15 (the cleavage and anchor sites), that is, those positions that generally have Watson–Crick binding. It has been speculated that the change in nucleotide sequence expands the target repertoire and/or enhances mRNA decay over translational repression by increasing or decreasing the degree of complementarity[128].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Intramural Research Program of the National Institute of Health, NCI-CCR.

Glossary

Fragile sites	Parts of a chromosome that are sensitive to break formation during metaphase when DNA replication is perturbed. The genes that lie within these regions are frequently deleted or rearranged in cancer.

Case–control study	An epidemiological study that compares two groups of individuals: those who have the condition under study (the cases) and those without the condition (the controls).
Passenger (3p) strand	Precursor miRNA sequences form a stem-loop structure. The single-stranded mature sequence lies at the 5′ end (5p strand). Generally the strand complementary to the mature miRNA at the 3′ end is degraded (3p strand) although in some cases it is not.
Minor allele frequency	The minor allele frequency of an sNP is the frequency of the least common allele in a population.
Genome-wide association studies (GWAS)	Large case–control studies in which genetic variation, in the form of sNPs, are examined across a genome to identify genetic associations with disease.
Tag SNPs	A genetic change that is in high linkage disequilibrium with other sNPs. The term ‘tag’ is used as these sNPs can be used to mark the genetic variations of all the sNPs they are associated with without sequencing all the sNPs. They are frequently used in genome-wide association studies.
Linkage disequilibrium	The non-random inheritance of alleles at two or more loci. The resulting haplotype is generally inherited from a single chromosome. Natural selection of a favourable phenotype can contribute to linkage disequilibrium between alleles.
Alternative splicing	Splicing is a post-transcriptional mechanism in which introns are removed and exons are joined together allowing the production of a specific protein product. Alternative splicing occurs when different combinations of exons (and introns) are cut together allowing genes to produce more than one mRNA isoform.
Expression Quantitative Trait Locus (eQTL) mapping	Quantitative trait loci (QTL) are regions of DNA that are closely linked to the genes that underlie the trait in question. Expression QTL (eQTL) are genetic loci that regulate gene expression traits. Because of the intricate association of miRNAs and gene expression, mir-SNPs are unique candidates for eQTL studies and eQTLs provide support for mir-SNP functionality.

References

1. Frazer KA, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851–861. [PubMed: 17943122] This paper provides a good background to HapMap and provides details of the most recent version of the database.
2. Kan YW, Dozy AM. Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relationship to sickle mutation. Proc. Natl Acad. Sci. USA 1978;75:5631–5635. [PubMed: 281713] A seminal study that describes the first detection of SNPs
3. International HapMap Consortium. A haplotype map of the human genome. Nature 2005;437:1299–1320. [PubMed: 16255080]
4. Bartel B. MicroRNAs directing siRNA biogenesis. Nature Struct. Mol. Biol 2005;12:569–571. [PubMed: 15999111] This is an excellent paper that covers miRNA biogenesis in a complete and descriptive manner.
5. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116:281–297. [PubMed: 14744438]

6. Vasudevan S, Tong Y, Steitz JA. Switching from repression to activation: microRNAs can up-regulate translation. Science 2007;318:1931–1934. [PubMed: 18048652]
7. Kim VN. MicroRNA biogenesis: coordinated cropping and dicing. Nature Rev. Mol. Cell Biol 2005;6:376–385. [PubMed: 15852042]
8. Lai EC. Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. Nature Genet 2002;30:363–364. [PubMed: 11896390]
9. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. Cell 2003;115:787–798. [PubMed: 14697198] One of the major studies used to delineate and define miRNA binding sites. The algorithms provided in this paper are widely used in target prediction programs.
10. Stark A, Brennecke J, Russell RB, Cohen SM. Identification of *Drosophila* MicroRNA targets. PLoS Biol 2003;1:E60. [PubMed: 14691535]
11. Grimson A, et al. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol. Cell 2007;27:91–105. [PubMed: 17612493]
12. Hobert O. Common logic of transcription factor and microRNA action. Trends Biochem. Sci 2004;29:462–468. [PubMed: 15337119]
13. Krek A, et al. Combinatorial microRNA target predictions. Nature Genet 2005;37:495–500. [PubMed: 15806104] Another of the very important descriptions of how miRNA binding sites can be predicted and the criteria one should consider.
14. Hon LS, Zhang Z. The roles of binding site arrangement and combinatorial targeting in microRNA repression of gene expression. Genome Biol 2007;8:R166. [PubMed: 17697356]
15. Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. Impaired microRNA processing enhances cellular transformation and tumorigenesis. Nature Genet 2007;39:673–677. [PubMed: 17401365]
16. Lu J, et al. MicroRNA expression profiles classify human cancers. Nature 2005;435:834–838. [PubMed: 15944708]
17. Garzon R, Calin GA, Croce CM. MicroRNAs in cancer. Annu. Rev. Med 2009;60:167–179. [PubMed: 19630570]
18. Kent OA, Mendell JT. A small piece in the cancer puzzle: microRNAs as tumor suppressors and oncogenes. Oncogene 2006;25:6188–6196. [PubMed: 17028598]
19. Esquela-Kerscher A, Slack FJ. Oncomirs — microRNAs with a role in cancer. Nature Rev. Cancer 2006;6:259–269. [PubMed: 16557279]
20. Fabbri M, Ivan M, Cimmino A, Negrini M, Calin GA. Regulatory mechanisms of microRNAs involvement in cancer. Expert Opin. Biol. Ther 2007;7:1009–1019. [PubMed: 17665990]
21. Calin GA, et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc. Natl Acad. Sci. USA 2002;99:15524–15529. [PubMed: 12434020] This paper provides the first demonstration that miRNA genes are deregulated in cancer.
22. Calin GA, et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc. Natl Acad. Sci. USA 2004;101:2999–3004. [PubMed: 14973191]
23. Sevignani C, et al. MicroRNA genes are frequently located near mouse cancer susceptibility loci. Proc. Natl Acad. Sci. USA 2007;104:8017–8022. [PubMed: 17470785]
24. Yang W, et al. Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. Nature Struct. Mol. Biol 2006;13:13–21. [PubMed: 16369484]
25. Seitz H, Ghildiyal M, Zamore PD. Argonaute loading improves the 5' precision of both MicroRNAs and their miRNA* strands in flies. Curr. Biol 2008;18:147–151. [PubMed: 18207740]
26. Abelson JF, et al. Sequence variants in SLITRK1 are associated with Tourette's syndrome. Science 2005;310:317–320. [PubMed: 16224024]
27. Saunders MA, Liang H, Li WH. Human polymorphism at microRNAs and microRNA target sites. Proc. Natl Acad. Sci. USA 2007;104:3300–3305. [PubMed: 17360642] This is an excellent paper that studies and describes the evolution of miRNA-related SNPs.
28. Duan R, Pak C, Jin P. Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA. Hum. Mol. Genet 2007;16:1124–1131. [PubMed: 17400653] This paper

provides a list of all the SNPs detected in miRNAs using older builds of the reference databases, including both pre and pri miRNA regions.

