DATABASE IN BRIEF

HUMAN MUTATION

OFFICIAL JOURNAL
HGVS
HUMAN GENOME VARIATION SOCIETY
www.hgvs.org

miRvar: A Comprehensive Database for Genomic Variations in microRNAs

Deeksha Bhartiya, Saurabh V. Laddha, Arijit Mukhopadhyay, and Vinod Scaria*

GN Ramachandran Knowledge Center for Genome Informatics, Institute of Genomics and Integrative Biology (CSIR), Mall Road, Delhi 110007

*Correspondence to Vinod Scaria, GN Ramachandran Knowledge Center for Genome Informatics, Institute of Genomics and Integrative Biology (CSIR), Mall Road, Delhi 110007. Tel: +91-11-25895615. Fax: +91-11-2766 7471. E-mail: vinods@igib.res.in

Contract grant sponsor: Council of Scientific and Industrial Research; Contract grant number: NWP0036.

Communicated by Peter K. Rogan

ABSTRACT: microRNAs are a recently discovered and well studied class of small noncoding functional RNAs. The regulatory role of microRNAs (miRNAs) has been well studied in a wide variety of biological processes but there have been no systematic effort to understand and analyze the genetic variations in miRNA loci and study its functional consequences. We have comprehensively curated genetic variations in miRNA loci in the human genome and established a computational pipeline to assess potential functional consequences of these variants along with methods for systematic curation and reporting of variations in these loci. The data is made available on the Leiden Open (source) Variation Database (LOVD) platform at http://genome.igib.res.in/mirlovd to provide ease of aggregation and analysis and is open for community curation efforts. ©2011 Wiley-Liss, Inc.

KEY WORDS: microRNA, variation, curation, LOVD

INTRODUCTION

The role of miRNAs ranges from host-pathogen interactions (Scaria et al., 2007), viral oncogenesis (Scaria and Jadhav, 2007) to regulation of development (Navarro and Lieberman, 2010). The mechanisms of many miRNAs and their association in diseases have now been studied in detail and is extensively reviewed (O'Connell et al., 2010; Casalini and Iorio, 2009; Singh et al., 2008; Lu et al., 2008). MicroRNAs (miRNAs) are well studied class of small endogenous non protein coding RNAs. miRNAs are transcribed as longer primary miRNAs (pri-miRNAs) and are generally polyadenylated and capped and processed into mature miRNA through an elaborate pathway. The biogenesis involves cropping of the hairpins, export into the cytoplasm, dicing and incorporation of the mature

Received 11 June 2010; accepted revised manuscript 21 January 2011.

DOI: 10.1002/humu.21482

miRNA into a ribonucleoprotein complex which regulates gene expression by binding to 3' UTR or 5' UTR of coding transcripts (Du and Zamore, 2005; Bartel, 2004; Meyers et al., 2010; Cai et al., 2009; Kim et al., 2005).

The availability of high-throughput DNA sequencing technology has enabled large-scale sequencing of humans and other model organisms revealing a wealth of information on human variability. The functional role of these variations has been studied mostly in the context of protein-coding genes (Frazer et al., 2009) whereas the functional consequences of variations in noncoding RNAs are just beginning to be reported and understood at the molecular level. Recently, it was shown that genetic variations in the loci encoding for precursor of hsa-mir-146a reduces the cellular level of mature miRNA and thereby confers genetic predisposition to papillary thyroid carcinoma (Jazdzewski et al., 2008). Later, the group also found an association between variations in the passenger strand of hsa-mir-146a and thyroid cancer (Jazdzewski et al., 2009). Similarly other studies have suggested that a single nucleotide variant rs11614913 in hsa-mir-196a-2 might be associated with survival of patients with non-small cell lung cancer and affect mature miRNA expression (Tian et al., 2009; Hu et al., 2009; Hu et al., 2008). Likewise rs895819 in hsa-mir-27a, an oncogenic miRNA, was showed to have association with a reduced risk of familial breast cancer (Yang et al., 2009). Mencia et al., 2009 showed that a point mutation in seed region of hsa-miR-96 result in significant reduction of mRNA targeting causing autosomal dominant, progressive hearing loss. These evidences suggest a greater role of genomic variations in modulating miRNA function and regulation manifesting as phenotypic correlates (Sun et al., 2009).

It is thus imperative to catalog the genomic variations in an interactive, web-based database along with necessary tools to mine the information to identify and study their potential pathogenic effects. This would necessitate not only a well-annotated and continuously updated set of variant information at these loci in terms of standard notations, phenotypes or disease associations along with features for the community to actively contribute variations that may be tagged with the clinical phenotypes. In addition, the resource should support interoperability with other tools and resources. To this end, we have used the Leiden Open (source) Variation Database (LOVD) Software (Fokkema et al., 2005) for creating a catalog of variations in miRNA loci, through manual curation of data from literature and from public databases available at http://genome.igib.res.in/mirlovd. With just a few studies on variations in miRNA, there is no standard ontology and format for them which make integration of data on a single platform strenuous. Thus we also propose a nomenclature standard for miRNA variations which fall in line with the already existing standards for variations in protein coding genes. To the best of our knowledge this is the first comprehensive collection and systematic genome-wide analysis of genomic variations in miRNA loci.

DATA COLLECTION AND ANALYSES

Data Collection

We collected and manually curated variations from public resources and published literature. The database of Single Nucleotide Polymorphisms (dbSNP build 130 on March 2006 assembly of the human genome) (Smigielski et al., 2000) was used to collect data for genetic variations in human genome which have been reported as validated by one or more methods. This data was mapped to human precursor miRNA (pre-miRNA) loci and sequence information was downloaded from hg18 version of human genome from UCSC genome browser (Kent et al., 2002). The data comparison including standardization of positions and analysis was performed using in-house Perl scripts available freely at http://genome.igib.res.in/mirvar/pipeline.html and the dataset was further manually curated to check for any possible error.

We also manually curated variation information from published literature which reported variations in miRNA loci. These include variations in 17 X-linked miRNAs from male patients diagnosed with autism or schizophrenia (Sun et al., 2009). In addition six post-transcriptionally modified miRNAs with adenosine to inosine RNA editing were also included (Blow et al., 2006). The information about the hairpin structure of pre-miRNAs was collected from miRNA registry (miRBase, release 13) (Griffiths-Jones et al., 2008; Griffiths-Jones et al., 2006) and was used for the analysis of structure and effects of variations on mature miRNAs.

Nomenclature

We standardized the nomenclature for reporting miRNAs, keeping in mind the standards laid down by the Human Genome Variation Society (HGVS) (http://www.hgvs.org/mutnomen/). We followed the HUGO Gene Nomenclature Committee (HGNC) (Bruford et al., 2008) nomenclature for miRNAs along with the miRBase nomenclature.

