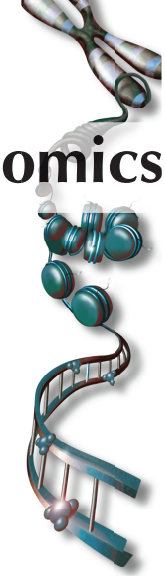
Review

For reprint orders, please contact: reprints@futuremedicine.com

Epigenomics



Classification of miRNA-related sequence variations

Karin Hrovatin[1] & Tanja Kunej*,[1]

[1]Department of Animal Science, Biotechnical Faculty, University of Ljubljana, Domžale, 1230, Slovenia

* Author for correspondence: Tel.: +386 1 320 3890; Fax: +386 1 721 7888; tanja.kunej@bf.uni-lj.si

miRNA regulome is whole set of regulatory elements that regulate miRNA expression or are under control of miRNAs. Its understanding is vital for comprehension of miRNA functions. Classification of miRNA-related genetic variability is challenging because miRNA interact with different genomic elements and are studied at different omics levels. In the present study, miRNA-associated genetic variability is presented at three levels: miRNA genes and their upstream regulation, miRNA silencing machinery and miRNA targets. Several types of miRNA-associated genetic variations are known, including short and structural polymorphisms and epimutations. Differential expression can also affect miRNA regulome function. Classification of miRNA-associated genetic variability presents a baseline for complementing sequence variant nomenclature, planning of experiments, protocols for multi-omics data integration and development of biomarkers.

First draft submitted: 3 October 2017; Accepted for publication: 8 January 2018; Published online: 23 March 2018

Keywords: epigenomics • expression profile • genetic disease • genetic variability • miRNA • miRNA regulome • miRNA target • silencing machinery • SNP • tools for miRNA

miRNAs are a class of short noncoding RNAs and recognized as one of the main regulators of gene expression [1]. They have been shown to be involved in complex regulatory networks, namely, one miRNA targets several targets and one gene could be under control of several miRNAs [2]. miRNAs are connected to numerous physiological processes, disease development and are also an emerging player in DNA damage response [3].

miRNA biogenesis consists of six main steps: miRNA genes are transcribed by RNA polymerase II as the long primary transcripts (pri-miRNA). However, when located downstream of Alu elements, tRNAs, mammalian-wide interspersed repeats or independent miRNA promoters, they are transcribed by RNA polymerase III. They are processed by the ribonuclease DROSHA/DGCR8 complex, releasing an approximately 60 bp hairpin precursor miRNA (pre-miRNA). Pre-miRNA is transported to the cytoplasm via nuclear pore with the aid of XPO5 and RAN. Pre-miRNAs are then processed by the ribonuclease DICER1 to approximately 22 nucleotides long functionally mature miRNAs. Mature miRNAs are complexed together with AGO family proteins, the core unit of the RNA-induced silencing complex (RISC). The miRNA-RISC complex binds target mRNAs and mediates the translational repression or degradation of target mRNAs [4].

Emerging evidence shows that some miRNAs possess nuclear localization signal and are thus translocated in the nucleus with the help of importin-8. Their nuclear functions are dependent on AGO proteins. Nuclear miRNAs recruit chromatin remodeling proteins and affect enhancers, leading to transcriptional gene activation or silencing [5]. For example, miRNA genes located in enhancer loci, such as *MIR24*, both increase the transcription of miRNA neighboring and distal genes via acting on enhancers and surrounding chromatin state in nucleus and silence target mRNAs in cytoplasm [6].

miRNA silencing is one of the epigenetic mechanisms and they have been shown to target very diverse set of targets, including coding regions, 3′ UTRs, 5′ UTRs and promotors [6,7]. Additionally, a subset of miRNA genes, named epi-miRNAs have been shown to be involved in regulation of genes encoding for epigenetic machinery [8]. Therefore, several epigenetic concepts, associated with miRNA exist; miRNAs are silencing/activating target genes, including genes encoding epigenetic machinery; additionally, they could be epigenetically regulated as any other protein-coding gene.

Future Medicine

miRNA-related genetic variability has been associated with predisposition of several diseases, including cancer [9]. However, selecting miRNA candidates for functional studies still poses a challenge. Nevertheless, miRNAs present a potential for biomarker development and therapy. For example, MIR34 is the first miRNA to reach Phase I clinical trials. It has been shown that *MIR34* is expressed in normal cells, but in cancer cells, the expression is downregulated. This dysregulation could be corrected by adding the MRX34; a double stranded mimic encapsulated in liposomes [10].

The field of miRNA variations is very complex and could be classified according to different criteria. For example, according to the location within the miRNA regulome, polymorphisms could be located within miRNA's upstream regions, they could affect downstream targets or they overlap miRNA genes and genes encoding for silencing machinery. miRNA-associated genetic variations could be studied at various omics levels, including genomics/DNA, transcriptomics/RNA, proteomics and epigenomics. miRNA-associated genetic variability could also be classified according to the biotype of sequence variants, such as SNPs and copy number variants (CNVs).

As miRNomics is a relatively new field, complete and systematic classification of miRNA-related polymorphisms has not yet been established. miRNA regulome is defined as a whole set of regulatory elements that either regulate miRNA expression or are under control by miRNA activity. Therefore, it is challenging to classify all the regulome elements and associated sequence variants as each gene encoding for the miRNA regulome component could be a subject of various classes of sequence variants, including short variants or larger structural variants. An additional challenge presents the fact that one class of miRNA-related variability could be placed under two categories. For example, there are two types of polymorphisms associated with silencing machinery: sequence variants within genes, encoding for components of miRNA biogenesis machinery like *DROSHA* and *DICER1*, and sequence variants within miRNA genes affecting DROSHA and DICER1 cleavage sites. Classification is also complicated by the fact that some types of variations (such as polymorphisms at the sites of CpG dinucleotides) affect various miRNA-related elements (e.g., promotor methylation). Therefore, changed miRNA expression profiles could be consequence of both differences in epigenetic mechanisms (DNA methylation or histone modifications) and DNA polymorphisms (short and structural). Furthermore, expression of miRNAs can be affected by their host genes. To solve these issues, two multi-omics data integration protocols for miRNA regulatory atlas development were introduced recently [11,12]. Nevertheless, expression profiles are often studied without factors causing differential expression. Additionally, research studies are not equally distributed across different miRNA regulome elements, for example, miRNA genes are much more explored than genes encoding for silencing machinery.

In the present review, miRNA-associated genetic variations are first presented in three categories according to their role in miRNA regulome: miRNA genes and their upstream regulation, silencing machinery and miRNA targets. Subsequently each of these three categories is then reviewed according to its genetic variability: short, structural and epigenetic mutations. As some of the changes in miRNA regulome expression are not explained on molecular level, we also include the category of expression profiles. Finally, effects of host genes on miRNAs are described. The overview of the miRNA regulome is presented in the Figure 1 and summary of reviewed miRNAs associated with genetic variability in Table 1. miRNA gene names used in this review are corrected in accordance with HGNC nomenclature [13] and [14].

miRNA genes & their upstream regulation

Genetic variability of miRNA genes and their expression profiles have been extensively examined in many diseases, especially cancer. Multiple studies by The Cancer Genome Atlas Research Network on molecular profiles of various cancers included miRNA analysis [15–18].

Short polymorphisms in miRNA genes

Short polymorphisms comprise of SNPs and indels, both of which have been extensively studied in the past. They affect miRNA expression and target recognition. In our previous study, we performed genome-wide screening of miRNA genes in 22 species; 15 animal and seven plant genomes. The number of discovered polymorphisms varied greatly [19,20]. The number of polymorphisms within miRNA genes will most probably increase with time, as current difference in the number of polymorphic miRNA genes is most likely not due to biological reason, but due to differences in activity of sequencing projects between species. Additionally, species with greater number of individuals present bigger genetic pool and thus greater number of identifiable genetic differences.