29. Wu M, et al. Genetic variations of microRNAs in human cancer and their effects on the expression of miRNAs. Carcinogenesis 2008;29:1710–1716. [PubMed: 18356149]
30. Chen K, Rajewsky N. Natural selection on human microRNA binding sites inferred from SNP data. Nature Genet 2006;38:1452–1456. [PubMed: 17072316]
31. Iwai N, Naraba H. Polymorphisms in human pre-miRNAs. Biochem. Biophys. Res. Commun 2005;331:1439–1444. [PubMed: 15883035]
32. Calin GA, et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N. Engl. J. Med 2005;353:1793–1801. [PubMed: 16251535]
33. Raveche ES, et al. Abnormal microRNA-16 locus with synteny to human 13q14 linked to CLL in NZB mice. Blood 2007;109:5079–5086. [PubMed: 17351108]
34. Yazici H, et al. Investigation of the miR16–11 (C > T) + 7 substitution in seven different types of cancer from three ethnic groups. J. Oncol. 2009 (doi:10.1155/2009/827532).
35. Yang H, et al. Evaluation of genetic variants in microRNA-related genes and risk of bladder cancer. Cancer Res 2008;68:2530–2537. [PubMed: 18381463]
36. Clague J, et al. Genetic variation in MicroRNA genes and risk of oral premalignant lesions. Mol. Carcinog 2009;49:183–189. [PubMed: 19851984]
37. Ye Y, et al. Genetic variations in microRNA-related genes are novel susceptibility loci for esophageal cancer risk. Cancer Prev. Res. (Phila. Pa.) 2008;1:460–469.
38. Hu Z, et al. Common genetic variants in pre-microRNAs were associated with increased risk of breast cancer in Chinese women. Hum. Mutat 2009;30:79–84. [PubMed: 18634034]
39. Tian T, et al. A functional genetic variant in microRNA-196a2 is associated with increased susceptibility of lung cancer in Chinese. Cancer Epidemiol. Biomarkers Prev 2009;18:1183–1187. [PubMed: 19293314]
40. Hoffman AE, et al. microRNA miR-196a-2 and breast cancer: a genetic and epigenetic association study and functional analysis. Cancer Res 2009;69:5970–5977. [PubMed: 19567675]
41. Peng S, et al. Association of MicroRNA-196a-2 Gene Polymorphism with Gastric Cancer Risk in a Chinese Population. Dig Dis. Sci 2009 Oct;16 (doi:10.1007/ s10620-009-1007-x).
42. Gottwein E, Cai X, Cullen BR. Expression and function of microRNAs encoded by Kaposi's sarcoma-associated herpesvirus. Cold Spring Harb. Symp. Quant. Biol 2006;71:357–364. [PubMed: 17381317]
43. Hu Z, et al. Genetic variants of miRNA sequences and non-small cell lung cancer survival. J. Clin. Invest 2008;118:2600–2608. [PubMed: 18521189]
44. Yang R, et al. A genetic variant in the pre-miR-27a oncogene is associated with a reduced familial breast cancer risk. Breast Cancer Res. Treat 2009 Nov;17 (doi: 10.1007/s10549-009-0633-5).
45. Kontorovich T, Levy A, Korostishevsky M, Nir U, Friedman E. SNPs in miRNA binding sites and miRNA genes as breast/ovarian cancer risk modifiers in Jewish high risk women. Int. J. Cancer 2009 Nov;30 [epub ahead of print].
46. Zeng Y, Cullen BR. Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. J. Biol. Chem 2005;280:27595–27603. [PubMed: 15932881]
47. Piskounova E, et al. Determinants of microRNA processing inhibition by the developmentally regulated RNA-binding protein Lin28. J. Biol. Chem 2008;283:21310–21314. [PubMed: 18550544]
48. Newman MA, Thomson JM, Hammond SM. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. RNA 2008;14:1539–1549. [PubMed: 18566191]
49. Yang J, et al. Analysis of sequence variations in 59 microRNAs in hepatocellular carcinomas. Mutat. Res 2008;638:205–209. [PubMed: 17900631]
50. Jazdzewski K, et al. Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. Proc. Natl Acad. Sci. USA 2008;105:7269–7274. [PubMed: 18474871]
51. Jazdzewski K, et al. Polymorphic mature microRNAs from passenger strand of pre-miR-146a contribute to thyroid cancer. Proc. Natl Acad. Sci. USA 2009;106:1502–1505. [PubMed: 19164563] This reference, and reference 50, are elegant studies that decipher how an SNP in an miRNA can

influence cancer susceptibility and provide insight into how to conduct these types of studies in the future.

52. Xu B, et al. A functional polymorphism in Pre-miR-146a gene is associated with prostate cancer risk and mature miR-146a expression *in vivo*. Prostate 2009;70:467–472. [PubMed: 19902466]
53. Xu T, et al. A functional polymorphism in the miR-146a gene is associated with the risk for hepatocellular carcinoma. Carcinogenesis 2008;29:2126–2131. [PubMed: 18711148]
54. Shen J, et al. A functional polymorphism in the miR-146a gene and age of familial breast/ovarian cancer diagnosis. Carcinogenesis 2008;29:1963–1966. [PubMed: 18660546]
55. Taganov KD, Boldin MP, Chang KJ, Baltimore D. NF-κB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. Proc. Natl Acad. Sci. USA 2006;103:12481–12486. [PubMed: 16885212]
56. Perry MM, et al. Rapid changes in microRNA-146a expression negatively regulate the IL-1β-induced inflammatory response in human lung alveolar epithelial cells. J. Immunol 2008;180:5689–5698. [PubMed: 18390754]
57. Quach H, et al. Signatures of purifying and local positive selection in human miRNAs. Am. J. Hum. Genet 2009;84:316–327. [PubMed: 19232555]
58. Chanock SJ, et al. Replicating genotype-phenotype associations. Nature 2007;447:655–660. [PubMed: 17554299]
59. Sethupathy P, Collins FS. MicroRNA target site polymorphisms and human disease. Trends Genet 2008;24:489–497. [PubMed: 18778868]
60. Jensen KP, et al. A common polymorphism in serotonin receptor 1B mRNA moderates regulation by miR-96 and associates with aggressive human behaviors. Mol. Psychiatry 2009;14:381–399. [PubMed: 18283276]
61. Martin MM, Lee EJ, Buckenberger JA, Schmittgen TD, Elton TS. MicroRNA-155 regulates human angiotensin II type 1 receptor expression in fibroblasts. J. Biol. Chem 2006;281:18277–18284. [PubMed: 16675453]
62. Sethupathy P, et al. Human microRNA-155 on chromosome 21 differentially interacts with its polymorphic target in the AGTR1 3' untranslated region: a mechanism for functional single-nucleotide polymorphisms related to phenotypes. Am. J. Hum. Genet 2007;81:405–413. [PubMed: 17668390]
63. Tan Z, et al. Allele-specific targeting of microRNAs to HLA-G and risk of asthma. Am. J. Hum. Genet 2007;81:829–834. [PubMed: 17847008]
64. Tan S, et al. Retained introns increase putative microRNA targets within 3' UTRs of human mRNA. FEBS Lett 2007;581:1081–1086. [PubMed: 17320082]
65. Lv K, et al. Allele-specific targeting of hsa-miR-657 to human IGF2R creates a potential mechanism underlying the association of ACAA-insertion/deletion polymorphism with type 2 diabetes. Biochem. Biophys. Res. Commun 2008;374:101–105. [PubMed: 18602895]
66. Tay Y, et al. Insights into the regulation of a common variant of HMGA2 associated with human height during embryonic development. Stem Cell Rev 2009;5:328–333. [PubMed: 20058197]
67. Weedon MN, et al. A common variant of HMGA2 is associated with adult and childhood height in the general population. Nature Genet 2007;39:1245–1250. [PubMed: 17767157]
68. Conner TS, et al. Functional polymorphisms in the serotonin 1B receptor gene (HTR1B) predict self-reported anger and hostility among young men. Am. J. Med. Genet. B Neuropsychiatr. Genet 2010;153B:67–78. [PubMed: 19350534]
69. Huang W, Li MD. Differential allelic expression of dopamine D1 receptor gene (*DRD1*) is modulated by microRNA miR-504. Biol. Psychiatry 2009;65:702–705. [PubMed: 19135651]
70. Cargill EJ, Nissing NJ, Grosz MD. Single nucleotide polymorphisms concordant with the horned/polled trait in Holsteins. BMC Res. Notes 2008;1:128. [PubMed: 19063733]
71. Clop A, et al. A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. Nature Genet 2006;38:813–818. [PubMed: 16751773]
72. Johnson SM, et al. RAS is regulated by the let-7 microRNA family. Cell 2005;120:635–647. [PubMed: 15766527]