The Database

The database, miRvar, is based on Leiden Open (source) Variation Database (LOVD) system which is web-based and user- interactive. LOVD provides individual home pages for all the genes and contains a list of variations for each gene.

miRvar contains gene pages for each of the miRNAs found to harbor variations and the variant information was converted following the guidelines described in the previous section. We have considered the pre-miRNA sequences as described in miRNA registry as the reference sequence because the transcriptional start sites of miRNAs are poorly understood. All variations were mapped to their respective reference sequences. The HGNC nomenclature of miRNAs was used and miRNA names were converted accordingly. The database is based on MySQL and additional links have been provided to enable easy access of related information in other databases like UCSC genome browser, miRBase (Griffiths-Jones et al., 2008; Griffiths-Jones et al., 2006), miR2Diseases (Jiang et al., 2009), miRecords (Xiao et al., 2009), and 1000 Genome Project (http://www.1000genomes.org). In addition, export of data in tab delimited formats and feeds for aggregation have been enabled for interoperability with other resources.

Data Analysis and Pipeline

A computational pipeline was established to predict potential effects of variations on miRNA biogenesis and its functions (Figure 1). This pipeline includes PHDcleav (http://imtech.res.in/raghava/phdcleav/submission.html) which predicts dicer processing sites using Support Vector Machine (SVM) models. We have used default parameters which is zero for positive dicer processing site predictions. The dicer cleavage site changes were predicted both for wild type pre-miRNAs and pre-miRNAs with variations.

In addition, RISCbinder (Ahmed et al., 2009) was used to predict the guide strand of miRNAs. We have used sequence based approach in RISCbinder with default cutoff. The strand with score more than the cutoff score (which was set zero) was predicted as positive (miRNA) else as negative (miRNA*). Guide strand of miRNAs were predicted both for wild type mature miRNAs and mature miRNAs with variations. Input files used for both the methods are in fasta format.

Further, we have analyzed the allele frequencies data of SNPs from the HapMap populations (http://hapmap.ncbi.nlm.nih.gov/biomart/martview) in order to evaluate the possible effect of the variations in the miRNA and its penetrance in the population.

RESULTS AND DISCUSSION

Mapping

A systematic mapping of validated SNP from dbSNP (build 130) from UCSC genome browser on genomic locations of pre-miRNA revealed 106 SNPs positions mapping to 85 miRNAs (Supp. Table S1). We mapped the position of genetic variants on pre-miRNA structure (Table 1) and found that 67 map to stem region of pre-miRNAs, six mapped to the terminal loop and 10 to the overhang (5’ or 3’), while 23 mapped to the mature miRNA loci of which five mapped to the 5' eight nucleotides or the seed region of the miRNAs (Table 2).

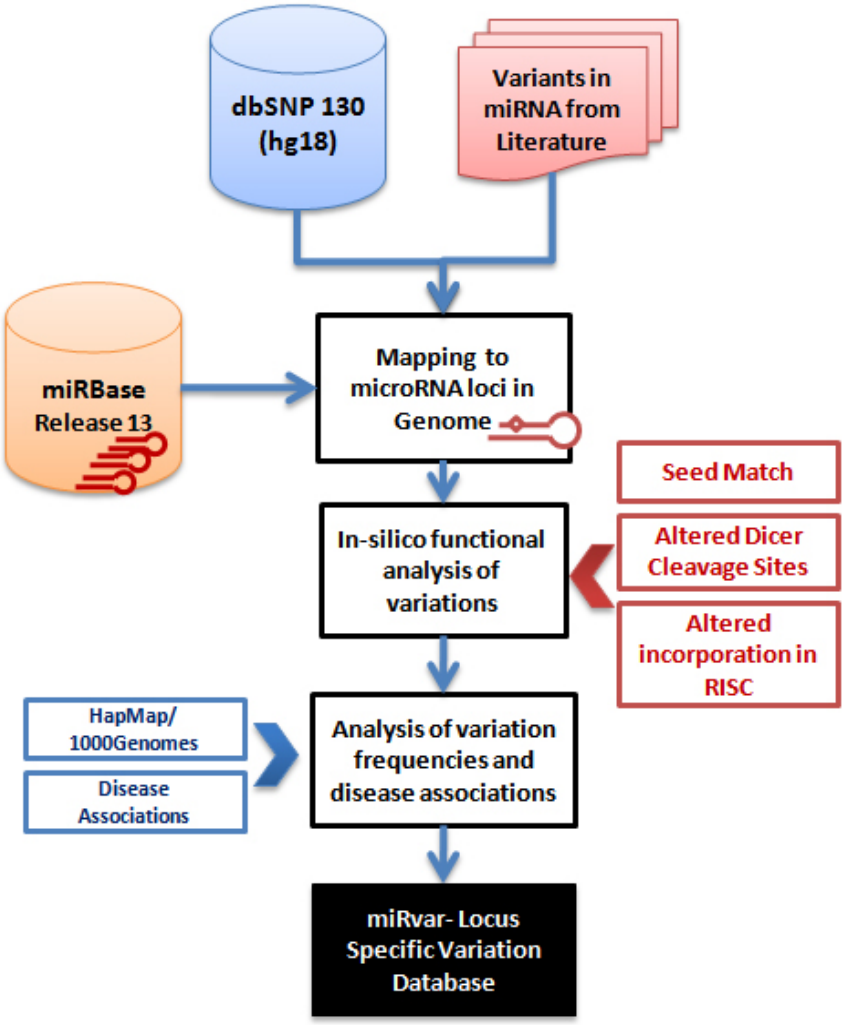


Figure 1. Workflow depicting analysis of variations in miRNA loci.

Table 1. Location of genetic variation on hairpin structure of precursor miRNAs from the mapped data from dbSNP

Location in precursor miRNA	Genetic variations (SNP) mapped from dbSNP
Terminal loop	6
Stem and Secondary Loop (excluding mature region)	67
Mature	23 (5 in seed region)
Overhang (5' or 3')	10

Table 2. miRNAs with variations in seed region

miRNAs	HGNC ID	Variation	dbSNP ID
hsa-mir-1977	37065	g.5A>G	rs2854138
hsa-mir-146a	31533	g.60C>G	rs2910164
hsa-mir-938	33681	g.16C>T	rs12416605
hsa-mir-518d	32121	g.59A>G	rs73602910
hsa-mir-499	32133	g.73C>T	rs7267163

According to miRBase, release 13, hsa-mir-1977 have been found to be present in chromosome 1 and Mitochondria (MT) which overlaps with Mt tRNA sequence so it has been removed from the current release of miRBase. The SNP was mapped on the same position on hsa-mir-1977 in both the chromosomes.

Functional consequences of variations

To investigate the potential functional effects of variants in the stem region in the biogenesis of miRNA, we established a protocol by using algorithms that predict functional features in a set of input miRNA sequences. PHDcleav predicted two miRNAs (Table 3) from the entire dataset having altered dicer processing sites as a result of variation suggesting the potential effect of variations on the processing of miRNAs.

RISCbinder predicted hsa-mir-608 (Table 3) from the 23 miRNAs having SNPs in mature sequence to have a different guide strand with the incorporation of variation in its sequence. The alteration in the selection of guide strand for RISC might hamper the target mRNA translation inhibition. All the variants, methods and potential functional effects are summarized in Table 3. Of all the miRNA variations we tested, these three variations in miRNAs could have potential impact in the biogenesis process and function.