Table 1. Examples of published studies describing different types of genetic variability associated with various elements of miRNA regulome.

miRNA symbol	Genetic variations or changes in gene expression	Validated targets	Associated phenotype	Species	Study	Year	PMID	Variation associated with gene for	Ref.
Multiple	miRNA-associated short polymorphism	–	–	Six plant species	Fekonja *et al.*	2015	–	miRNA	[20]
Multiple	miRNA-associated short polymorphism	–	–	Various animal species	Zorc *et al.*	2015	25874014	miRNA	[19]
cel-miR-48	miR-rSNP	–	Defects in larval development	*Caenorhabditis elegans*	Li *et al.*	2005	16139229	miRNA	[23]
Multiple	miR-rSNP in TF (*lin-42*), miRNA overexpression	–	Development	*C. elegans*	Perales *et al.*	2014	25032706	TF	[25]
MIR34A, *MIR34B*, *MIR34C*	miR-rSNP, epigenetic regulation	–	–	Human	Strmšek and Kunej	2014	–	miRNA	[11]
MIR137	miR-rSNP	–	Schizophrenia	Human	Warburton *et al.*	2016	26429811	miRNA	[21]
MIR133A1	miR-rSNP	–	Asthma	Human	Zhou *et al.*	2016	27383317	miRNA	[22]
MIR146A	Heterozygosity	Multiple	Papillary thyroid carcinoma	Human	Jazdzewski *et al.*	2009	19164563	miRNA	[30]
MIR627	miR-SNP in seed region	*ATP6V0E1*	–	Human	Gong *et al.*	2012	22045659	miRNA	[26]
multiple	miR-SNP in seed region	–	–	Vertebrates	Zorc *et al.*	2012	22303453	miRNA	[28]
MIR346	miR-SNP in seed region	–	–	Cattle	Zorc *et al.*	2015	25874014	miRNA	[19]
MIR6578	miR-SNP	–	–	Chicken	Zorc and Kunej	2016	26800695	miRNA	[98]
MIR510, *MIR934*	miR-SM-SNP in miRNA affects processing speed	–	Schizophrenia	Human	Sun *et al.*	2009	19617315	miRNA	[33]
MIR96	miR-SM-SNP	–	–	Human	Gong *et al.*	2012	22045659	miRNA	[26]
MIR618	miR-SM-SNP in miRNA affects processing speed	Multiple	Non-Hodgkin lymphoma	Human	Fu *et al.*	2014	24503492	miRNA	[34]
Multiple	miR-SM-SNP	–	–	Human	Obsteter *et al.*	2015	25629077	miRNA	[32]
MIR15, *MIR16*	Deletion	–	Chronic lymphocytic leukemia	Human	Calin *et al.*	2002	12434020	miRNA	[36]
MIR146B	Deletion of TF-regulating miRNA expression	*Irak1*	Rett syndrome	Mouse, human	Urdinguio *et al.*	2010	20716963	TF	[40]
MIR1306, *MIR3618*	Microdeletion	–	Schizophrenia	Human, mouse	Merico *et al.*	2014	25484875	miRNA	[38]
MIR145, *MIR146A*	Chromosomal deletion	–	Thrombocytosis	Human, mouse	Pellagatti and Boultwood	2015	26075044	miRNA	[37]

CNV: Copy number variation; ID: Identification number; LOH: Loss of heterozygosis; miR-rSNP: miRNA regulatory SNP; miR-SM-SNP: miRNA silencing machinery SNP; miR-TS-SNP: miRNA target site SNP; TF: Transcription factor.

10.2217/epi-2017-0126

Table 1. Examples of published studies describing different types of genetic variability associated with various elements of miRNA regulome (cont.).

miRNA symbol	Genetic variations or changes in gene expression	Validated targets	Associated phenotype	Species	Study	Year	PMID	Variation associated with gene for	Ref.
MIR548AQ	Microdeletion	–	Syndromic intellectual disability	Human, zebrafish	Labonne *et al.*	2016	27106595	miRNA	[39]
MIR23A	CNV	*MT2A*	Gastric cancer	Human, nude mouse	An *et al.*	2013	23553990	miRNA	[45]
MIR124	CNV	–	Autism	Human	Vaishnavi *et al.*	2013	23451085	miRNA	[44]
–	CNV	–	Schizophrenia	Human	Warnica *et al.*	2015	25034949	miRNA	[43]
MIR650	CNV	*ING4*	–	Human	Yun *et al.*	2015	26622897	miRNA	[42]
MIR29A, *MIR1256*	miRNA promotor methylation	TRIM68, PGK-1	Prostate cancer	Human	Li *et al.*	2012	22805767	miRNA	[49]
MIR10A	miRNA promotor methylation	–	Cancer (hepatocellular carcinoma)	Human	Shen *et al.*	2012	22976466	miRNA	[50]
MIR941, *MIR1247*	miRNA promotor methylation	*MIR941: TAOK1*, *KDM6B* and *MIR1247: RARA*, *STX1B*, *RCC2*	Gastric cancer	Human	Kim *et al.*	2014	24785261	miRNA	[52]
MIR15, *MIR16*	Underexpression	–	Chronic lymphocytic leukemia	Human	Calin *et al.*	2002	12434020	miRNA	[36]
Multiple	Over/underexpression	–	Breast cancer	Human	Iorio *et al.*	2005	16103053	miRNA	[56]
MIR106A	Over/underexpression	*RB1*	Gastric, prostate and lung tumor and breast tumor, respectively	Human, mouse	Volinia *et al.*	2006	16461460	miRNA	[61]
MIR21 and others	Overexpression	–	Tumors	Human	Schetter *et al.*	2008	18230780	miRNA	[55]
Multiple	Over/underexpression	–	Cancer	Human	Ferdin *et al.*	2010	20218735	miRNA	[65]
Multiple	Over/underexpression	–	Primary breast tumor	Human	Enerly *et al.*	2011	21364938	miRNA	[76]
MIR210	Overexpression	–	Breast cancer	Human	Jung *et al.*	2012	22370716	miRNA	[60]
MIR21, *MIR210*, *MIR126*, *MIR30A*	Over/underexpression	–	Lung cancer	Human	Vosa *et al.*	2013	23225545	miRNA	[66]
MIR21, *MIR155*, *MIR375*	Over/underexpression	–	Pancreatic ductal adenocarcinoma	Human	Ma *et al.*	2013	24289824	miRNA	[75]
MIR34	Underexpression	–	Cancer	Human, mouse	Agostini and Knight	2014	24657911	miRNA	[10]
Multiple	Over/underexpression	–	Breast cancer	Human	Liu *et al.*	2014	25195131	miRNA	[72]

CNV: Copy number variation; ID: Identification number; LOH: Loss of heterozygosis; miR-rSNP: miRNA regulatory SNP; miR-SM-SNP: miRNA silencing machinery SNP; miR-TS-SNP: miRNA target site SNP; TF: Transcription factor.

Table 1. Examples of published studies describing different types of genetic variability associated with various elements of miRNA regulome (cont.).

miRNA symbol	Genetic variations or changes in gene expression	Validated targets	Associated phenotype	Species	Study	Year	PMID	Variation associated with gene for	Ref.
MIR18B, *MIR103*, *MIR107*, *MIR652*	Overexpression	–	Breast cancer	Human	Sahlberg *et al.*	2015	25547678	miRNA	[1]
Multiple	Over/underexpression	–	Nasopharyngeal cancer	Human	Wang *et al.*	2014	25450278	miRNA	[69]
Multiple	Over/underexpression	–	Chronic kidney disease	Human	Zawada *et al.*	2014	24184689	miRNA	[77]
Multiple	Over/underexpression	–	Bladder cancer	Human	Cheng *et al.*	2015	26316777	miRNA	[74]
MIR145, *MIR155*, *MIR382*	Over/underexpression	–	Breast cancer	Human	Cui *et al.*	2015	25296735	miRNA	[73]
MIR130A	Overexpression of *HDAC3* and downregulation of miRNA	*TNF-α*	Spinal cord injury	Human	Ma *et al.*	2015	25973054	Epigenetic machinery	[47]
Multiple	Over/underexpression	–	Renal cell carcinoma	Human	Song *et al.*	2015	25436016	miRNA	[67]
Multiple	Over/underexpression due to heroin use and hepatitis C infection	–	Response to hepatitis C infection	Human	Zhou *et al.*	2015	25572448	miRNA	[62]
MIR1226	Underexpression	–	Colorectal tumors	Human	Butkyte	2016	27019673	miRNA	[80]
Multiple	Over/underexpression	–	Alzheimer's disease	Human	Hu *et al.*	2016	26903857	miRNA	[63]
MIR630	Overexpression of miRNA	*DICER1*	Tumor	Human	Rupaimoole *et al.*	2016	26725326	miRNA	[78]
MIR188	Overexpression	*MLLT4*	Colorectal cancer	Human	Pichler *et al.*	2017	27601590	miRNA	[59]
Multiple	miR-SM-SNP in *DICER1*	–	Nonepithelial ovarian cancer	Human	Anglesio *et al.*	2013	23132766	SM	[83]
Multiple	Mutation in *alg-1*	–	Development	*C. elegans*	Perales *et al.*	2014	25032706	SM	[25]
Multiple	Mutation in signaling protein affects SM expression	–	Medullary thyroid carcinoma		Puppin *et al.*	2014	24569963	SM	[85]
MIRLET7 family and others	miR-SM-SNP in *DROSHA* and *DICER1*	–	Wilms' tumors	Human	Rakheja *et al.*	2014	25190313	SM	[82]
Multiple	miR-SM-SNP	–	Alzheimer's disease	Human	Yilmaz *et al.*	2016	26796812	SM	[84]

CNV: Copy number variation; ID: Identification number; LOH: Loss of heterozygosis; miR-rSNP: miRNA regulatory SNP; miR-SM-SNP: miRNA silencing machinery SNP; miR-TS-SNP: miRNA target site SNP; TF: Transcription factor.