73. Takamizawa J, et al. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Res 2004;64:3753–3756. [PubMed: 15172979]
74. Chin LJ, et al. A SNP in a let-7 microRNA complementary site in the KRAS 3' untranslated region increases non-small cell lung cancer risk. Cancer Res 2008;68:8535–8540. [PubMed: 18922928] This is one of the only studies on miRNA-related SNPs to assemble several lines of epidemiological and functional evidence that relates miR-SNPs to cancer.
75. Nelson HH, et al. KRAS mutation, KRAS-LCS6 polymorphism, and non-small cell lung cancer. Lung Cancer 2009 Oct;24 (doi:10.1016/j. lungcan.2009.09.008).
76. Heo I, et al. Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Mol. Cell 2008;32:276–284. [PubMed: 18951094]
77. Viswanathan SR, Daley GQ, Gregory RI. Selective blockade of microRNA processing by Lin28. Science 2008;320:97–100. [PubMed: 18292307]
78. Christensen BC, et al. A let-7 microRNA-binding site polymorphism in the KRAS 3' UTR is associated with reduced survival in oral cancers. Carcinogenesis 2009;30:1003–1007. [PubMed: 19380522]
79. Landi D, et al. Polymorphisms within micro-RNA-binding sites and risk of sporadic colorectal cancer. Carcinogenesis 2008;29:579–584. [PubMed: 18192692]
80. Freeman GJ, et al. B7-1 and B7-2 do not deliver identical costimulatory signals, since B7-2 but not B7-1 preferentially costimulates the initial production of IL-4. Immunity 1995;2:523–532. [PubMed: 7538442]
81. Georges M, et al. Polymorphic microRNA-target interactions: a novel source of phenotypic variation. Cold Spring Harb. Symp. Quant. Biol 2006;71:343–350. [PubMed: 17381315] This paper describes how the online tool Patrocles was developed and is currently organised.
82. Yu Z, et al. Aberrant allele frequencies of the SNPs located in microRNA target sites are potentially associated with human cancers. Nucleic Acids Res 2007;35:4535–4541. [PubMed: 17584784]
83. Song F, et al. An miR-502-binding site single-nucleotide polymorphism in the 3'-untranslated region of the *SET8* gene is associated with early age of breast cancer onset. Clin. Cancer Res 2009;15:6292–6300. [PubMed: 19789321]
84. Landi D, Gemignani F, Barale R, Landi S. A catalog of polymorphisms falling in microRNA-binding regions of cancer genes. DNA Cell Biol 2008;27:35–43. [PubMed: 17941804]
85. Gao Y, et al. An insertion/deletion polymorphism at miRNA-122-binding site in the interleukin-1alpha 3' untranslated region confers risk for hepatocellular carcinoma. Carcinogenesis 2009;30:2064–2069. [PubMed: 19917630]
86. Saetrom P, et al. A risk variant in an miR-125b binding site in BMPR1B is associated with breast cancer pathogenesis. Cancer Res 2009;69:7459–7465. [PubMed: 19738052]
87. Chen JM, Ferec C, Cooper DN. A systematic analysis of disease-associated variants in the 3' regulatory regions of human protein-coding genes I: general principles and overview. Hum. Genet 2006;120:1–21. [PubMed: 16645853]
88. Tchatchou S, et al. A variant affecting a putative miRNA target site in estrogen receptor (ESR) 1 is associated with breast cancer risk in premenopausal women. Carcinogenesis 2009;30:59–64. [PubMed: 19028706]
89. Adams BD, Furneaux H, White BA. The micro-ribonucleic acid (miRNA) miR-206 targets the human estrogen receptor-α (ERα) and represses ERα messenger RNA and protein expression in breast cancer cell lines. Mol. Endocrinol 2007;21:1132–1147. [PubMed: 17312270]
90. Mishra PJ, Humeniuk R, Longo-Sorbello GS, Banerjee D, Bertino JR. A miR-24 microRNA binding-site polymorphism in dihydrofolate reductase gene leads to methotrexate resistance. Proc. Natl Acad. Sci. USA 2007;104:13513–13518. [PubMed: 17686970] This is one of the first studies to describe the relationship between miR-SNPs and pharmacogenomics.
91. Mishra PJ, et al. MiR-24 tumor suppressor activity is regulated independent of p53 and through a target site polymorphism. PLoS ONE 2009;4:e8445. [PubMed: 20041160]
92. Wurdinger T, Costa FF. Molecular therapy in the microRNA era. Pharmacogenomics J 2007;7:297–304. [PubMed: 17189960]
93. Spitz MR, Wu X, Mills G. Integrative epidemiology: from risk assessment to outcome prediction. J. Clin. Oncol 2005;23:267–275. [PubMed: 15637390]

94. Savas S, Liu G. Genetic variations as cancer prognostic markers: review and update. Hum. Mutat 2009;30:1369–1377. [PubMed: 19639655]

95. Mandola MV, et al. A 6 bp polymorphism in the thymidylate synthase gene causes message instability and is associated with decreased intratumoral TS mRNA levels. Pharmacogenetics 2004;14:319–327. [PubMed: 15115918]

96. Lu JW, et al. Polymorphism in the 3'-untranslated region of the thymidylate synthase gene and sensitivity of stomach cancer to fluoropyrimidine-based chemotherapy. J. Hum. Genet 2006;51:155–160. [PubMed: 16424979]

97. Murphy G, Pfeiffer R, Camargo MC, Rabkin CS. Meta-analysis shows that prevalence of Epstein-Barr virus-positive gastric cancer differs based on sex and anatomic location. Gastroenterology 2009;137:824–833. [PubMed: 19445939]

98. Lung RW, et al. Modulation of LMP2A expression by a newly identified Epstein-Barr virus-encoded microRNA miR-BART22. Neoplasia 2009;11:1174–1184. [PubMed: 19881953] This paper describes how an Epstein–Barr virus-encoded miRNA regulates the expression on an immunogenic protein and in doing so describes a principle that can be applied to other virally related cancers.

99. Lytle JR, Yario TA, Steitz JA. Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR. Proc. Natl Acad. Sci. USA 2007;104:9667–9672. [PubMed: 17535905]

100. Ehrenreich IM, Purugganan MD. Sequence variation of MicroRNAs and their binding sites in *Arabidopsis*. Plant Physiol 2008;146:1974–1982. [PubMed: 18305205]

101. Kloosterman WP, Wienholds E, Ketting RF, Plasterk RH. Substrate requirements for let-7 function in the developing zebrafish embryo. Nucleic Acids Res 2004;32:6284–6291. [PubMed: 15585662]

102. Sunkar R, Zhu JK. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. Plant Cell 2004;16:2001–2019. [PubMed: 15258262]

103. Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. Science 2005;309:1577–1581. [PubMed: 16141076]

104. Tay Y, Zhang J, Thomson AM, Lim B, Rigoutsos I. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. Nature 2008;455:1124–1128. [PubMed: 18806776] This paper neatly describes how miRNAs can target the coding regions of important biological genes. Moreover, it shows how variation in sequence structure can affect miRNA binding and that SNPs in coding regions might be viewed as more relevant in future studies.