In order to evaluate the possible effects of the variations in the miRNA sequences in the mapped data from UCSC, we have analyzed the allele frequencies of the 41 SNPs from the HapMap database, and found the median allele frequency to be less than 0.1. In addition, for 23 SNPs the minor allele was absent in at least one of the populations. Both these facts point towards a possible negative selection for the minor allele implicating that there might be functional consequences for the variant allele (Supp. Fig. S1).

Table 3. Functional analysis using PHDcleav and RISCbinder

miRBase ID	HGNC ID	Variation	Effect of variation
hsa-mir-608	32864	g.37C>G	Might alter incorporation in RISC$^{(RI)}$
hsa-mir-449b	32794	g.27A>G	Might alter Dicer Cleavage Site$^{(D)}$
hsa-mir-1324	35377	g.31C>G	Might alter Dicer Cleavage Site$^{(D)}$

(RI): RISCbinder used to predict guide strand of miRNA. (D): PHDcleav used to predict Dicer cleavage sites in pre-miRNA
For the details about the scores of both the tools used, see Supp. Table S2.

miRvar

The data was then uploaded on miRvar, a community curation resource for genomic variation in miRNA loci. miRvar provides a user friendly interface to search variants in different miRNAs and has information about the variant position and the allele change in the respective miRNA, in addition to making it easy to browse for information through linkouts like miRBase, UCSC Genome Browser, miR2Disease, 1000 Genomes Project and miRecords. The database comprises of a total 108 unique entries including 85 miRNAs from UCSC-dbSNP mapping and 25 from published literature (Supp. Table S1). The 25 miRNAs from literature include 19 X-linked miRNAs having variations (Sun et al., 2009) and 6 post-transcriptionally modified miRNAs with adenosine to inosine RNA editing (Blow et al., 2006). Eight of the systematic mapped genetic variations have been previously reported to be associated with diseases (Supp. Table S3). The variant files can be downloaded from the database in standard meta-tagged format to enable interoperability with other resources. We have also enabled the data for visualization on genome browser including the UCSC genome browser. The database also provides a search page to search for miRNAs using their names, disease associations, location of SNP on the hairpin structure and chromosome. The database is open for curation and data suggestions from the community.

FUTURE PERSPECTIVES

miRvar is a starting point for systematic curation and analysis of genetic variants in miRNA loci. Since the understanding of miRNAs and their effects in biological processes, including diseases are just emerging and so is the understanding on the potential functional role of the genomic variations in these loci, we foresee this database to be enriched with new variations and information on disease associations. Though there have been a paucity of tools for predicting functional impact of variations in miRNAs, a curated and up-to-date dataset, as we foresee would also enable researchers to create and test newer methodologies to predict pathogenic effects of genetic variation in miRNAs. In support to the community effort towards creating high-quality annotations and standards for data exchange, we have enabled a standardized format for collection and reporting of variations in these loci. We feel the database would be immensely enriched by data generated through the ongoing efforts internationally to understand variations in human populations, including disease associations. In addition to the existing data, we have already started porting variation data from high-throughput small-RNA sequencing experiments using next-generation sequencing technology. We would work towards making the resource compatible and linked with other major resources and tools in the field. We also feel the standards and methods laid forward in developing the resource would be widely used by the community in the future and would be a guide to curation of variations in other noncoding RNA loci.

ACKNOWLEDGMENTS

The authors acknowledge group members of RNA@IGIB for stimulating discussions and Dr. Anshu Bhardwaj for reviewing the manuscript. We thank Mr. Anupam Kumar Mondal and Ms. Saakshi Jalali for evaluating the database and Mr. Hemant Suryawanshi for his valuable inputs. The work was supported by funding from the Council of Scientific and Industrial Research (CSIR), India through project "NWP0036" (Comparative genomics of non-coding RNAs). The authors also acknowledge Dr. Debasis Dash and the Gen2Phen project (FP7/2007-2013) for considerable support and inputs in the terms of designing and standardizing the data and annotations.

REFERENCES

Ahmed F, Ansari HR, Raghava GP. 2009. Prediction of guide strand of microRNAs from its sequence and secondary structure. BMC Bioinformatics 10:105.

Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281-297.

Blow MJ, Grocock RJ, van DS, Enright AJ, Dicks E, Futreal PA, Wooster R, Stratton MR. 2006. RNA editing of human microRNAs. Genome Biol 7:R27.

Bruford EA, Lush MJ, Wright MW, Sneddon TP, Povey S, Birney E. 2008. The HGNC Database in 2008: a resource for the human genome. Nucleic Acids Res 36:D445-D448.

Cai Y, Yu X, Hu S, Yu J. 2009. A brief review on the mechanisms of miRNA regulation. Genomics Proteomics Bioinformatics 7:147-154.

Casalini P, Iorio MV. 2009. MicroRNAs and future therapeutic applications in cancer. J BUON 14 Suppl 1:S17-S22.

Chatzikyriakidou A, Voulgari PV, Georgiou I, Drosos AA. 2010. The role of microRNA-146a (miR-146a) and its target IL-1R-associated kinase (IRAK1) in psoriatic arthritis susceptibility. Scand J Immunol 71:382-385.

Christensen BC, Avissar-Whiting M, Ouellet LG, Butler R, Nelson HH, McClean MD, Marsit CJ, Kelsey KT. 2010. Mature miRNA sequence polymorphism in MIR196A2 is associated with risk and prognosis of head and neck cancer. Clin Cancer Res 16:3713-3720.

Dou T, Wu Q, Chen X, Ribas J, Ni X, Tang C, Huang F, Zhou L, Lu D. 2010. A polymorphism of microRNA196a genome region was associated with decreased risk of glioma in Chinese population. J Cancer Res Clin Oncol 136:1853-1859..

Du T, Zamore PD. 2005. microPrimer: the biogenesis and function of microRNA. Development 132:4645-4652.

Fokkema IF, den Dunnen JT, Taschner PE. 2005. LOVD: easy creation of a locus-specific sequence variation database using an "LSDB-in-a-box" approach. Hum Mutat 26:63-68.

Frazer KA, Murray SS, Schork NJ, Topol EJ. 2009. Human genetic variation and its contribution to complex traits. Nat Rev Genet 10:241-251.

Griffiths-Jones S, Grocock RJ, van DS, Bateman A, Enright AJ. 2006. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 34:D140-D144.

Griffiths-Jones S, Saini HK, van DS, Enright AJ. 2008. miRBase: tools for microRNA genomics. Nucleic Acids Res 36:D154-D158.

Hoffman AE, Zheng T, Yi C, Leaderer D, Weidhaas J, Slack F, Zhang Y, Paranjape T, Zhu Y. 2009. microRNA miR-196a-2 and breast cancer: a genetic and epigenetic association study and functional analysis. Cancer Res 69:5970-5977.

Hu Z, Chen J, Tian T, Zhou X, Gu H, Xu L, Zeng Y, Miao R, Jin G, Ma H, Chen Y, Shen H. 2008. Genetic variants of miRNA sequences and non-small cell lung cancer survival. J Clin Invest 118:2600-2608.