Table 1. Examples of published studies describing different types of genetic variability associated with various elements of miRNA regulome (cont.).

miRNA symbol	Genetic variations or changes in gene expression	Validated targets	Associated phenotype	Species	Study	Year	PMID	Variation associated with gene for	Ref.
MIR1, *MIR206*	miR-TS-SNP	*GDF8*	Muscular hypertrophy	Sheep	Clop *et al.*	2006	16751773	miRNA target	[88]
Multiple	miR-TS-SNP	–	Cancer	Human	Landi *et al.*	2008	17941804	miRNA target	[91]
MIR187, *MIR138*, *MIR638*, *MIR628*	miR-TS-SNP	*TGFB1*, *XRCC1*, *BRCA1*, *TGFBR1* and others	Breast cancer	Human	Nicoloso *et al.*	2010	20332227	miRNA target	[87]
MIR627	miR-TS-SNP	–	–	Human	Gong *et al.*	2012	22045659	miRNA target	[26]
MIR195	miR-TS-SNP	*MRAS*	Coronary artery disease	Human	Haas *et al.*	2012	22664914	miRNA target	[2]
MIR155	miR-TS-SNP	*AGTR1*	Cardiovascular diseases	Human	Haas *et al.*	2012	22664914	miRNA target	[2]
MIR214	miR-TS-SNP	*STAT3*	Hepatocellular carcinoma	Human	Fan *et al.*	2016	27619679	miRNA target	[90]
Multiple	miR-TS-SNP or indel	*SFTPA1*, *SFTPA2*	Lung diseases	Human	Silveyra *et al.*	2014	24793167	miRNA target	[93]
MIR33	miRNA expression dependent on host gene	–	–	Human	Rodriguez *et al.*	2004	15364901	miRNA-host gene	[99]

CNV: Copy number variation; ID: Identification number; LOH: Loss of heterozygosis; miR-rSNP: miRNA regulatory SNP; miR-SM-SNP: miRNA silencing machinery SNP; miR-TS-SNP: miRNA target site SNP; TF: Transcription factor.

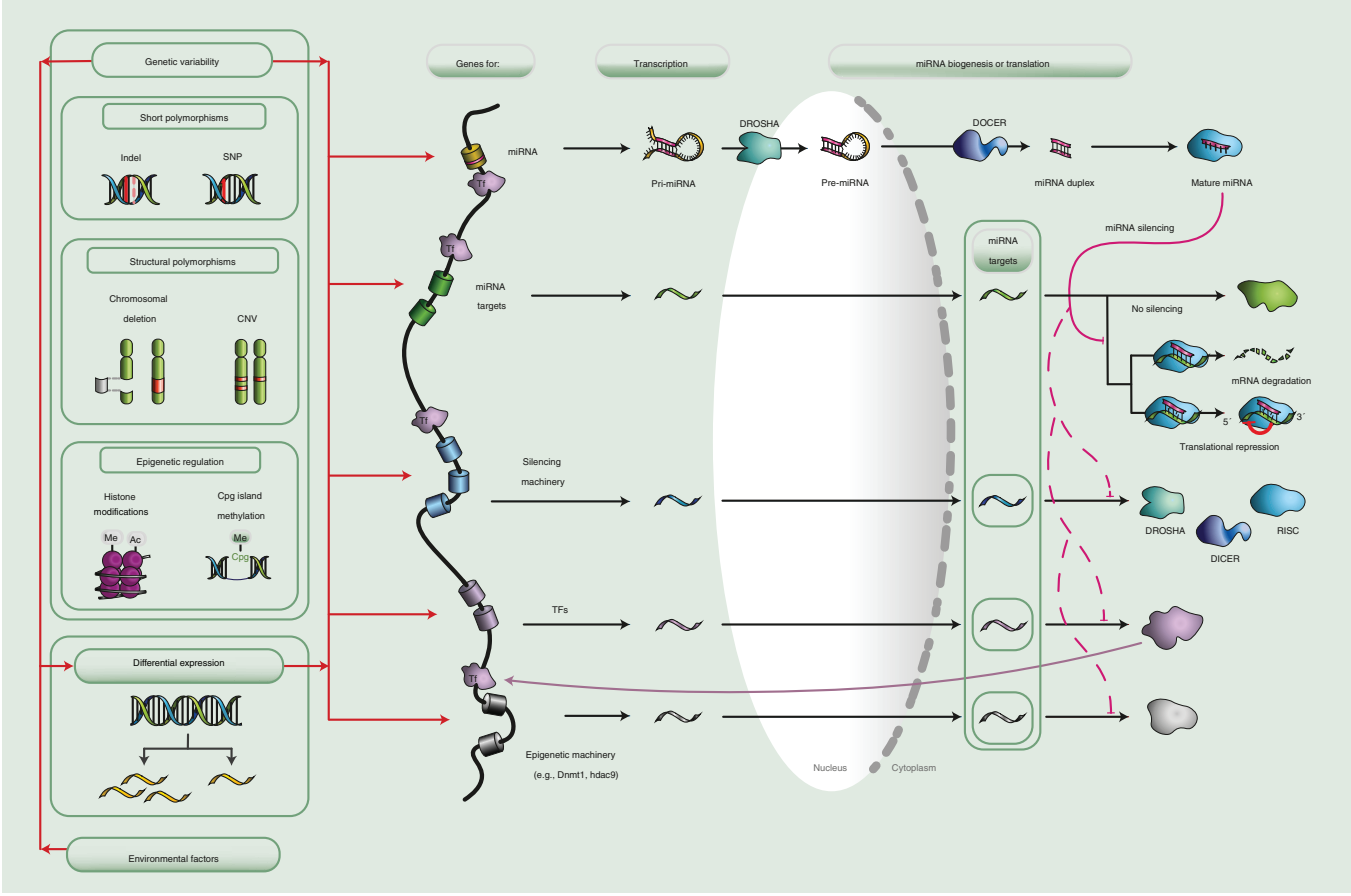


Figure 1. Overview of the miRNA regulome. The figure presents genes encoding for various miRNA regulatory elements: miRNA genes, genes encoding for miRNA targets, components for silencing machinery and genes for miRNA upstream regulators, including TF and enzymes for epigenetic machinery. Various types of genetic variability could affect expression profiles or final products of these genes; short variants, structural variants and epigenetic changes.
Ac: Acetyl group; CNV: Copy number variation; Me: Methyl group; TF: Transcription factor.

miRNA regulatory SNPs

Regulatory SNPs (rSNP) are located within upstream or downstream miRNA regions and are termed as miR-rSNPs. For example, promoter polymorphism rs2660304 may act as predisposing factor for schizophrenia through altering the levels of *MIR137* expression in a genotype-dependent manner [21]. The C alleles of polymorphisms rs8089787 and rs9948906 in promotor region of *MIR133A1* are associated with increased risk for asthma [22]. Defects in larval development of *Caenorhabditis elegans* are observed due to an SNP in upstream region of *cel-mir-48* leading to its premature upregulation [23]. In our previous *in silico* study, we analyzed genetic variability within upstream regions of the *MIR34* family, discovering nine SNPs in each; having *MIR34A* and *MIR34B/C* genes that are located within CpG dinucleotides and could cause gain or loss of the CpG dinucleotides. Additionally, two SNPs resided within transcription factor (TF)-binding sites [11]. These prioritized miRNA polymorphisms present candidates for functional experiments. rSNPs associated with miRNA genes could be associated with various regulatory elements and common regions enriched in TF- and RNA-binding protein binding sites [24]. Mutations in *lin-42*, transcription repressor of *C. elegans*, lead to overexpression of various miRNAs. However, defective *lin-42* does not alter cyclic expression patterns of miRNAs, which have been shown to be regulated by miRNA upstream sequences [25].