105. Andersson MK, et al. The multifunctional FUS, EWS and TAF15 proto-oncoproteins show cell type-specific expression patterns and involvement in cell spreading and stress response. BMC Cell Biol 2008;9:37. [PubMed: 18620564]

106. Duursma AM, Kedde M, Schrier M, le Sage C, Agami R. miR-148 targets human DNMT3b protein coding region. RNA 2008;14:872–877. [PubMed: 18367714]

107. Forman JJ, Legesse-Miller A, Coller HA. A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. Proc. Natl Acad. Sci. USA 2008;105:14879–14884. [PubMed: 18812516]

108. Brennecke J, Stark A, Russell RB, Cohen SM. Principles of microRNA-target recognition. PLoS Biol 2005;3:e85. [PubMed: 15723116]

109. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661–678. [PubMed: 17554300]

110. Merritt WM, et al. Dicer, Drosha, and outcomes in patients with ovarian cancer. N. Engl. J. Med 2008;359:2641–2650. [PubMed: 19092150] The first paper to show that DICER and DROSHA expression were related to cancer survival.

111. Horikawa Y, et al. Single nucleotide polymorphisms of microRNA machinery genes modify the risk of renal cell carcinoma. Clin. Cancer Res 2008;14:7956–7962. [PubMed: 19047128]

112. Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U. Nuclear export of microRNA precursors. Science 2004;303:95–98. [PubMed: 14631048]

113. Chiosea S, et al. Overexpression of Dicer in precursor lesions of lung adenocarcinoma. Cancer Res 2007;67:2345–2350. [PubMed: 17332367]

114. Chiosea S, et al. Up-regulation of dicer, a component of the MicroRNA machinery, in prostate adenocarcinoma. Am. J. Pathol 2006;169:1812–1820. [PubMed: 17071602]

115. Kumar MS, et al. Dicer1 functions as a haploinsufficient tumor suppressor. Genes Dev 2009;23:2700–2704. [PubMed: 19903759]
116. Melo SA, et al. A TARBP2 mutation in human cancer impairs microRNA processing and DICER1 function. Nature Genet 2009;41:365–370. [PubMed: 19219043]
117. Wan D, et al. Two variants of the human hepatocellular carcinoma-associated *HCAP1* gene and their effect on the growth of the human liver cancer cell line Hep3B. Genes Chromosom. Cancer 2004;39:48–58. [PubMed: 14603441]
118. Suzuki HI, et al. Modulation of microRNA processing by p53. Nature 2009;460:529–533. [PubMed: 19626115]
119. Mechanic LE, et al. Polymorphisms in XPD and TP53 and mutation in human lung cancer. Carcinogenesis 2005;26:597–604. [PubMed: 15564288]
120. Hollstein M, Sidransky D, Vogelstein B, Harris CC. p53 mutations in human cancers. Science 1991;253:49–53. [PubMed: 1905840]
121. Whibley C, Pharoah PD, Hollstein M. p53 polymorphisms: cancer implications. Nature Rev. Cancer 2009;9:95–107. [PubMed: 19165225]
122. Morin RD, et al. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Res 2008;18:610–621. [PubMed: 18285502] This paper describes annotated miRNA sequence variations as ‘isomiRs’. It is a very detailed analysis of isomiR differences between human embryonic stem cells and embroid bodies. Moreover, it has useful sequence information.
123. Ruby JG, et al. Large-scale sequencing reveals 21 U-RNAs and additional microRNAs and endogenous siRNAs. C. elegans. Cell 2006;127:1193–1207.
124. Landgraf P, et al. A mammalian microRNA expression atlas based on small RNA library sequencing. Cell 2007;129:1401–1414. [PubMed: 17604727]
125. Kuchenbauer F, et al. In-depth characterization of the microRNA transcriptome in a leukemia progression model. Genome Res 2008;18:1787–1797. [PubMed: 18849523]
126. Ebhardt HA, et al. Meta-analysis of small RNA-sequencing errors reveals ubiquitous post-transcriptional RNA modifications. Nucleic Acids Res 2009;37:2461–2470. [PubMed: 19255090] Although they are were discovered relatively recently, this paper provides intriguing evidence that 3′ and 5′ variations of miRNAs may affect their function and subcellular localization.
127. Fu H, et al. Identification of human fetal liver miRNAs by a novel method. FEBS Lett 2005;579:3849–3854. [PubMed: 15978578]
128. Reid JG, et al. Mouse let-7 miRNA populations exhibit RNA editing that is constrained in the 5’-seed/ cleavage/anchor regions and stabilize predicted mmu-let-7a:mRNA duplexes. Genome Res 2008;18:1571–1581. [PubMed: 18614752]
129. Thomson JM, et al. Extensive post-transcriptional regulation of microRNAs and its implications for cancer. Genes Dev 2006;20:2202–2207. [PubMed: 16882971]
130. Chuang JC, Jones PA. Epigenetics and microRNAs. Pediatr. Res 2007;61 24R–29R.
131. Lehmann U, et al. Epigenetic inactivation of microRNA gene hsa-mir-9-1 in human breast cancer. J. Pathol 2008;214:17–24. [PubMed: 17948228]
132. Yanaihara N, et al. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. Cancer Cell 2006;9:189–198. [PubMed: 16530703]
133. Brueckner B, et al. The human let-7a-3 locus contains an epigenetically regulated microRNA gene with oncogenic function. Cancer Res 2007;67:1419–1423. [PubMed: 17308078]
134. Clark TA, et al. Discovery of tissue-specific exons using comprehensive human exon microarrays. Genome Biol 2007;8:R64. [PubMed: 17456239]
135. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB. Proliferating cells express mRNAs with shortened 3’ untranslated regions and fewer microRNA target sites. Science 2008;320:1643–1647. [PubMed: 18566288] The authors describe the effects of alternative splicing of mRNA on miRNA function, showing that proliferating cells have shorter 3′ UTRs and therefore less miRNA regulation.
136. Makeyev EV, Zhang J, Carrasco MA, Maniatis T. The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. Mol. Cell 2007;27:435–448. [PubMed: 17679093]