Hu Z, Liang J, Wang Z, Tian T, Zhou X, Chen J, Miao R, Wang Y, Wang X, Shen H. 2009. Common genetic variants in pre-microRNAs were associated with increased risk of breast cancer in Chinese women. Hum Mutat 30:79-84.

Jazdzewski K, Liyanarachchi S, Swierniak M, Pachucki J, Ringel MD, Jarzab B, de la Chapelle A. 2009. Polymorphic mature microRNAs from passenger strand of pre-miR-146a contribute to thyroid cancer. Proc Natl Acad Sci U S A 106:1502-1505.

Jazdzewski K, Murray EL, Franssila K, Jarzab B, Schoenberg DR, de la Chapelle A. 2008. Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. Proc Natl Acad Sci U S A 105:7269-7274.

Jiang Q, Wang Y, Hao Y, Juan L, Teng M, Zhang X, Li M, Wang G, Liu Y. 2009. miR2Disease: a manually curated database for microRNA deregulation in human disease. Nucleic Acids Res 37:D98-104.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. Genome Res 12:996-1006.

Kim VN. 2005. MicroRNA biogenesis: coordinated cropping and dicing. Nat Rev Mol Cell Biol 6:376-385.

Kontorovich T, Levy A, Korostishevsky M, Nir U, Friedman E. 2010. Single nucleotide polymorphisms in miRNA binding sites and miRNA genes as breast/ovarian cancer risk modifiers in Jewish high-risk women. Int J Cancer 127:589-597.

Li H, Ruan J, Durbin R. 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 18:1851-1858.

Lu M, Zhang Q, Deng M, Miao J, Guo Y, Gao W, Cui Q. 2008. An analysis of human microRNA and disease associations. PLoS One 3:e3420.

Mencia A, Modamio-Hoybjor S, Redshaw N, Morin M, Mayo-Merino F, Olavarrieta L, Aguirre LA, del C, I, Steel KP, Dalmay T, Moreno F, Moreno-Pelayo MA. 2009. Mutations in the seed region of human miR-96 are responsible for nonsyndromic progressive hearing loss. Nat Genet 41:609-613.

Meyers BC, Simon SA, Zhai J. 2010. MicroRNA processing: battle of the bulge. Curr Biol 20:R68-R70.

Navarro F, Lieberman J. 2010. Small RNAs Guide Hematopoietic Cell Differentiation and Function. J Immunol 184:5939-5947.

O'Connell RM, Rao DS, Chaudhuri AA, Baltimore D. 2010. Physiological and pathological roles for microRNAs in the immune system. Nat Rev Immunol 10:111-122.

Peng S, Kuang Z, Sheng C, Zhang Y, Xu H, Cheng Q. 2009. Association of MicroRNA-196a-2 Gene Polymorphism with Gastric Cancer Risk in a Chinese Population. Dig Dis Sci 55:2288-2293.

Qi P, Dou TH, Geng L, Zhou FG, Gu X, Wang H, Gao CF. 2010. Association of a variant in MIR 196A2 with susceptibility to hepatocellular carcinoma in male Chinese patients with chronic hepatitis B virus infection. Hum Immunol 71:621-626.

Scaria V, Hariharan M, Brahmachari SK, Maiti S, Pillai B. 2007. microRNA: an emerging therapeutic. ChemMedChem 2:789-792.

Scaria V, Jadhav V. 2007. microRNAs in viral oncogenesis. Retrovirology 4:82.

Singh SK, Pal BM, Girschick HJ, Bhadra U. 2008. MicroRNAs--micro in size but macro in function. FEBS J 275:4929-4944.

Shen J, Ambrosone CB, DiCioccio RA, Odunsi K, Lele SB, Zhao H. 2008. A functional polymorphism in the miR-146a gene and age of familial breast/ovarian cancer diagnosis. Carcinogenesis 29:1963-1966.

Smigielski EM, Sirotkin K, Ward M, Sherry ST. 2000. dbSNP: a database of single nucleotide polymorphisms. Nucleic Acids Res 28:352-355.

Sun G, Yan J, Noltner K, Feng J, Li H, Sarkis DA, Sommer SS, Rossi JJ. 2009. SNPs in human miRNA genes affect biogenesis and function. RNA 15:1640-1651.

Tian T, Shu Y, Chen J, Hu Z, Xu L, Jin G, Liang J, Liu P, Zhou X, Miao R, Ma H, Chen Y, Shen H. 2009. A functional genetic variant in microRNA-196a2 is associated with increased susceptibility of lung cancer in Chinese. Cancer Epidemiol Biomarkers Prev 18:1183-1187.

Wang M, Ye Y, Qian H, Song Z, Jia X, Zhang Z, Zhou J, Ni C. 2010. Common genetic variants in pre-microRNAs are associated with risk of coal workers' pneumoconiosis. J Hum Genet 55:13-17.

Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T. 2009. miRecords: an integrated resource for microRNA-target interactions. Nucleic Acids Res 37:D105-D110.

Xu B, Feng NH, Li PC, Tao J, Wu D, Zhang ZD, Tong N, Wang JF, Song NH, Zhang W, Hua LX, Wu HF. 2010. A functional polymorphism in Pre-miR-146a gene is associated with prostate cancer risk and mature miR-146a expression in vivo. Prostate 70:467-472.