Short polymorphisms within miRNA genes

miRNA gene polymorphisms are believed to have a profound effect on target recognition and phenotypic variability because each miRNA is predicted to target hundreds of mRNAs [26]. Sequence variants within miRNA genes include various short mutations, such as indels and SNPs (miR-SNP). miRNA sequence variants have been associated with

several diseases, including cancer [27]. Sequence variants could be located within mature miRNA seed regions, which is responsible for target binding; however, they could also be located within pri-miRNA, pre-miRNA and mature miRNA regions.

Short polymorphisms in miRNA seed regions

Some very polymorphic miRNA genes have been identified, for example, cattle miRNA gene *MIR346* comprises highly polymorphic mature miRNA seed region; six out of seven nucleotides composing seed region are polymorphic [19]. Examples of miRNA genes with seed SNPs, which cause a formation of a seed region annotated to another miRNA, have been reported [28]. Bioinformatics prediction revealed that polymorphisms within mature miRNA seed regions could change the number of target genes [29] and create new targets [26]. A SNP may give rise to heterozygosity of seed regions and thus production of two miRNAs, each with distinct set of suppressed target genes. For example, heterozygosity for an SNP in *MIR146A* leads to change in cell transcriptomes (predicted suppression of more genes than either of homozygotes) and is associated with papillary thyroid carcinoma [30].

Short polymorphisms in miRNA genes affecting processing by silencing machinery

miRNAs processed by miRNA processing machinery have specific sites cleaved by microprocessor machinery [31]. Mutations in those sites may lead to changes in cleavage patterns or processing speed. miRNA genes with polymorphisms overlapping both DROSHA and DICER1 cleavage sites have been reported [32]. DROSHA and DICER1 cutting sites located at the SNP in *MIR934* leads to cleavage offset and thus changed products. Wild-type MIR934 guide strand is 5p, but in mutated allele the 3p has lower 5′ end thermodynamics due to change in cleavage site, leading to it being the preferred guide [33].

In a genome-wide study, SNPs in pre-miRNA stem region were identified. In general, SNPs with decreased stem stability led to reduced production of mature miRNA, as G to A transition in *MIR96*, and vice versa [26]. An SNP, which is likely to lead to changes in secondary structure of *MIR618*, is related to lower processing rate, producing less mature MIR618 and deregulating lymphoma-related genes. However, levels of primary *MIR618* transcript are constant. Therefore, these changes in stem-loop formation may affect processing by DROSHA or DICER1 [34].

Furthermore, most pri-miRNAs tend to produce one main miRNA, highest in abundance, and other miRNA isoforms – isomiRs, which are less common. Sequence and structure features of DROSHA and DICER1 cleavage sites affect cleavage patterns, thus producing isomiRs with variable lengths. DICER1 shows structure and sequence preferences while DROSHA possesses only sequence bias. For example, DICER1 is apt to cleave after U residue even when this results in longer product [31]. However, isomiRs may be also produced due to nontemplate nucleotide addition, limited degradation by exonucleases and AGO2 loading preferences. Such diverse set of miRNAs provides higher regulatory potential although it originates from one miRNA gene [31].

Structural mutations overlapping miRNA genes

Structural mutations comprise various variants, including chromosomal deletions and CNVs. Both lead to differential expression and are related to many human diseases, including numerous cancer types and psychological disorders.

Deletions including miRNA genes & their regulatory elements

Many human miRNAs have been shown to be located in cancer-associated genomic regions, often exactly in minimal regions of loss of heterozygosis or minimal regions of amplification [35]. Several chromosomal deletions associated with various diseases have been shown to include miRNA genes, for example, 5q, 22q11.2 and 12q24.31. A miRNA tumor suppressor cluster *MIR15A/MIR16-1* has been shown to be involved in 13q deletions in chronic lymphocytic leukemia [36]. Loss of the miRNA genes *MIR145* and *MIR146A* in 1.5 Mbp deletion identified by molecular mapping and fluorescent *in situ* hybridization has been associated with the thrombocytosis observed in 5q- syndrome patients [37]. Microdeletions in 22q11.2 are recurrent structural variants that impart a high risk for schizophrenia and are found in up to 1% of all patients with schizophrenia. This annotated 2.6 Mbp deletion region includes seven validated miRNA genes, including *MIR1306* and *MIR3618*. Functional enrichment profiles of the 22q11.2 region miRNA target genes suggested a role in neuronal processes and broader developmental pathways [38]. The 12q24.31 region includes microdeletions associated with syndromic intellectual disability and has been reported to contain multiple genes, including *MIR548AQ*. Deletions spanning 360 kb have been detected with microarrays and quantitative PCR [39].

Furthermore, deletions of transcriptional regulators may also affect miRNA expression. Rett syndrome model mice with knock out of exons 3 and 4 of *Mecp2* have aberrant miRNA levels. To illustrate, *MIR146A* and *MIR146B* are downregulated, leading to lower repression of *Irak1* and thus inflammation of brain tissue [40].

Copy number variations overlapping miRNA genes, CNV-miRNAs

Copy number variations represent gene mutations, wherein level of gene expression is decreased by a deletion or increased by duplication [41]. Therefore, CNV loci encompassing genes may potentially affect gene expression, which can subsequently affect phenotypes and disease development. Beside protein-coding genes, CNVs also overlap miRNA genes. Comparative analysis of genomic locations of 9388 CNVs and 1871 miRNA genes resulted in 38 miRNAs located in CNV regions [42]. The study showed that *MIR650* located in a copy number-variable region is functional in osteosarcoma. This miRNA plays a significant role in the production of IL-6 by regulating *ING4* expression and NF-κB signaling in IL1B-stimulated MG-63 osteosarcoma cells. In schizophrenia, roles of both specific CNVs and miRNAs have been demonstrated in a variety of studies [41]. It has also been shown using a genome-wide approach that in schizophrenia, miRNA loci are enriched in rare CNVs [43]. Multiple miRNAs, located in autism-related CNVs, have been shown to be a part of regulatory network, comprised of miRNAs and their targets, some of them being TFs and miRNA processing machinery components. Such is *MIR124* that regulates the expression of *DICER1* and *XPO5*; therefore, its deregulation may affect other miRNAs [44]. Gastric cancer-specific changes in miRNA expression, for example, *MIR23A*, are result of CNVs overlapping with miRNA genes [45]. Therefore, the discovery of an enrichment of miRNAs across the genome in CNVs suggests that reducing or increasing gene dosage of miRNAs results in perturbation in coordinate gene expression at a genome-wide level [41].

Epigenetic regulation

Since miRNAs are part of complex epigenetic loops, there are several epigenetic concepts associated with miRNA regulatory network, namely, miRNAs post-transcriptionally regulate several target genes, including genes encoding for epigenetic machinery (*DNMT1*, *HDAC* and *PRC*). Additionally, expression of miRNA genes is frequently deregulated via CpG methylation and/or histone modifications [7]. The latter will be discussed below. Mutations affecting epigenetic regulation (epimutations) could also change expression of miRNA genes and may contribute to disease development.

In our recent review [46], we reviewed 63 miRNA genes shown to be epigenetically regulated in association with 21 diseases, including 11 cancer types. Beside DNA methylation, which is further described in next section, histone modifications have also been shown to control miRNA gene expression [47]. Both mechanisms, DNA methylation and histone modifications, cooperate in miRNA regulation. However, a systematic review regarding interplay of mechanisms in miRNA regulation is needed.

DNA methylation of miRNA promoter sequences has been shown to inhibit transcription and thereby interferes with downstream functions of miRNA molecules [48]. Mutations in CpG sites affecting methylation are described in the section 'miRNA regulatory SNPs', whereas methylation profiles are presented in this article. For example, upstream regulatory sequences of *MIR29A* and *MIR1256*, whose downregulation has been associated with prostate cancer, contain many CpG dinucleotide methylation targets [49]. It has been shown that in cancerous tissue, there is more methylation leading to lower miRNA expression and upregulation of targets *TRIM68* and *PGK-1* [49]. Similarly, hepatocellular carcinoma-related hypermethylation of *MIR10A* host gene *HOXB4* leads to downregulation of this miRNA. Higher methylation levels were associated to some hepatocellular carcinoma risk factors, namely, alcohol consumption and viral infection [50]. Hence, substances, such as enoxacin, that induce miRNA expression may be useful in cancer treatment [51]. Ectopic expression of *MIR941* and *MIR1247*, hypermethylated in patients suffering from gastric cancer, was shown to inhibit proliferation and migration of cancer cell lines [52].