137. Mayr C, Bartel DP. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell 2009;138:673–684. [PubMed: 19703394]
138. Passetti F, Ferreira CG, Costa FF. The impact of microRNAs and alternative splicing in pharmacogenomics. Pharmacogenomics J 2009;9:1–13. [PubMed: 19156160]
139. Ponder BA. Cancer genetics. Nature 2001;411:336–341. [PubMed: 11357140]
140. Pharoah PD, Dunning AM, Ponder BA, Easton DF. Association studies for finding cancer-susceptibility genetic variants. Nature Rev. Cancer 2004;4:850–860. [PubMed: 15516958]
141. Pharoah PD, et al. Polygenic susceptibility to breast cancer and implications for prevention. Nature Genet 2002;31:33–36. [PubMed: 11984562]
142. Hwang HW, Wentzel EA, Mendell JT. A hexanucleotide element directs microRNA nuclear import. Science 2007;315:97–100. [PubMed: 17204650]
143. Valadi H, et al. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nature Cell Biol 2007;9:654–659. [PubMed: 17486113] The first evidence that miRNAs can be packaged into exosome microvesicles is described here.
144. Gibbings DJ, Ciaudo C, Erhardt M, Voinnet O. Multivesicular bodies associate with components of miRNA effector complexes and modulate miRNA activity. Nature Cell Biol 2009;11:1143–1149. [PubMed: 19684575]
145. Berezikov E, et al. Evolutionary flux of canonical microRNAs and mirtrons in *Drosophila*. Nature Genet 2010;42:6–9. [PubMed: 20037610]
146. Okamura K, Chung WJ, Lai EC. The long and short of inverted repeat genes in animals: microRNAs, mirtrons and hairpin RNAs. Cell Cycle 2008;7:2840–2845. [PubMed: 18769156]
147. Ruby JG, Jan CH, Bartel DP. Intronic microRNA precursors that bypass Drosha processing. Nature 2007;448:83–86. [PubMed: 17589500]
148. Brendle A, et al. Polymorphisms in predicted microRNA-binding sites in integrin genes and breast cancer: ITGB4 as prognostic marker. Carcinogenesis 2008;29:1394–1399. [PubMed: 18550570]
149. Chen K, et al. Polymorphisms in microRNA targets: a gold mine for molecular epidemiology. Carcinogenesis 2008;29:1306–1311. [PubMed: 18477647]
150. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 1993;75:843–854. [PubMed: 8252621]
151. Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation. C. elegans. Cell 1993;75:855–862.
152. Berezikov E, et al. Diversity of microRNAs in human and chimpanzee brain. Nature Genet 2006;38:1375–1377. [PubMed: 17072315]
153. Aravin A, Tuschl T. Identification and characterization of small RNAs involved in RNA silencing. FEBS Lett 2005;579:5830–5840. [PubMed: 16153643]
154. Stark MS, et al. Characterization of the Melanoma miRNAome by Deep Sequencing. PLoS ONE 2010;5:e9685. [PubMed: 20300190]
155. Heo I, et al. TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell 2009;138:696–708. [PubMed: 19703396]
156. Li J, Yang Z, Yu B, Liu J, Chen X. Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in *Arabidopsis*. Curr. Biol 2005;15:1501–1507. [PubMed: 16111943]
157. Wu H, et al. miRNA profiling of naive, effector and memory CD8 T cells. PLoS ONE 2007;2:e1020. [PubMed: 17925868]
158. Waidner LA, et al. MicroRNAs of Gallid and Meleagrid herpesviruses show generally conserved genomic locations and are virus-specific. Virology 2009;388:128–136. [PubMed: 19328516]
159. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 2006;34:D140–D144. [PubMed: 16381832]
160. O'Toole AS, Miller S, Haines N, Zink MC, Serra MJ. Comprehensive thermodynamic analysis of 3' double-nucleotide overhangs neighboring Watson-Crick terminal base pairs. Nucleic Acids Res 2006;34:3338–3344. [PubMed: 16820533]

161. Qi P, et al. Association of a variant in MIR 196A2 with susceptibility to hepatocellular carcinoma in male Chinese patients with chronic hepatitis B virus infection. Hum. Immunol 2010 Mar;12 (doi: 10.1016/j.humimm.2010.02.017).

162. Dou T, et al. A polymorphism of microRNA196a genome region was associated with decreased risk of glioma in Chinese population. J. Cancer Res. Clin. Oncol 2010 Mar;14 (doi: 10.1007/s00432-010-0844-5).

163. Chen S, et al. An insertion/deletion polymorphism in the 3′ untranslated region of beta-transducin repeat-containing protein (betaTrCP) is associated with susceptibility for hepatocellular carcinoma in Chinese. Biochem. Biophys. Res. Commun 2010;391:552–556. [PubMed: 19931512]

164. Sun J, et al. Sequence variants at 22q13 are associated with prostate cancer risk. Cancer Res 2009;69:10–15. [PubMed: 19117981]