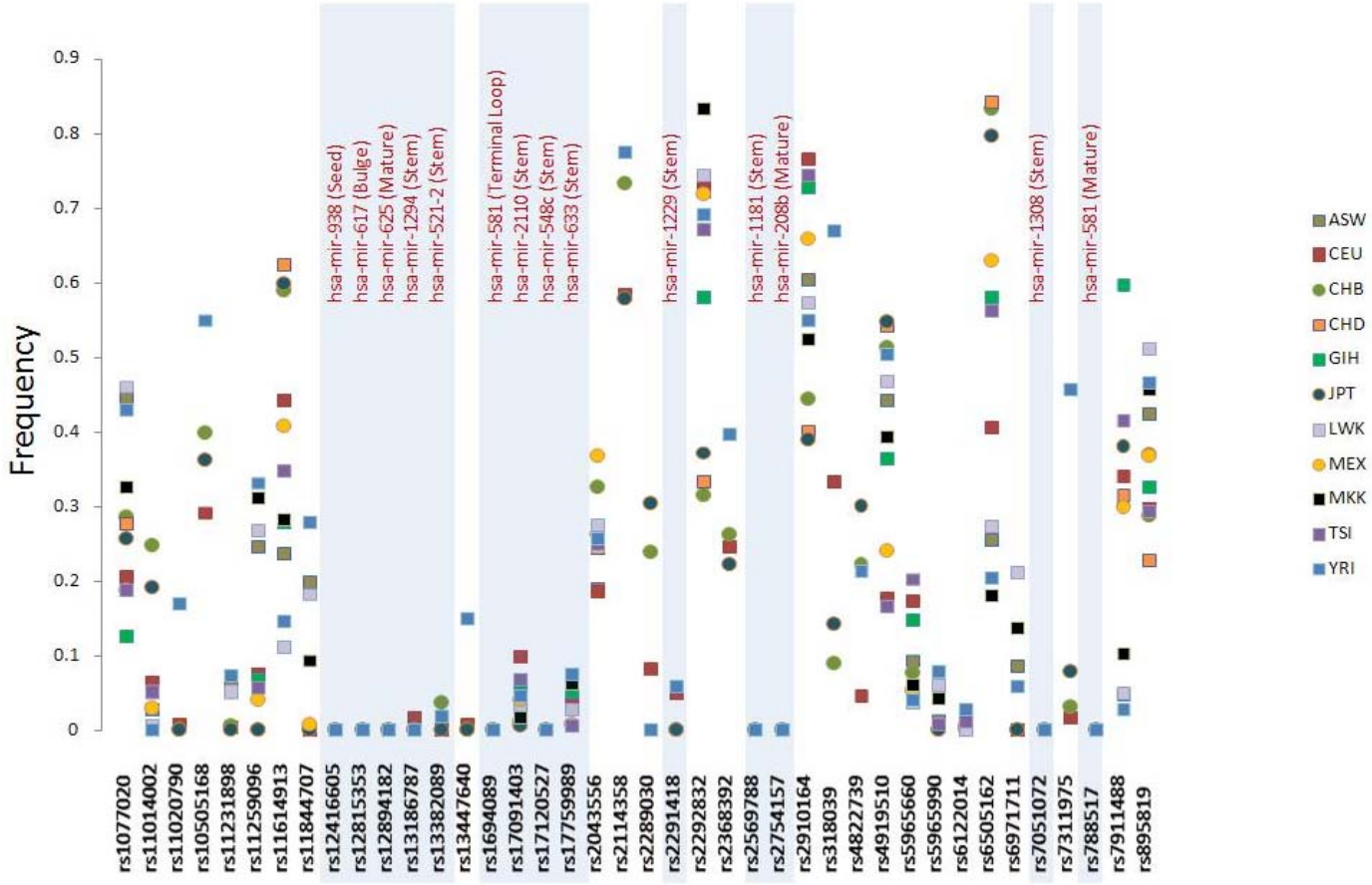
Yang H, Dinney CP, Ye Y, Zhu Y, Grossman HB, Wu X. 2008. Evaluation of genetic variants in microRNA-related genes and risk of bladder cancer. Cancer Res 68:2530-2537.

Yang R, Schlehe B, Hemminki K, Sutter C, Bugert P, Wappenschmidt B, Volkmann J, Varon R, Weber BH, Niederacher D, Arnold N, Meindl A, Bartram CR, Schmutzler RK, Burwinkel B. 2009. A genetic variant in the pre-miR-27a oncogene is associated with a reduced familial breast cancer risk. Breast Cancer Res Treat. 121:693-702.

Ye Y, Wang KK, Gu J, Yang H, Lin J, Ajani JA, Wu X. 2008. Genetic variations in microRNA-related genes are novel susceptibility loci for esophageal cancer risk. Cancer Prev Res (Phila Pa) 1:460-469.

Zhou B, Rao L, Peng Y, Wang Y, Chen Y, Song Y, Zhang L. 2010. Common genetic polymorphisms in pre-microRNAs were associated with increased risk of dilated cardiomyopathy. Clin Chim Acta 411:1287-1290.

SUPPORTING INFORMATION



Supp. Figure S1. Minor Allele Frequencies of the Variations in HapMap Populations.

Supp. Table S1. Genetic variations mapped on precursor miRNAs using two strategies: UCSC-dbSNP Mapping and Literature surveys

Chr	miRBase ID***	HGNC ID	Variation*	dbSNP ID**	Location on hairpin structure	SNP Validation status	Strategy Used
chr1	hsa-mir-1977	37065	g.5A>G	rs2854138	Seed	by-cluster	UCSC-dbSNP Mapping
chr1	hsa-mir-1977	37065	g.62C>T	rs9783068	Stem	by-cluster	UCSC-dbSNP Mapping
chr1	hsa-mir-1977	37065	g.61A>G	rs41453547	Stem	by-cluster	UCSC-dbSNP Mapping
chr1	hsa-mir-1977	37065	g.53C>T	rs9701099	Stem	by-cluster, by-2hit-2allele	UCSC-dbSNP Mapping
chr2	hsa-mir-1285-2	35278	g.87A>C	rs72904307	3' Overhang	by-1000genomes	UCSC-dbSNP Mapping
chr2	hsa-mir-1302-3	35295	g.59G>T	rs2441621	Stem	by-2hit-2allele	UCSC-dbSNP Mapping
chr2	hsa-mir-1302-3	35295	g.11A>G	rs6542147	Stem	by-cluster, by-2hit-2allele	UCSC-dbSNP Mapping
chr2	hsa-mir-149	31536	g.86C>T	rs2292832	Stem	by-cluster,by-frequency,by-2hit-2allele,by-hapmap,by-1000genomes	UCSC-dbSNP Mapping
chr2	hsa-mir-559	32815	g.53C>T	rs58450758	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr3	hsa-mir-1248	35314	g.72C>T	rs73063489	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr3	hsa-mir-1324	35377	g.93C>T	rs7614638	3' Overhang	by-2hit-2allele	UCSC-dbSNP Mapping
chr3	hsa-mir-1324	35377	g.31C>G	rs3008994	Stem	by-cluster, by-2hit-2allele	UCSC-dbSNP Mapping
chr3	hsa-mir-1324	35377	g.32A>G	rs3008993	Stem	by-cluster	UCSC-dbSNP Mapping
chr3	hsa-mir-569	32825	g.96A>C	rs73037390	3' Overhang	by-1000genomes	UCSC-dbSNP Mapping

Chr	miRBase ID***	HGNC ID	Variation*	dbSNP ID**	Location on hairpin structure	SNP Validation status	Strategy Used
chr3	hsa-mir-570	32826	g.34C>T	rs9860655	Stem	by-2hit-2allele	UCSC-dbSNP Mapping
chr4	hsa-mir-1255a	35320	g.71C>T	rs28664200	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr4	hsa-mir-1255b-1	35366	g.3A>G	rs6841938	Mature	by-2hit-2allele	UCSC-dbSNP Mapping
chr4	hsa-mir-1269	35337	g.79A>G	rs73239138	Mature	by-1000genomes	UCSC-dbSNP Mapping
chr4	hsa-mir-943	33689	g.12C>T	rs1077020	Stem	by-cluster, by-frequency, by-hapmap	UCSC-dbSNP Mapping
chr5	hsa-mir-1229	33924	g.23C>T	rs2291418	Stem	by-frequency, by-hapmap, by-1000genomes	UCSC-dbSNP Mapping
chr5	hsa-mir-1274a	35341	g.33C>T	rs318039	Terminal loop	by-frequency, by-2hit-2allele, by-hapmap,by-1000genomes	UCSC-dbSNP Mapping
chr5	hsa-mir-1294	35287	g.104A>G	rs13186787	Stem	by-frequency, by-2hit-2allele, by-hapmap	UCSC-dbSNP Mapping
chr5	hsa-mir-146a	31533	g.60C>G	rs2910164	Seed	by-cluster, by-frequency, by-submitter, ,by-hapmap,by-1000genomes	UCSC-dbSNP Mapping
chr5	hsa-mir-449b	32794	g.27A>G	rs10061133	Mature	by-cluster, by-2hit-2allele	UCSC-dbSNP Mapping
chr5	hsa-mir-580	32836	g.37G>T	rs73080005	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr5	hsa-mir-581	32837	g.36A>G	rs788517	Mature	by-hapmap	UCSC-dbSNP Mapping
chr5	hsa-mir-581	32837	g.44G>T	rs1694089	Terminal loop	by-hapmap	UCSC-dbSNP Mapping
chr5	hsa-mir-585	32841	g.87A>G	rs62376934	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr7	hsa-mir-590	32846	g.72C>T	rs6971711	Mature	by-frequency, by-hapmap	UCSC-dbSNP Mapping
chr7	hsa-mir-595	32851	g.3C>T	rs4909237	Stem	by-2hit-	UCSC-