Epigenetic silencing of some miRNAs is cancer specific, so those genes have a potential for biomarker development. A catalog of epigenetically regulated miRNA genes has been reported consisting of the following data: miRNA gene, cancer type, cell lines, miRNA targets, DNA methylation status in adjacent tissues and the reference [53]. Integrated data from 150 papers showed 180 miRNA genes reported to be silenced with DNA methylation in 36 cancer types [54]. The data from those papers were fragmented, presentation of the results is not standardized and studies were methodologically heterogeneous, and the study presented the first systematic review toward integration of diverse sets of information. miRNA genes were sorted into two groups: genes that have been shown to be methylated in several cancer types, which have a potential for a general cancer biomarker and genes

10.2217/epi-2017-0126

shown to be methylated in one cancer type and thus present cancer-specific biomarker potential. miRNA gene family *MIR34* has been found to be associated with the highest number of cancer types. Several miRNA–cancer associations have been confirmed in more than one study. However, it is most likely that these data will change with adding information from novel studies. The comparison between two of our publications in 2011 and 2015 revealed that the research on this field is still not improving optimally and more systematic monitoring of the field is needed. Therefore, a decision tree for identification of miRNA genes silenced by DNA methylation in cancer has been proposed [11].

miRNA expression profiles

miRNA expression profiles, also called transcriptional fingerprints or miRNA signatures, have been extensively studied in association with a large number of diseases. In chronic lymphocytic leukemia, *MIR15* and *MIR16* are often underexpressed [36]. On the contrary, high expression of *MIR21* was associated with more advanced tumors and poor therapeutic outcome with higher mortality rate [55]. Similarly, connections between bio-pathological features of breast cancer (such as receptor expression and metastasis) and specific miRNA signatures were analyzed by Iorio *et al.* [56]. Disease-specific miRNA expression profiles have been reported; they also classify human cancers and correlate with the timespan of ongoing treatment [57,58]. For example, high expression of *MIR188* was proposed as diagnostic marker for colorectal cancer [59]. Overexpression of *MIR210* was associated with breast cancer and trastuzumab resistance [60]. Epigenetic changes (described in the 'Epigenetic regulation' section) are one of the key factors affecting expression. Other reasons may be TF mutations, changes in promotor sequences, structural polymorphisms, external factors and short polymorphisms affecting miRNA processing leading to mature miRNAs. Thus, expression profiles are result of various changes in the genome and environment. In this article, we review expression differences that are not explained on the molecular level.

As miRNA expression profiles greatly vary between different tissue types, diagnostic miRNAs for multiple tumor types should be first proven to be characteristic for one particular tumor rather than averaged from combined miRNA cancer profiles. For example, *MIR21*, *MIR17* and *MIR191* are all differentially expressed in multiple cancer types. Also, numerous cancer characteristic miRNAs have univocal signatures (higher or lower expression), indicating a similar tumorigenic mechanism. Many of their targets are cancer genes, for example, *RB1* targeted by MIR106A [61]. Moreover, certain miRNA transcriptional profiles can be associated with external factors, such as heroin use and hepatitis C infection [62].

The results of expression experiments differ between publications and are weakly reproducible. This may be due to different methods for quantitative detection and cut-off values of miRNA level [63,64]. On the other hand, expression studies in cancer have been reviewed and examples of miRNAs reported to be disregulated in more than one paper have been identified [65]. Similarly, *MIR21*, *MIR210*, *MIR126* and *MIR30A* were shown to be unidirectionally dysregulated in lung cancer in more than ten studies. Their target genes are associated with cell signaling [66]. A network of differentially expressed miRNAs, their TFs and target genes has been constructed for renal cell carcinoma. The data were extracted from online databases, such as TarBase for miRNA–target interactions and TransmiR for TF–miRNA interactions, and manually from publications. To unify the gene and miRNA symbols, the NCBI nomenclature was used [67].

Circulating miRNAs, also called fluid-expressed miRNAs, present a good source for biomarker discovery. miRNAs can be packaged in microparticles (exosomes, microvesicles and apoptotic bodies) or associated with RNA-binding proteins (AGO2) or lipoprotein complexes (high-density lipoprotein) to prevent their degradation. Since deregulation of miRNA expression is an early event in tumorigenesis, measuring circulating miRNA levels may be useful for early cancer detection, especially when cancer symptoms are less pronounced. Early diagnosis can contribute to the success of treatment [68]. Furthermore, diagnostics tests based on circulating miRNA are relatively noninvasive, easy to perform and inexpensive [69]. For example, overexpression of serum miRNAs encoded by *MIR18B*, *MIR103*, *MIR107* and *MIR652* can predict poor outcome of breast cancer [1]. However, many miRNAs previously reported as potentially diagnostic are also expressed in blood cells. This poses a problem as white blood cell counts and hemolysis importantly affect the overall circulating miRNA concentration [70].

One of the possible strategies for biomarker development is integrated analysis of heterogeneous gene expression profiles for development of robust disease-specific transcriptional fingerprints. Multiple biological markers were identified for the metastasis of melanoma by collecting the miRNA expression datasets from different platforms deposited in PubMed and Gene Expression Omnibus [71]. Various studies have integrated data presenting miRNA expression profiles in patients (Alzheimer's disease and different cancers) compared with healthy specimens.

Diagnostic miRNA sets consisting of multiple miRNA expression profiles from tissue, blood and urine were constructed [63,69,72–74]. miRNAs show promise in diagnostics, but are often not yet precise enough. High specificity and sensitivity was observed in breast cancer determination assay based on *MIR145*, *MIR155* and *MIR382* concentrations [73]. Diagnostic power of the assays is dependent upon sample types, although studies are not in agreement regarding plasma or serum superiority [72,73]. Meta-analysis of miRNAs expression profiles in pancreatic ductal adenocarcinoma collected from various studies, followed by clinical validation, showed that overexpression of *MIR21* and *MIR155* and underexpression of *MIR375* may be related to poor survival [75]. Combining miRNA–mRNA integrated analyses were performed in association with several diseases, for example, in primary breast tumors [76] and chronic kidney disease [77]. A key challenge in disease classification using miRNA expression profiles is in low overlap between biomarker gene sets obtained in different studies. Several meta-analyses need to be performed for identification of more accurate disease-specific markers.

miRNA silencing machinery

Canonical miRNA biogenesis requires the microprocessor components DROSHA and DGCR8 to generate pre-miRNA and DICER1 to form mature miRNA [14]. Polymorphisms associated with miRNA microprocessor complex, also called miRNA processing machinery, are a source of potential biomarkers. They include polymorphisms, which are either located within: genes encoding for components of miRNA biogenesis, such as *DROSHA*, *DGCR8*, *DICER1* or *XPO5* or miRNA genes overlapping DROSHA/DICER1 cleavage sites (described in the 'Short polymorphisms in miRNA genes affecting processing by silencing machinery' section). Silencing machinery can also be regulated by miRNAs. In hypoxic conditions, *MIR630* is upregulated, leading to higher silencing of its target *DICER1* and subsequent cell metastasis [78].

However, the microprocessor is not required for processing of some miRNAs, originating from DROSHA-independent pathways (including miRtrons) and/or DICER1-independent pathways [14,79]. miRtrons are located within host gene introns and possess 5′ and 3′ end spliceosome splice sites which can overlap with cleavage sites for DROSHA/DGCR8 complex, leading to competitive interactions. Further miRtron processing follows classical pathway in the cytoplasm [79]. Some miRtrons were reported to be associated with human diseases; for example, *MIR1226* is significantly underexpressed in colorectal tumors [80]. Underexpression of miRtrons may be a result of spliceosome mutations, such as in patients with myelodysplastic syndromes; however, miRNAs processed by canonical Drosha-dependent pathway were also differentially expressed [81]. Therefore, more research regarding miRtron-related mutations and their importance in phenotypic variations should be conducted.