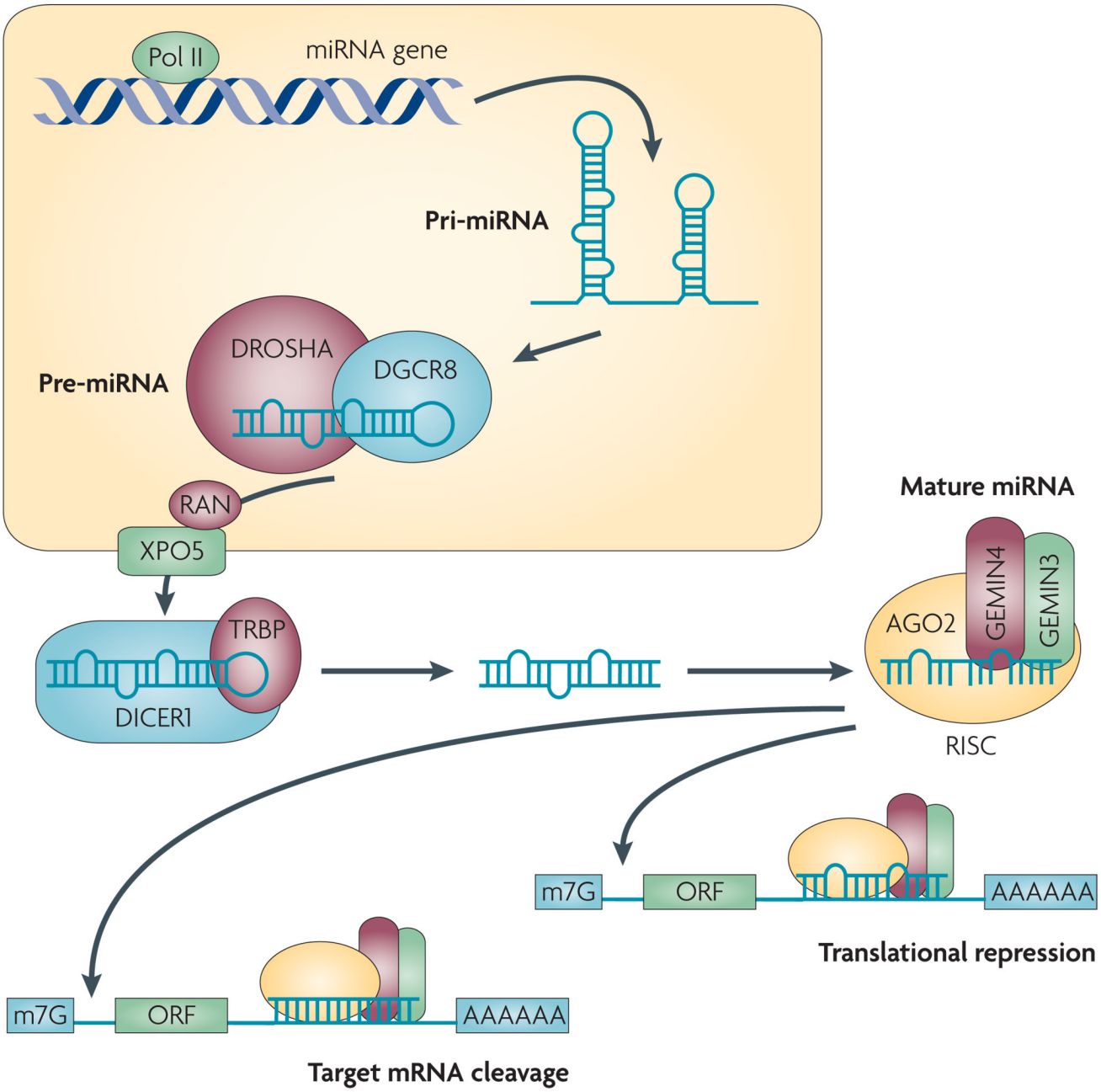

Figure 1. Illustrative overview of the miRNA network

RNA polymerase II (Pol II) produces a 500–3,000 nucleotide transcript, called the primary microRNA (miRNA), or pri-miRNA, that is then cropped to form a pre-miRNA hairpin by a multi-protein complex that includes DROSHA (~60–100 nucleotides) (a simplified view is shown here). This double-stranded hairpin structure is exported from the nucleus by RAN GTPase and exportin 5 (XPO5)112. Finally, the pre-miRNA is cleaved by DICER1 to produce two miRNA strands, a mature miRNA sequence, approximately 20 nucleotides in length, and its short-lived complementary sequence, which is denoted miR* and sometimes called the passenger strand or 3p strand159. The thermodynamic stability of the miRNA duplex termini and the identity of the nucleotides in the 3′ overhang determines which of the strands is incorporated into the RNA-inducing silencing complex (RISC)160. In some cases, in which

both the lead and passenger strands have a similar thermodynamic stability, both strands will be loaded. The single stranded miRNA is incorporated into RISC, which then targets it to the target 3′ untranslated region mRNA sequence to facilitate repression and cleavage. AA, poly A tail; m7G, 7-methylguanosine cap; ORF, open reading frame.

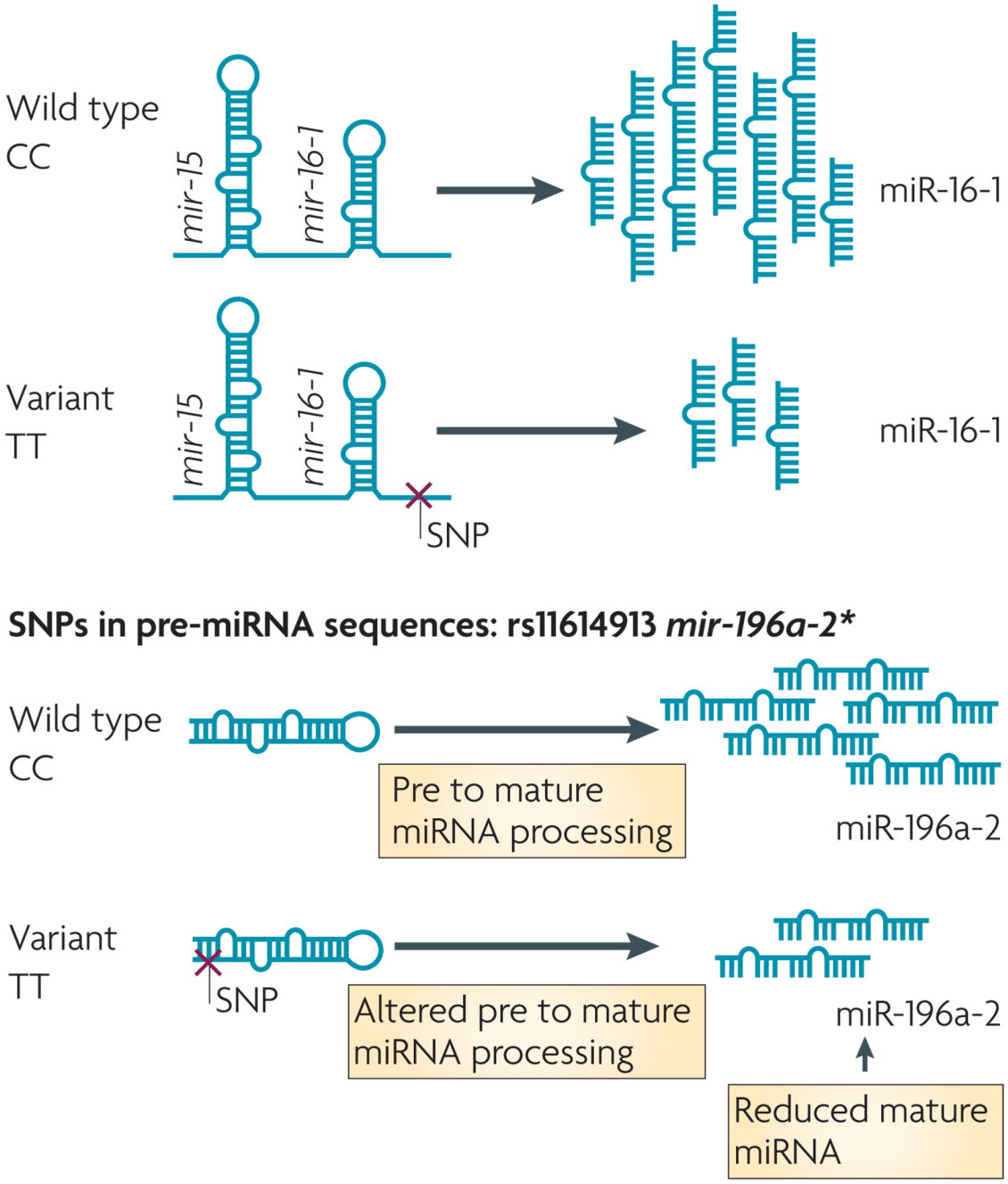

Figure 2. Diagrammatic representation of SNPs in pri-miRNA and pre-miRNA sequences

Single nucleotide polymorphisms (SNPs) can occur in the pri-miRNA and pre-miRNA strands and are likely to affect miRNA processing and subsequent mature miRNA levels. Such SNPs can lead to either an increase or decrease in processing.