Chr	miRBase ID***	HGNC ID	Variation*	dbSNP ID**	Location on hairpin structure	SNP Validation status	Strategy Used
						2allele, by-1000genomes	dbSNP Mapping
chr7	hsa-mir-96	31648	g.42A>G	rs73159662	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr8	hsa-mir-1206	35272	g.36C>T	rs2114358	Stem	by-cluster, by-frequency, by-2hit-2allele, by-hapmap,by-1000genomes	UCSC-dbSNP Mapping
chr8	hsa-mir-1322	35374	g.60C>T	rs59878596	Mature	by-1000genomes	UCSC-dbSNP Mapping
chr8	hsa-mir-2053	37069	g.31A>G	rs10505168	Stem	by-cluster, by-frequency, by-hapmap, by-1000genomes	UCSC-dbSNP Mapping
chr8	hsa-mir-548h-4	35345	g.79C>T	rs73235381	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr8	hsa-mir-548h-4	35345	g.44A>G	rs73235382	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr10	hsa-mir-1265	35332	g.44C>T	rs11259096	Terminal loop	by-cluster, by-frequency, by-2hit-2allele, by-hapmap,by-1000genomes	UCSC-dbSNP Mapping
chr10	hsa-mir-1307	35372	g.70A>G	rs7911488	Terminal loop	by-cluster, by-frequency, by-2hit-2allele, by-hapmap	UCSC-dbSNP Mapping
chr10	hsa-mir-202	32080	g.13C>T	rs12355840	Stem	by-2hit-2allele	UCSC-dbSNP Mapping
chr10	hsa-mir-2110	37071	g.34C>T	rs17091403	Stem	by-cluster, by-frequency, by-hapmap	UCSC-dbSNP Mapping
chr10	hsa-mir-603	32859	g.40C>T	rs11014002	Stem	by-cluster, by-frequency, by-hapmap	UCSC-dbSNP Mapping
chr10	hsa-mir-604	32860	g.29C>T	rs2368393	Stem	by-2hit-2allele, by-1000genomes	UCSC-dbSNP Mapping
chr10	hsa-mir-604	32860	g.24C>T	rs2368392	Stem	by-frequency, by-hapmap, by-1000genomes	UCSC-dbSNP Mapping
chr10	hsa-mir-605	32861	g.74A>G	rs2043556	Stem	by-cluster, by-frequency,	UCSC-dbSNP

Chr	miRBase ID***	HGNC ID	Variation*	dbSNP ID**	Location on hairpin structure	SNP Validation status	Strategy Used
						by-submitter, by-2hit-2allele,by-hapmap,by-1000genomes	Mapping
chr10	hsa-mir-608	32864	g.37C>G	rs4919510	Mature	by-cluster, by-frequency, by-2hit-2allele, by-hapmap,by-1000genomes	UCSC-dbSNP Mapping
chr10	hsa-mir-938	33681	g.16C>T	rs12416605	Seed	by-cluster, by-hapmap	UCSC-dbSNP Mapping
chr11	hsa-mir-1304	35302	g.65A>C	rs2155248	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr11	hsa-mir-1908	35392	g.5C>T	rs174561	Stem	by-frequency	UCSC-dbSNP Mapping
chr11	hsa-mir-194-2	31565	g.76A>G	rs11231898	Stem	by-cluster, by-frequency, by-hapmap	UCSC-dbSNP Mapping
chr11	hsa-mir-548l	35292	g.26C>T	rs11020790	Mature	by-cluster, by-frequency, by-2hit-2allele, by-hapmap	UCSC-dbSNP Mapping
chr11	hsa-mir-548l	35292	g.37C>T	rs13447640	Stem	by-cluster, by-frequency, by-hapmap	UCSC-dbSNP Mapping
chr11	hsa-mir-612	32868	g.51G>T	rs550894	Stem	by-2hit-2allele	UCSC-dbSNP Mapping
chr11	hsa-mir-612	32868	g.12A>G	rs12803915	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr12	hsa-mir-1178	35259	g.37C>T	rs7311975	Stem	by-cluster, by-frequency, by-2hit-2allele ,by-hapmap,by-1000genomes	UCSC-dbSNP Mapping
chr12	hsa-mir-196a-2	31568	g.78C>T	rs11614913	Mature	by-cluster, by-frequency, by-hapmap, by-1000genomes	UCSC-dbSNP Mapping
chr12	hsa-mir-492	32081	g.113C>G	rs2289030	Stem	by-cluster, by-frequency, by-hapmap	UCSC-dbSNP Mapping
chr12	hsa-mir-548c	32800	g.12A>G	rs17120527	Stem	by-frequency, by-hapmap	UCSC-dbSNP Mapping

Chr	miRBase ID***	HGNC ID	Variation*	dbSNP ID**	Location on hairpin structure	SNP Validation status	Strategy Used
chr12	hsa-mir-617	32873	g.83C>G	rs12815353	3' Overhang	by-hapmap	UCSC-dbSNP Mapping
chr12	hsa-mir-618	32874	g.77A>C	rs2682818	Stem	by-cluster, by-frequency, by-2hit-2allele	UCSC-dbSNP Mapping
chr13	hsa-mir-92a-1	31643	g.22A>G	rs9589207	Mature	by-cluster, by-frequency, by-hapmap	UCSC-dbSNP Mapping
chr14	hsa-mir-1185-2	35254	g.79A>G	rs11844707	Stem	by-cluster, by-frequency, by-hapmap	UCSC-dbSNP Mapping
chr14	hsa-mir-1260	35325	g.68->T	rs28909969	Stem	by-frequency	UCSC-dbSNP Mapping
chr14	hsa-mir-208b	33669	g.54A>T	rs2754157	Mature	by-hapmap	UCSC-dbSNP Mapping
chr14	hsa-mir-300	33636	g.28C>T	rs12894467	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr14	hsa-mir-624	32880	g.66C>G	rs73251987	Mature	by-1000genomes	UCSC-dbSNP Mapping
chr14	hsa-mir-625	32881	g.70A>C	rs12894182	Mature	by-hapmap	UCSC-dbSNP Mapping
chr15	hsa-mir-1282	35360	g.49G>T	rs11269	Stem	by-frequency	UCSC-dbSNP Mapping
chr15	hsa-mir-147b	33655	g.8C>G	rs56073218	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr15	hsa-mir-631	32887	g.61->CT	rs5745925	Stem	by-frequency	UCSC-dbSNP Mapping
chr17	hsa-mir-193a	31563	g.19G>T	rs60406007	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr17	hsa-mir-423	31880	g.87A>C	rs6505162	3' Overhang	by-cluster, by-frequency, by-2hit-2allele ,by-hapmap,by-1000genomes	UCSC-dbSNP Mapping
chr17	hsa-mir-548h-3	35344	g.40A>G	rs9913045	Mature	by-1000genomes	UCSC-dbSNP Mapping
chr17	hsa-mir-633	32889	g.36A>G	rs17759989	Stem	by-frequency	UCSC-dbSNP