Mutations in miRNA processing machinery can importantly reduce expression of tumor suppressing miRNAs, such as *MIRLET7* family. Therefore, they are common in various tumors. Yet, complete depletion of miRNA biogenesis is more unusual, as certain miRNAs are still required for tumor viability [82]. Mutations in *DICER1*, occurring in nonepithelial ovarian cancer, lead to significantly lowered production of all 5p miRNA strands and thus upregulation of their target genes [83]. Missense mutations of *DROSHA*, found in Wilms' tumors, impair miRNA biogenesis more than null mutations, thus acting in dominant-negative manner. As DROSHA functions in pri-miRNA biogenesis, no 5p/3p miRNA strand processing difference has been observed [82]. Mutations in AGO family protein of *C. elegans* reduce levels of mature miRNAs, leading to accumulation of pre-miRNAs and therefore impair adult-specific genetic regulatory programs [25]. As short polymorphisms in silencing machinery have such profound effects on miRNA regulome, they may be useful as biomarkers [84].

Dysregulated expression of silencing machinery components may be a result of polymorphisms within signaling proteins. For example, signaling protein produced from mutated *RET*, related to medullary thyroid carcinoma, can lead to elevated levels of miRNA processing proteins, namely, DICER, DGCR8 and XPO-5 [85].

miRNA targets

A large number of miRNA targets have been reported in several species; however, many of them are related to disease development. To illustrate, targets of MIR618, related to non-Hodgkin lymphoma, have been shown to be included in interaction network enriched in neoplasia-connected genes and centered around *TP53* tumor suppressor gene [34]. In this section, target mutations are described that affect miRNA binding and subsequent epigenetic regulation. However, post-transcriptional modifications of mRNAs can also affect target binding. For example, the use of alternative polyadenylation sites affects the length of 3′ UTRs and thus presence of miRNA target sites. Shorter 3′ UTRs are related to malignant transformation [86].

Short polymorphisms within miRNA targets

Polymorphisms within miRNA target sites affect silencing by destroying or creating new targets or by altering degree of silencing. For example, miRNA target SNPs have been reported in breast cancer [87]. Additionally, changes in miRNA–target interactions have been shown in livestock species. Texel sheep has an illegitimate MIR1 and MIR206 target site in *GDF8*, whose loss-of-function mutations were also proven to cause double muscling in cattle [88]. Furthermore, SNPs near or in target regions might affect miRNA binding due to changed 3D structure of target mRNA [2]. It has been shown that mutated alleles could modulate gene expression by differential interaction with miRNAs. This can have serious consequences as miR-TS-SNP sites influence several diseases, including tumor susceptibility [89]. Hepatocellular carcinoma-related SNP in *STAT3* leads to MIR214 target destruction [90]. A catalog of polymorphisms falling in miRNA-binding regions of cancer genes has been reported [91]; however, an update study in this study field is needed. A set of SNPs in human genome that may lead to target creation was identified in [26]; however, experimental validation is needed.

Furthermore, indels can also affect miRNA–target binding. Therefore, indels are often selected against, being less common in mRNA target sequences than in their 3′ UTRs [92]. An 11 bp indel in *SFTPA1* and *SFTPA2* affects binding of various miRNAs [93]. In a genome-wide study, experimentally proven indels and SNPs in target sites of 213 genes were associated with human disease pathways [92]. Since reporting in this field is very heterogeneous, minimal standards for reporting miRNA–target interactions have been suggested [94].

Structural polymorphisms overlapping target genes

Beside miRNA genes, target genes can be as well found in CNV regions. Moreover, there is significantly more miRNAs with target genes in CNV regions than in non-CNV regions, as well as there is more target sequences in CNV-located genes. Human CNV genes are even more tightly regulated by miRNAs than CNV genes in other animals. Also, human CNV genes host more miRNAs than non-CNV genes and miRNAs often regulate expression of their hosts. This indicates that miRNAs might have evolved as a compensatory mechanism for changes in genetic dosage [95]. Moreover, some transposable elements contain miRNA target sites and their integration in 3′ UTRs of certain genes created novel targets through evolution. A few of possibly active transposable elements possess such miRNA-binding sites, presenting a chance for novel target creation in future [96].

miRNA host genes

Certain miRNAs are located within host genes, which leads to common regulation of host gene and miRNA (discussed further in this section) and specific miRNA processing patterns, for example, regarding miRtrons addressed above in the 'miRNA silencing machinery' section. Genome-wide and interspecies-wide *in silico* screening for miRNA genes located within host genes has been performed. The results showed that miRNA genes could be located within introns and exons of protein-coding genes, two overlapping protein-coding genes, genes encoding for other noncoding RNAs like long intergenic noncoding RNAs or within region overlapping protein-coding gene and gene for small nucleolar RNA [97]. For example, co-location of *MIR6578* and *HADHB* genes has been reported in chicken, which also included a missense polymorphism located within a mature miRNA seed region [98]. Expression of host genes and miRNAs is often correlated [97]. Therefore, the signals that affect host transcription also influence the miRNA. For instance, *MIR33* is located in *SREBP2*, whose expression is dependent on sterol levels [99].

Certain human and mouse miRNAs are located in genes encoding silencing machinery. For example, *MIR3173* resides in *DICER1* and *MIR593* in RISC component *SND1* [97]. Two schizophrenia-related miRNA genes, *MIR1306* and *MIR3618*, overlap *DGCR8*, which encodes a subunit of the miRNA microprocessor complex [38]. It has also been shown that human host genes are enriched for functions associated with regulation of transcription [100].

Tools for analysis of miRNA regulome & its genetic variability

Big datasets that are often produced during miRNA regulome analysis are nearly impossible to analyze manually. Thus, multiple tools for miRNA-related analysis have been developed. Recently, more than 160 tools for miRNA analysis have been collected in tools4miRs platform [101]. Similarly, OMICtools is a database listing various omics tools, including a large number of programs for miRNomic analysis [102]. Tools for miRNA and/or target site prediction are extensively discussed by Rasal *et al.* [103], Riffo-Campos *et al.* [104], Shinre and Bhadra [105] and Afonso-Grunz and Müller [86]. Such reviews greatly benefit the scientific community as specific advantages and shortcomings of individual tools are identified. Additionally, scientists are challenged with frequent new releases of

Table 2. Examples of tools used for miRNA regulome analysis.

Name of the tool	Use	Species	Study	Year	PMID	Ref.
dPORE	Database of SNPs affecting TF-binding sites	Human	Schmeier *et al.*	2011	21326606	[110]
miRviewer	Visualization of homologous miRNAs	Multiple	Kiezun *et al.*	2012	22330228	[107]
omiRas	miRNA differential expression analysis, miRNA–target interactions	Human, mouse, wild boar and several plant species	Müller *et al.*	2013	23946503	[113]
PolymiRTS	Database of polymorphisms affecting miRNAs and miRNA target sites	Human, mouse	Bhattacharya *et al.*	2014	24163105	[111]
OMICtools	Collection of tools for miRNA regulome analysis	–	Henry *et al.*	2014	25024350	[102]
miRBase	Database of known miRNAs	Multiple	Kozomara and Griffiths-Jones	2014	24275495	[106]
BioVLAB-MMIA-NGS	miRNA differential expression analysis	Human, mouse, rhesus monkey, rice	Chae *et al.*	2015	25270639	[114]
miRVaS	Location of polymorphisms within miRNAs, miRNA structural changes due to miRNA polymorphisms	Human and others	Cammaerts *et al.*	2016	26384425	[109]
Tools4miRs	Collection of tools for miRNA and miRNA target analysis	–	Lukasik *et al.*	2016	27153626	[101]
ImiRP	Prediction of illegitimate target sites due to miRNA target polymorphisms	Multiple	Ryan *et al.*	2016	27122020	[112]

TF: Transcription factor.

genomic resources, and large efforts are needed to keep these tools updated with increasing amounts of genomic data.

Examples of programs and databases for analysis of miRNA regulome genetic variability are presented in Table 2. The best known database for miRNA analysis is miRBase [106]. miRNA regulome sequence variations can be studied between species or on multiple individuals within one species. Homologous miRNA sequences of multiple species could be visualized using miRviewer [107]. Many tools allow *in silico* target prediction; however, such approaches are more error prone. Nevertheless, they present a starting point for designing experiments.