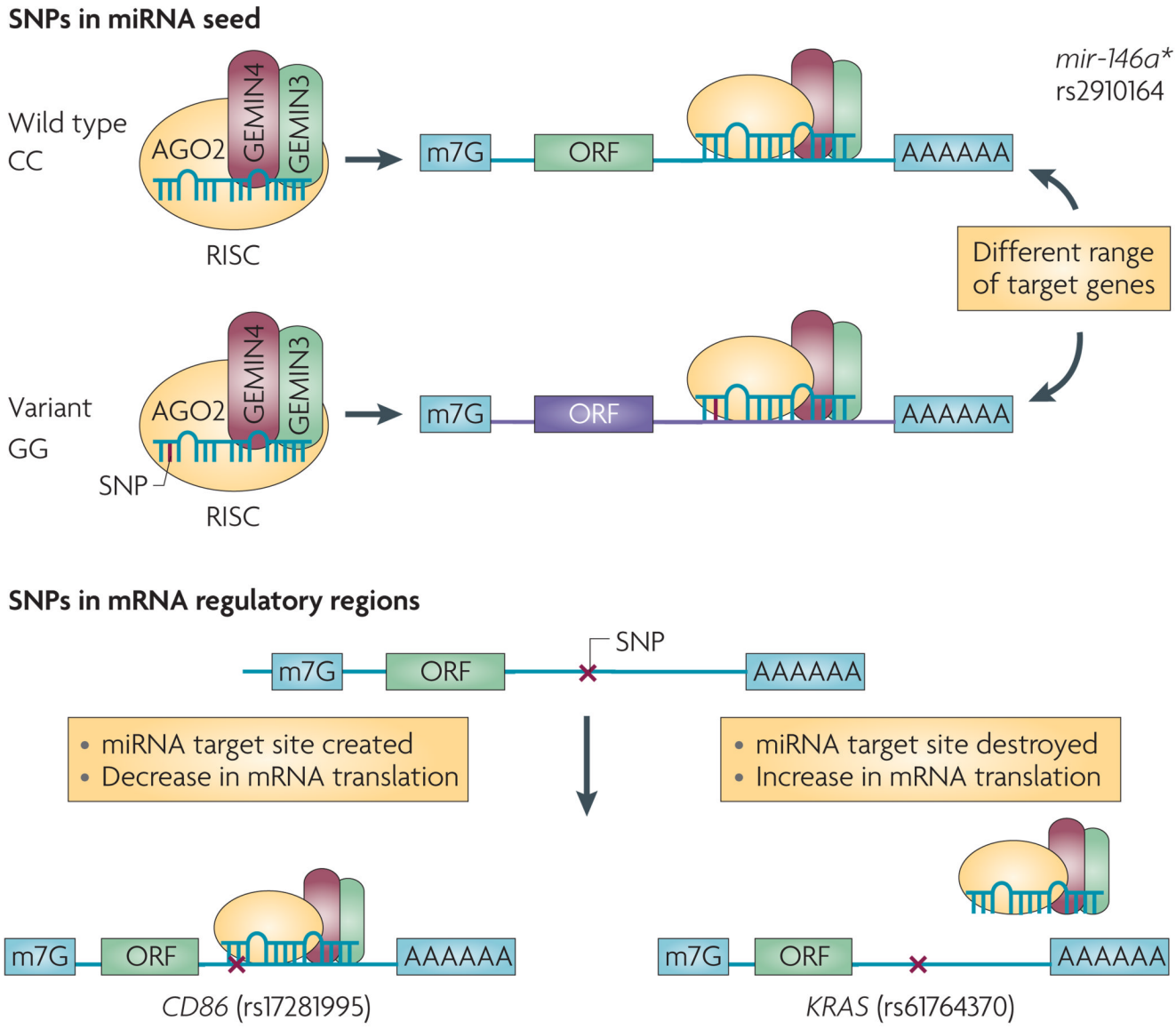

Figure 3. Diagrammatic representation of SNPs in miRNA seed and regulatory regions

Single nucleotide polymorphisms (SNPs) in mature microRNAs (miRNAs) within the seed sequence can strengthen or reduce binding between the miRNA and its mRNA target. Moreover, such SNPs can create or destroy target binding sites, as is the case for *mir-146a**. SNPs located within the 3′ untranslated region miRNA binding sites function analogously to seed region SNPs and modulate the miRNA–mRNA interaction. They can create or destroy miRNA binding sites and affect subsequent mRNA translation. AA, poly A tail; m7G, 7-methylguanosine cap; ORF, open reading frame.

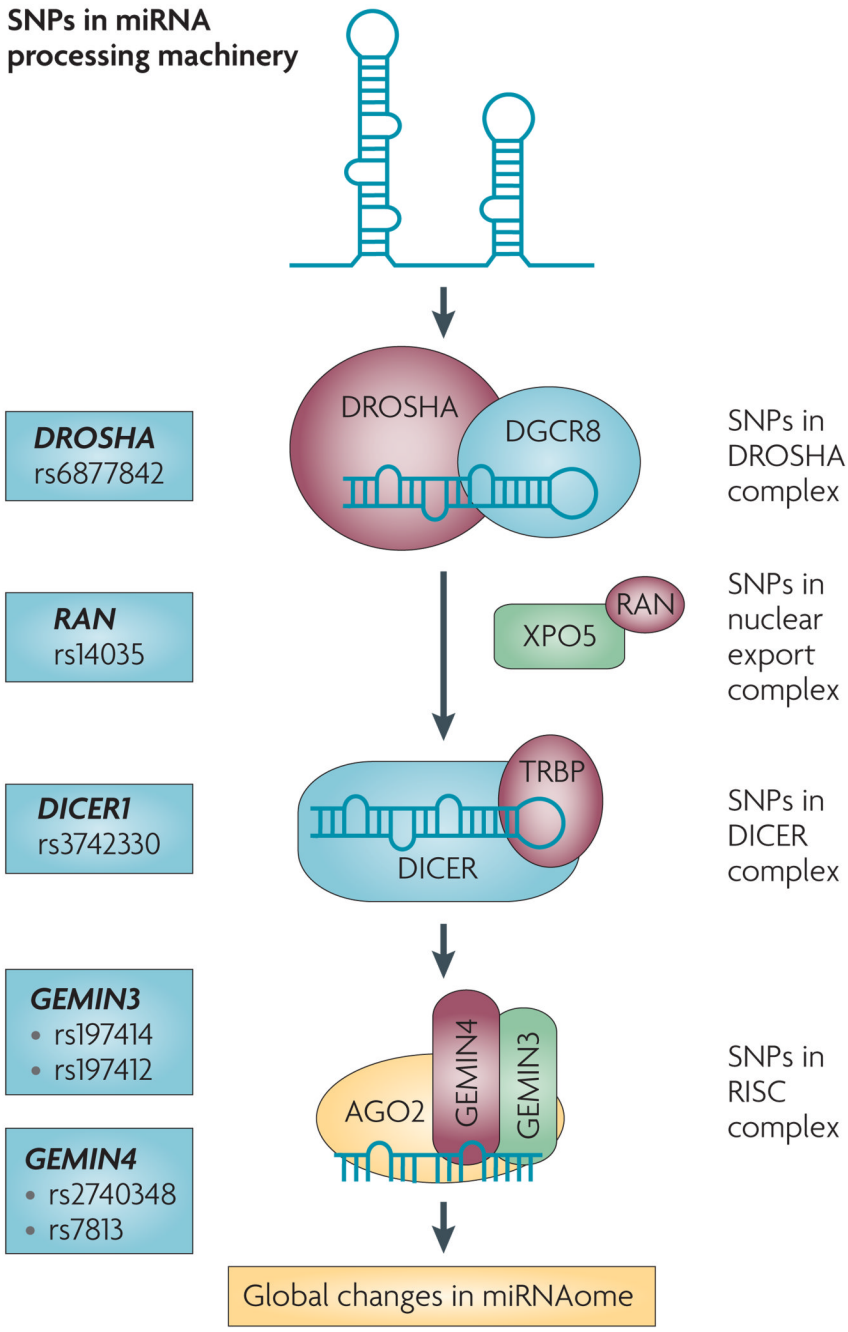

Figure 4. Diagrammatic representation of SNP sin miRNA processing machinery

Single nucleotide polymorphisms (SNPs) can also occur within the processing machinery. Although still lacking biological validation, these SNPs are likely to affect the microRNAome (miRNAome) as a whole, possibly leading to the overall suppression of miRNA output. In addition, SNPs in cofactors of miRNA processing, such as p53, may indirectly affect miRNA maturation.