Chr	miRBase ID***	HGNC ID	Variation*	dbSNP ID**	Location on hairpin structure	SNP Validation status	Strategy Used
							Mapping
chr19	hsa-mir-1181	35262	g.56C>G	rs2569788	Stem	by-hapmap	UCSC-dbSNP Mapping
chr19	hsa-mir-220b	33640	g.87C>G	rs1053262	3' Overhang	by-frequency	UCSC-dbSNP Mapping
chr19	hsa-mir-27a	31613	g.40C>T	rs895819	Terminal loop	by-cluster, by-frequency, by-hapmap, by-1000genomes	UCSC-dbSNP Mapping
chr19	hsa-mir-518d	32121	g.59A>G	rs73602910	Seed	by-1000genomes	UCSC-dbSNP Mapping
chr19	hsa-mir-521-2	32113	g.7G>T	rs13382089	Stem	by-frequency, by-hapmap	UCSC-dbSNP Mapping
chr20	hsa-mir-1-1	31499	g.3C>T	rs6122014	Stem	by-hapmap	UCSC-dbSNP Mapping
chr20	hsa-mir-1292	35364	g.44C>T	rs73576045	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr20	hsa-mir-499	32133	g.73C>T	rs3746444	Seed	by-cluster, by-frequency, by-1000genomes	UCSC-dbSNP Mapping
chr20	hsa-mir-499	32133	g.98C>T	rs7267163	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr20	hsa-mir-646	32902	g.74C>T	rs6513496	Mature	by-cluster, by-2hit-2allele, by-1000genomes	UCSC-dbSNP Mapping
chr20	hsa-mir-646	32902	g.3G>T	rs6513497	Stem	by-cluster, by-2hit-2allele	UCSC-dbSNP Mapping
chr20	hsa-mir-647	32903	g.74A>G	rs73147065	Stem	by-1000genomes	UCSC-dbSNP Mapping
chr20	hsa-mir-663	32919	g.91C>T	rs28670321	3' Overhang	by-1000genomes	UCSC-dbSNP Mapping
chr20	hsa-mir-663	32919	g.3A>C	rs7266947	5' Overhang	by-cluster, by-2hit-2allele	UCSC-dbSNP Mapping
chr20	hsa-mir-941-1	33684	g.13C>G	rs7268785	Stem	by-cluster	UCSC-dbSNP Mapping
chr20	hsa-mir-941-1	33684	g.31A>G	rs2427556	Stem	by-cluster, by-2hit-2allele	UCSC-dbSNP

Chr	miRBase ID***	HGNC ID	Variation*	dbSNP ID**	Location on hairpin structure	SNP Validation status	Strategy Used
							Mapping
chr20	hsa-mir-941-1	33684	g.87A>G	rs6089780	Stem	by-cluster, by-2hit-2allele	UCSC-dbSNP Mapping
chr20	hsa-mir-941-3	33686	g.69C>G	rs12625445	Mature	by-cluster, by-2hit-2allele	UCSC-dbSNP Mapping
chr20	hsa-mir-941-3	33686	g.13C>G	rs12625454	Stem	by-cluster, by-2hit-2allele	UCSC-dbSNP Mapping
chr22	hsa-mir-548j	35276	g.105C>G	rs4822739	Stem	by-cluster, by-frequency, by-2hit-2allele, by-hapmap,by-1000genomes	UCSC-dbSNP Mapping
chr22	hsa-mir-650	32906	g.71C>G	rs5996397	Stem	by-2hit-2allele	UCSC-dbSNP Mapping
chrX	hsa-mir-1308	35369	g.23A>G	rs7051072	Stem	by-hapmap	UCSC-dbSNP Mapping
chrX	hsa-mir-532	32795	g.79A>G	rs456615	Stem	by-cluster, by-1000genomes	UCSC-dbSNP Mapping
chrX	hsa-mir-532	32795	g.82A>G	rs456617	Stem	by-cluster, by-1000genomes	UCSC-dbSNP Mapping
chrX	hsa-mir-548f-5	35309	g.85C>T	rs60180387	3’ Overhang	by-1000genomes	UCSC-dbSNP Mapping
chrX	hsa-mir-888	33648	g.77G>T	rs5965660	3’Overhang	by-cluster, by-frequency, by-hapmap ,by-1000genomes	UCSC-dbSNP Mapping
chrX	hsa-mir-891a	33635	g.35A>G	rs5965990	Terminal loop	by-frequency, by-hapmap	UCSC-dbSNP Mapping
chrX	hsa-mir-934	33677	g.9A>G	rs73558572	Stem	by-1000genomes	UCSC-dbSNP Mapping
MT*	hsa-mir-1977	37065	g.5A>G	rs2854138	Seed	by-cluster	UCSC-dbSNP Mapping
MT*	hsa-mir-1977	37065	g.62C>T	rs9783068	Stem	by-cluster	UCSC-dbSNP Mapping
MT*	hsa-mir-1977	37065	g.61A>G	rs41453547	Stem	by-cluster	UCSC-dbSNP Mapping