On the contrary, experimentally validated miRNA–target pairs can be retrieved from scientific literature or from databases, recently evaluated by Lee *et al.* [108]. However, the authors note that even databases of experimentally validated interactions are often fallible and not regularly updated. Location of polymorphisms within miRNAs greatly affects the resulting molecular phenotype. miRNA functional regions in which polymorphisms occur can be predicted with miRVaS [109]. dPORE-miRNA is an online database for investigation of the potential effects of SNPs on miRNA gene regulation [110]. Polymorphisms in miRNAs and target sites as well as their effects on phenotypes are collected in database PolymiRTS, last updated in 2013 [111]. As intentional changes in target sequence used as controls in interaction tests may create novel target sites (e.g., illegitimate miRNA target sites), a tool for their prediction has been designed. ImiRP generates mutated target sites excluding illegitimate target sites and can be used in miRNA–target interaction reporter assays [112].

Changes in miRNA expression are often studied in relation to various diseases. Differential expression of miRNAs between two sequencing datasets can be analyzed with omiRas, which also constructs mRNA–target interaction networks [113]. As investigation of only miRNA–target sequence complementarily can lead to false results, combined miRNA–mRNA analysis can be applied [86]. BioVLAB-MMIA enables integrated mRNA–miRNA analysis and is compatible with high-throughput platforms, including next-generation sequencing data (e.g., RNA-seq) [114].

Conclusion & future perspective

The present study presents a systematic review of classes of miRNA-associated genetic variability sorted according to genes' function and genetic variation type. This collected and systematized knowledge could contribute to development of several research areas. It could contribute to extend the current version of the sequence variant nomenclature, which is under development by the Sequence Ontology (SO) project [115]. In the SO database, miRNAs are deposited under the SO accession SO:0000276. Integrated information from this review also enables more efficient planning of experimental designs and setting more targeted hypotheses. Classified miRNA-associated

genetic variants will enable development of more efficient prioritization of potential biomarkers and their experimental validation. miRNA-related variants sorted according to single-omics types will also enable development of protocols for multi-omics data integration. The present study could increase awareness of the scientific community for the possibility that their projects are associated with miRNAs, namely, miRNAs are involved in fine-tuning of gene expression via several types of interactions. No matter if analyzing one gene, a set of genes or a whole genome, it is possible that the study is associated with a complex miRNA-related regulatory network.

Systematization of knowledge from this complex field still presents a challenge. One of the possibilities to start a miRNA-related project is to first develop an atlas of miRNA regulatory elements. A protocol for development of miRNA regulatory atlas connecting all known miRNA interactions has been proposed [12]. It could be developed when researching miRNA genes in order to better plan laboratory experiments and to reveal complete miRNA regulatory network. The protocol integrates miRNA upstream regulators, overlapping elements and downstream targets. However, species-specific protocols should be developed, because of the differences of available sources between species. Moreover, significant advance in development of potential biomarkers has been discovered using integrated omics analysis and combining miRNA expression profiles with other data types.

The presented classification of miRNA-associated genetic variability is not final and will further develop in accordance with novel discoveries in miRNomics field. However, this integrated review will enable faster revelation of complete miRNA regulatory network, associations with disease development, biomarker development and targets for therapy.

Financial & competing interests disclosure

This work was supported by the Slovenian Research Agency (ARRS) through the Research program P4-0220. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Executive summary

Classes of miRNA-related sequence variants

- miRNA-related sequence variants can be classified according to various criteria, such as component of miRNA regulome, biotype of a genetic variant and omics level.
- As miRNomics is a relatively new field, systematic classification of miRNA-related polymorphisms needs to be established.

miRNA genes & their upstream regulation

- miRNA genes can be associated with short or structural mutations and epigenetic regulation affecting transcription, miRNA processing and target binding.
- Changes in transcriptional regulation of miRNA genes could lead to development of diseases. miRNA expression profiles could be used for diagnostic purposes.

miRNA silencing machinery

- Mutations in genes encoding for silencing machinery affect processing of a great number of miRNAs.
- Some miRNAs are processed by Drosha/Dicer independent pathways and thus affected by genetic variability of other cellular components such as spliceosomes.

miRNA targets

- Short variations in miRNA target genes may lead to changed interactions between the target and miRNA, and thus altering the level of silencing. Structural variations can cause differential expression of targets, which can be further regulated by miRNAs.
- There are many bioinformatic tools for miRNA target prediction, but results need to be experimentally validated.

miRNA host genes

- miRNAs are often located within host genes and can be affected by their transcription and vice versa.

Tools for analysis of miRNA regulome & its genetic variability

- A great number of bioinformatic tools for miRNA regulome analysis have been designed.
- Major challenges of the field are related with nonregular updates of databases and bioinformatic tools.

Conclusion & future perspective

- The present review enables a systematic overview of sequence variants associated with miRNA regulome, and thus assist in study planning and multi-omics data integration protocols.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1 Kleivi Sahlberg K, Bottai G, Naume B *et al.* A serum microRNA signature predicts tumor relapse and survival in triple-negative breast cancer patients. *Clin. Cancer Res.* 21(5), 1207–1214 (2015).

2 Haas U, Sczakiel G, Laufer SD. MicroRNA-mediated regulation of gene expression is affected by disease-associated SNPs within the 3′-UTR via altered RNA structure. *RNA Biol.* 9(6), 924–937 (2012).

3 Zhang C, Peng G. Non-coding RNAs: an emerging player in DNA damage response. *Mutat. Res. Rev. Mutat. Res.* 763, 202–211 (2015).

4 Ha M, Kim VN. Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* 15(8), 509–524 (2014).

5 Kotipalli A, Gutti R, Mitra CK. Dynamics of miRNA biogenesis and nuclear transport. *J. Integr. Bioinform.* 13(5), 22–34 (2016).

6 Xiao M, Li J, Li W *et al.* MicroRNAs activate gene transcription epigenetically as an enhancer trigger. *RNA Biol.* 14(10), 1326–1334 (2016).

7 Kunej T, Godnic I, Horvat S, Zorc M, Calin GA. Cross talk between microRNA and coding cancer genes. *Cancer J.* 18(3), 223–231 (2012).

8 Fabbri M, Calore F, Paone A, Galli R, Calin GA. Epigenetic regulation of miRNAs in cancer. *Adv. Exp. Med. Biol.* 754, 137–148 (2013).

9 Wojcicka A, de la Chapelle A, Jazdzewski K. MicroRNA-related sequence variations in human cancers. *Hum. Genet.* 133(4), 463–469 (2014).

10 Agostini M, Knight RA. miR-34: from bench to bedside. *Oncotarget* 5(4), 872–881 (2014).

11 Strmsek Z, Kunej T. Data integration of 104 studies related with microRNA epigenetics revealed that miR-34 gene family is silenced by DNA methylation in the highest number of cancer types. *Discoveries* 2(2), e18 (2014).

12 Zorc M, Kunej T. Development of integrative map of microRNA gene regulatory elements. *Microrna* 4(3), 205–208 (2015).

13 Gray KA, Yates B, Seal RL, Wright MW, Bruford EA. Genenames.org: the HGNC resources in 2015. *Nucleic Acids Res.* 43, D1079–D1085 (2015).

14 Desvignes T, Batzel P, Berezikov E *et al.* miRNA nomenclature: a view incorporating genetic origins, biosynthetic pathways, and sequence variants. *Trends Genet.* 31(11), 613–626 (2015).

15 Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* 490(7418), 61–70 (2012).

16 Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487(7407), 330–337 (2012).

17 Cancer Genome Atlas Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 511(7511), 543–550 (2014).

18 Cancer Genome Atlas Network. Genomic and epigenomic landscapes of adult *de novo* acute myeloid leukemia. *N. Engl. J. Med.* 368(22), 2059–2074 (2013).

19 Zorc M, Obsteter J, Dovc P, Kunej T. Genetic variability of microRNA genes in 15 animal species. *J. Genomics* 3, 51–56 (2015).

20 Fekonja S, Zorc M, Kunej T. Genome-wide *in silico* screening of genetic variability of microRNA genes in seven plant species. *Plant Omics* 8(6), 544–548 (2015).

21 Warburton A, Breen G, Bubb VJ, Quinn JP. A GWAS SNP for schizophrenia is linked to the internal MIR137 promoter and supports differential allele-specific expression. *Schizophr. Bull* 42(4), 1003–1008 (2016).

22 Zhou PP, Li Y, Ma ZD, Li ZY, Chen FY, Jiang YX. Single nucleotide polymorphisms in the promoter region of mir-133a-1 and in pre-mir-152 rs1707 may contribute to the risk of asthma in a Chinese Han population. *Eur. Rev. Med. Pharmacol. Sci.* 20(12), 2642–2649 (2016).