Table 1

SNPs in miRNAs and cancer risk

SNP ID	miRNA (allele)	Cancer site	Genotype	OR (95% CI)	Refs
Pre-miR					
rs2910164	*mir-146a* (G/C)	Papillary thyroid carcinoma	GC GG/CC	Ref 1.62 (1.30; 2.00)	50
		Hepatocellular carcinoma	CC GC GG	Ref 1.19 (0.77; 1.85) 2.02 (1.06; 3.85)*	53
rs11614913	*mir-196a-2* (T/C)	Lung cancer	TT/TC CC	Ref 1.25 (1.01; 1.54)	39
		Oesophageal cancer	CC/CT TT	Ref 1.73 (1.16; 2.56)	37
		Breast cancer	CC CT TT	Ref 0.84 (0.63; 1.12) 0.44 (0.28; 0.70)	40
		Gastric cancer	TT/TC CC	Ref 1.57 (1.03; 2.39)	41
		Hepatocellular carcinoma	TT CT CC	Ref 1.01 (0.68; 0.76) 2.00 (1.07; 3.76)*	161
		Glioma	TT CT CC CC vs.CT/TT	Ref 1.24 (0.96; 1.59) 0.84 (0.61; 1.16) 0.74 (0.56; 0.98)	162
rs3746444	*mir-499* (A/G)	Breast cancer	AA AG GG AG/GG	Ref 1.19 (0.97; 1.46) 1.75 (1.07; 2.85) 1.25 (1.02; 1.51)	38
rs6505162	*mir-423* (C/A)	Bladder cancer	AA/ACM CC	Ref 1.34 (1.00; 1.79)*	35
		Oesophageal cancer	CC CA/AA	Ref 0.58 (0.41; 0.82)	37
		Breast cancer	AA AC CC	Ref 2.84 (1.17; 6.85) 2.77 (1.11; 6.90)	45
rs2289030	*mir-492* (C/G)	Bladder cancer	GG GC/CC	Ref 2.67 (1.26; 5.62)‡	35
rs895819	*mir-27a*	Breast cancer	AA AG GG	Ref 0.77 (0.66; 0.91) 0.88 (0.68; 1.15)	44

SNP ID	miRNA (allele)	Cancer site	Genotype	OR (95% CI)	Refs
		Breast cancer	CT TT	Ref 1.96 (1.16; 3.33)	45
rs5745925	*mir-631* (CT/-)	Oesophageal cancer	CT CT−/—	Ref 1.59 (1.02; 2.49)	37
Pri-miR					
rs7372209	*mir-26a-1* (C/T)	Bladder cancer	CC/CT TT	Ref 0.36 (0.13; 0.94)‡	35
		Premalignant oral lesions	CC CT/TT	Ref 2.09 (1.23; 3.56)	36
				Oesophageal cancer	CC CT/TT	Ref 1.33 (1.01; 1.74)	37
rs531564	*mir-124-1* (C/G)	Bladder cancer	CC/CG GG	Ref 4.85 (1.02; 23.01)§	35
rs213210	*mir-219-1* (T/C)	Oesophageal cancer	TT/CT CC	Ref 1.75 (1.10; 2.90)	37

CI, confidence interval; miRNA, microRNA; OR, odds ratio; Ref, reference allele; SNP, single nucleotide polymorphism.

*Males only.

‡Females only.

§Old cases only.

Table 2

Summary of SNPs in miRNA binding sites and their association with cancer risk

Gene	SNP ID	miRNA	Allele	Allele that miRNA binds	Genotype	OR (95% CI)	Cancer site	Refs
Candidate gene approach								
KRAS	rs61764370	Let-7	T-G	Ancestral allele	TT TG/GG	Ref 2.30 (1.1; 4.6)*	Lung cancer	74
ESR1	rs2747648	miR-453, miR-181 and miR-219	T-C	Derived allele	TT TC CC	Ref 0.74 (0.54; 0.97) 0.48 (0.09; 2.47)	Breast cancer	88
SETD8	rs16917496	miR-502	T-C	Derived allele	TT TC CC	Ref 0.94 (0.71; 1.24) 1.66 (1.06; 2.61)‡	Breast cancer	83
BTRCP	rs16405	miR-920	9 nt Ins/Del	Deletion alleles (derived)	Ins/Ins Ins/Del Del/Del	Ref 0.56 (0.31; 1.00) 0.44 (0.24; 0.83)	Hepatocellular carcinoma	163
IL1A	rs3783553	miR-122 and miR-378	TTCA Del/Ins	Deletion alleles (ancestral)	Del/Del Ins/Del Ins/Ins	Ref 0.79 (0.68; 0.92) 0.62 (0.49; 0.78)	Hepatocellular carcinoma	85
Systematic approach								
RYR3	rs1044129	Not specified	A-G	Not specified		1.75 (1.30; 2.36)	Not specified	82
C14orf101	rs4901706	Not specified	A-G	Not specified		8.37 (3.34; 21.00)	Not specified	82
DAG1	rs12583	Not specified	G-T	Not specified		0.47 (0.33; 0.68)	Not specified	82
SETD8	rs16917496	miR-502	C-T	Derived allele		1.94 (1.41; 2.68)	Not specified	82
AFF1	rs17703261	miR-19a, miR-19b, miR-585 and miR-648	A-T	Ancestral allele		0.34 (0.20; 0.58)	Not specified	82
KIAA0423	rs1053667	miR-19a and miR-19b	C-T	Ancestral allele		3.29 (1.72; 6.32)	Not specified	82
GOLGA7	rs11337	Not specified	G-T	Not specified		2.86 (1.45; 5.66)	Not specified	82
MATR3	rs14109	miR-420	C-T	Ancestral allele		2.74 (1.36; 5.53)	Not specified	82
KRT81	rs3660	miR-17, miR-93 miR-20b, miR-519d, miR-520g, miR-520h, miR-519c-3p, miR-519b-3p, miR-519a and miR-765	C-G	Ancestral allele and derived allele		1.39 (1.07; 1.80)	Not specified	82
USP9X	rs10463	Not specified	A-G	Not specified		2.04 (1.11; 3.74)	Not specified	82

Gene	SNP ID	miRNA	Allele	Allele that miRNA binds	Genotype	OR (95% CI)	Cancer site	Refs
BMPR1B	rs1434536	miR-125b	C-T	C allele	CC CT TT	Ref 1.29 (0.95; 1.74) 1.94 (1.40; 2.71)	Breast cancer	86
CD86	rs17281995	miR-337, miR-582, miR-200a, miR-184and miR-212	G-C	Ancestral allele	GG GC CC	Ref 1.33 (1.00; 1.76) 2.93 (1.29; 6.67)	Colorectal cancer	79

CI, confidence interval; Del, deletion; Ins, insertion; miRNA, microRNA; nt, nucleotide; OR, odds ratio; Ref, reference allele; SNP, single nucleotide polymorphism.

*Moderate smoker.

‡Pre-menopausal women only.

Table 3

Relationship between SNPs in miRNA-related machinery and cancer risk

Gene	SnP ID	Allele	Cancer type	OR (95% CI)	Refs
XPO5	rs11077	A-C	Oesophageal	1.58 (1.03; 2.45)	37
RAN	rs14035	C-T	Oesophageal	1.99 (1.17; 3.38)	37
TRBP	rs784567	C-T	Bladder	0.69 (0.48; 0.98)*	35
DICER1	rs3742330	A-G	Premalignant oral lesions	2.09 (1.03; 4.24)	36
GEMIN4	rs2740348	G-C	Renal cell carcinoma	0.67 (0.47; 0.96)	111
	rs7813	T- C	Renal cell carcinoma	0.68; (0.47; 0.98)	111
GEMIN3	rs197414	C-A	Oesophageal Bladder	1.45 (1.02; 2.06) 2.50 (1.08; 5.78)	35,37
	rs197412	T- C	Premalignant oral lesions	0.58 (0.33; 0.99)	36
TNRC6B	rs9623117	C-T	Prostate	1.18 (1.11; 1.26)	164

CI, confidence interval; miRNA, microRNA; OR, odds ratio; SNP, single nucleotide polymorphism.

*Significant in young compared with old.