Chr	miRBase ID***	HGNC ID	Variation*	dbSNP ID**	Location on hairpin structure	SNP Validation status	Strategy Used
MT*	hsa-mir-1977	37065	g.53C>T	rs9701099	Stem	by-cluster, by-2hit-2allele	UCSC-dbSNP Mapping
MT*	hsa-mir-1978	37066	g.44A>G	rs3937039	Mature	by-cluster	UCSC-dbSNP Mapping
chr1	hsa-mir-197	31569	r.14a>i	-	Stem	RNA Editing	Blow et al., 2006
chr8	hsa-mir-151	31762	r.49a>i	-	Seed	RNA Editing	Blow et al., 2006
chr14	hsa-mir-376a-1	31869	r.9a>i	-	Seed	RNA Editing	Blow et al., 2006
chr14	hsa-mir-376a-1	31869	r.49a>i	-	Seed	RNA Editing	Blow et al., 2006
chr14	hsa-mir-379	31872	r.10a>i	-	Seed	RNA Editing	Blow et al., 2006
chr21	hsa-mir-99a	31650	r.13a>i	-	Mature	RNA Editing	Blow et al., 2006
chrX	hsa-let-7f-2	31484	g.11G>A	-	Stem	X linked	Sun et al., 2009
chrX	hsa-mir-188	31559	g.60C>T	-	Seed	X linked	Sun et al., 2009
chrX	hsa-mir-18b	32025	g.32A>G	-	Stem	X linked	Sun et al., 2009
chrX	hsa-mir-223	31603	r.20A>I	-	Stem	RNA Editing	Blow et al., 2006
chrX	hsa-mir-224	31604	g.41G>A	-	Terminal loop	X linked	Sun et al., 2009
chrX	hsa-mir-325	31768	g.66G>A	-	Stem	X linked	Sun et al., 2009
chrX	hsa-mir-421	32793	g.73G>A	-	Stem	X linked	Sun et al., 2009
chrX	hsa-mir-421	32793	4 nt down from 3' end of the stem	-	3’ Overhang	X linked	Sun et al., 2009
chrX	hsa-mir-450a-2	32137	g.4T>C	-	Seed	X linked	Sun et al., 2009
chrX	hsa-mir-502	32136	g.13C>G	-	Stem	X linked	Sun et al., 2009
chrX	hsa-mir-505	32140	g.8C>T	-	Stem	X linked	Sun et al., 2009
chrX	hsa-mir-509-1	32146	g.54insTGA	-	Stem	X linked	Sun et al., 2009
chrX	hsa-mir-509-2	33641	g.9G>T	-	Stem	X linked	Sun et al., 2009
chrX	hsa-mir-509-2	33641	g.11->A	-	Mature	X linked	Sun et al., 2009
chrX	hsa-mir-509-3	33675	g.22G>A	-	Mature	X linked	Sun et al., 2009

Chr	miRBase ID***	HGNC ID	Variation*	dbSNP ID**	Location on hairpin structure	SNP Validation status	Strategy Used
chrX	hsa-mir-509-3	33675	g.19C>G	-	Mature	X linked	Sun et al., 2009
chrX	hsa-mir-509-3	33675	g.13C>T	-	Seed	X linked	Sun et al., 2009
chrX	hsa-mir-510	32147	g.48T>C	-	Stem	X linked	Sun et al., 2009
chrX	hsa-mir-510	32147	g.6G>A	-	Stem	X linked	Sun et al., 2009
chrX	hsa-mir-660	32916	g.15C>T	-	Mature	X linked	Sun et al., 2009
chrX	hsa-mir-888	33648	g.77A>C	-	3' Overhang	X linked	Sun et al., 2009
chrX	hsa-mir-890	33644	g.66G>C	-	Stem	X linked	Sun et al., 2009
chrX	hsa-mir-891b	33645	g.35C>G	-	Terminal loop	X linked	Sun et al., 2009
chrX	hsa-mir-892b	33649	g.15T>C	-	Stem	X linked	Sun et al., 2009
chrX	hsa-mir-934	33677	g.1T>G	-	Mature	X linked	Sun et al., 2009

*Genomic positions are from build hg18 (UCSC)

**All rsIDs are based on dbSNP (build130)

*** All miRNA names (ID) are from miRBase.

· MT*= Mitochondria

Dead Entry in miRBase new release:

- According to miRBase, release 13, hsa-mir-1977 have been found to be present in chromosome 1 and Mitochondria (MT) which overlaps with Mt tRNA sequence so it has been removed from the current release of miRBase.

Supp. Table S2. Functional analysis using PHDcleav and RISCbinder

miRBase ID	HGNC ID	Variation	Wild miRNA Score	Variant miRNA Score	Effect of Variation	Method	Strategy
hsa-mir-608	32864	g.37C>G	0.183	-0.132	Altered incorporation in RISC	RISCbinder	UCSC-dbSNP mapping
hsa-mir-449b	32794	g.27A>G	0.8014721	No site predicted	Altered Dicer Cleavage Site	PHDcleav	UCSC-dbSNP mapping
hsa-mir-1324	35377	g.31C>T	1.513702	No site predicted	Altered Dicer Cleavage Site	PHDcleav	UCSC-dbSNP mapping

Supp. Table S3. Systematic mapped genetic variations previously reported in diseases

Chr	miRbase ID	HGNC ID	dbSNP ID	Disease Associations	References
chr2	hsa-mir-149	31536	rs2292832	Breast cancer	Hu et al., 2008
chr2	hsa-mir-149	31536	rs2292832	Lung cancer	Tian et al., 2009
chr2	hsa-mir-149	31536	rs2292832	Pneumoconiosis (Lung)	Wang et al., 2010
chr5	hsa-mir-146a	31533	rs2910164	Papillary thyroid carcinoma	Jazdzewski et al., 2008
chr5	hsa-mir-146a	31533	rs2910164	Hepatocellular carcinoma	Hu T et al., 2008
chr5	hsa-mir-146a	31533	rs2910164	Psoriasic arthritis	Chatzikyriakidou et al., 2010
chr5	hsa-mir-146a	31533	rs2910164	Prostate cancer	Xu et al., 2010
chr5	hsa-mir-146a	31533	rs2910164	Breast cancer	Shen et al., 2008
chr5	hsa-mir-146a	31533	rs2910164	Ovarian cancer	Shen et al., 2008
chr12	hsa-mir-196a-2	31568	rs11614913	Lung cancer	Tian et al., 2009
chr12	hsa-mir-196a-2	31568	rs11614913	Oesophageal cancer	Ye et al., 2008
chr12	hsa-mir-196a-2	31568	rs11614913	Breast cancer	Hoffman et al., 2009
chr12	hsa-mir-196a-2	31568	rs11614913	Gastric cancer	Peng et al., 2009
chr12	hsa-mir-196a-2	31568	rs11614913	Hepatocellular carcinoma	Qi et al., 2010
chr12	hsa-mir-196a-2	31568	rs11614913	Glioma	Dou et al., 2010
chr12	hsa-mir-196a-2	31568	rs11614913	Dilated cardiomyopathy (Heart)	Zhou et al., 2010
chr12	hsa-mir-196a-2	31568	rs11614913	Head and neck squamous cell carcinoma	Christensen et al., 2010
chr12	hsa-mir-492	32081	rs2289030	Bladder cancer	Yang et al., 2008
chr15	hsa-mir-631	32887	rs5745925	Esophageal cancer	Ye et al., 2008
chr17	hsa-mir-423	31880	rs6505162	Bladder cancer	Yang et al., 2008
chr17	hsa-mir-423	31880	rs6505162	Breast Cancer	Kontorovich et al., 2010
chr17	hsa-mir-423	31880	rs6505162	Oesophageal cancer	Ye et al., 2008
chr19	hsa-mir-27a	31613	rs895819	Breast cancer	Kontorovich et al., 2009; Yang et al., 2009
chr20	hsa-mir-499	32133	rs3746444	Breast cancer	Hu et al., 2008

Chr	miRbase ID	HGNC ID	dbSNP ID	Disease Associations	References
chr20	hsa-mir-499	32133	rs3746444	Lung cancer	Tian et al., 2009
chr20	hsa-mir-499	32133	rs3746444	Dilated cardiomyopathy (Heart)	Zhou et al., 2010