23 Li M, Jones-Rhoades MW, Lau NC, Bartel DP, Rougvie AE. Regulatory mutations of mir-48, a *C. elegans* let-7 family microRNA, cause developmental timing defects. *Dev. Cell* 9(3), 415–422 (2005).

24 Jha A, Mehra M, Shankar R. The regulatory epicenter of miRNAs. *J. Biosci.* 36(4), 621–638 (2011).

25 Perales R, King DM, Aguirre-Chen C, Hammell CM. LIN-42, the *Caenorhabditis elegans* PERIOD homolog, negatively regulates microRNA transcription. *PLoS Genetics* 10(7), e1004486 (2014).

26 Gong J, Tong Y, Zhang HM *et al.* Genome-wide identification of SNPs in microRNA genes and the SNP effects on microRNA target binding and biogenesis. *Hum. Mut.* 33(1), 254–263 (2012).

27 Pipan V, Zorc M, Kunej T. MicroRNA polymorphisms in cancer: a literature analysis. *Cancers* 7(3), 1806–1814 (2015).

28 Zorc M, Skok DJ, Godnic I *et al.* Catalog of microRNA seed polymorphisms in vertebrates. *PLoS ONE* 7(1), e30737 (2012).

29 Jevsinek Skok D, Godnic I, Zorc M *et al.* Genome-wide *in silico* screening for microRNA genetic variability in livestock species. *Anim. Genet.* 44(6), 669–677 (2013).

30 Jazdzewski K, Liyanarachchi S, Swierniak M *et al.* Polymorphic mature microRNAs from passenger strand of pre-miR-146a contribute to thyroid cancer. *Proc. Natl Acad. Sci. USA* 106(5), 1502–1505 (2009).

31 Starega-Roslan J, Witkos TM, Galka-Marciniak P, Krzyzosiak WJ. Sequence features of Drosha and Dicer cleavage sites affect the complexity of isomiRs. *Int. J. Mol. Sci.* 16(4), 8110–8127 (2015).

32 Obsteter J, Dovc P, Kunej T. Genetic variability of microRNA regulome in human. *Mol. Genet. Genomic Med.* 3(1), 30–39 (2015).

33 Sun G, Yan J, Noltner K *et al.* SNPs in human miRNA genes affect biogenesis and function. *RNA* 15(9), 1640–1651 (2009).

- **The role of SNPs in miRNA biogenesis.**

34 Fu A, Hoffman AE, Liu R, Jacobs DI, Zheng T, Zhu Y. Targetome profiling and functional genetics implicate miR-618 in lymphomagenesis. *Epigenetics* 9(5), 730–737 (2014).

35 Calin GA, Sevignani C, Dumitru CD *et al.* Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc. Natl Acad. Sci. USA* 101(9), 2999–3004 (2004).

36 Calin GA, Dumitru CD, Shimizu M *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* 99(24), 15524–15529 (2002).

37 Pellagatti A, Boultwood J. Recent advances in the 5q- syndrome. *Mediterr. J. Hematol. Infect. Dis.* 7(1), e2015037 (2015).

38 Merico D, Costain G, Butcher NJ *et al.* MicroRNA dysregulation, gene networks, and risk for schizophrenia in 22q11.2 deletion syndrome. *Front. Neurol.* 5, 238 (2014).

39 Labonne JD, Lee KH, Iwase S *et al.* An atypical 12q24.31 microdeletion implicates six genes including a histone demethylase KDM2B and a histone methyltransferase SETD1B in syndromic intellectual disability. *Hum. Genet.* 135(7), 757–771 (2016).

40 Urdinguio RG, Fernandez AF, Lopez-Nieva P *et al.* Disrupted microRNA expression caused by Mecp2 loss in a mouse model of Rett syndrome. *Epigenetics* 5(7), 656–663 (2010).

41 Morrow EM. MicroRNAs in copy number variants in schizophrenia: misregulation of genome-wide gene expression programs. *Biol. Psychiatry* 77(2), 93–94 (2015).

42 Yun JH, Moon S, Lee HS *et al.* MicroRNA-650 in a copy number-variable region regulates the production of interleukin 6 in human osteosarcoma cells. *Oncol. Lett.* 10(4), 2603–2609 (2015).

43 Warnica W, Merico D, Costain G *et al.* Copy number variable microRNAs in schizophrenia and their neurodevelopmental gene targets. *Biol. Psychiatry* 77(2), 158–166 (2015).

44 Vaishnavi V, Manikandan M, Tiwary BK, Munirajan AK. Insights on the functional impact of microRNAs present in autism-associated copy number variants. *PLoS ONE* 8(2), e56781 (2013).

45 An J, Pan Y, Yan Z *et al.* MiR-23a in amplified 19p13.13 loci targets metallothionein 2A and promotes growth in gastric cancer cells. *J. Cell. Biochem.* 114(9), 2160–2169 (2013).

46 Piletič K, Kunej T. MicroRNA epigenetic signatures in human disease. *Arch. Toxicol.* 90(10), 2405–2419 (2016).

47 Ma YD, Fang J, Liu H, Zhou L. Increased HDAC3 and decreased miRNA-130a expression in PBMCs through recruitment HDAC3 in patients with spinal cord injuries. *Int. J. Clin. Exp. Pathol.* 8(2), 1682–1689 (2015).

48 Baer C, Claus R, Plass C. Genome-wide epigenetic regulation of miRNAs in cancer. *Cancer Res.* 73(2), 473–477 (2013).

49 Li Y, Kong D, Ahmad A, Bao B, Dyson G, Sarkar FH. Epigenetic deregulation of miR-29a and miR-1256 by isoflavone contributes to the inhibition of prostate cancer cell growth and invasion. *Epigenetics* 7(8), 940–949 (2012).

50 Shen J, Wang S, Zhang YJ *et al.* Genome-wide aberrant DNA methylation of microRNA host genes in hepatocellular carcinoma. *Epigenetics* 7(11), 1230–1237 (2012).

51 Sousa E, Graca I, Baptista T *et al.* Enoxacin inhibits growth of prostate cancer cells and effectively restores microRNA processing. *Epigenetics* 8(5), 548–558 (2013).

52 Kim JG, Kim TO, Bae JH *et al.* Epigenetically regulated MIR941 and MIR1247 target gastric cancer cell growth and migration. *Epigenetics* 9(7), 1018–1030 (2014).

53 Kunej T, Godnic I, Ferdin J, Horvat S, Dovc P, Calin GA. Epigenetic regulation of microRNAs in cancer: an integrated review of literature. *Mutat. Res.* 717(1–2), 77–84 (2011).

54 Strmsek Z, Kunej T. MicroRNA silencing by DNA methylation in human cancer: a literature analysis. *Non-Coding RNA* 1(1), 44–52 (2015).

- **Epigenetic regulation of miRNA genes.**

55 Schetter AJ, Leung SY, Sohn JJ *et al.* MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *JAMA* 299(4), 425–436 (2008).

56 Iorio MV, Ferracin M, Liu CG *et al.* MicroRNA gene expression deregulation in human breast cancer. *Cancer Res.* 65(16), 7065–7070 (2005).

- **Correlation between miRNA expression and cancer bio-pathological features.**

57 Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nat. Rev. Cancer* 6(11), 857–866 (2006).

58 Calin GA, Ferracin M, Cimmino A *et al.* A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N. Engl. J. Med.* 353(17), 1793–1801 (2005).

59 Pichler M, Stiegelbauer V, Vychytilova-Faltejskova P *et al.* Genome-wide miRNA analysis identifies miR-188-3p as a novel prognostic marker and molecular factor involved in colorectal carcinogenesis. *Clin. Cancer Res.* 23(5), 1323–1333 (2017).

60 Jung EJ, Santarpia L, Kim J *et al.* Plasma microRNA 210 levels correlate with sensitivity to trastuzumab and tumor presence in breast cancer patients. *Cancer* 118(10), 2603–2614 (2012).

61 Volinia S, Calin GA, Liu CG *et al.* A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl Acad. Sci. USA* 103(7), 2257–2261 (2006).

•• **Regulation of cancer-related genes by differentially expressed miRNA genes.**

62 Zhou Y, Sun L, Wang X *et al.* Heroin use promotes HCV infection and dysregulates HCV-related circulating microRNAs. *J. Neuroimmune Pharmacol.* 10(1), 102–110 (2015).

63 Hu YB, Li CB, Song N *et al.* Diagnostic value of microRNA for Alzheimer's disease: a systematic review and meta-analysis. *Front. Aging Neurosci.* 8, 13 (2016).

64 Fu X, Han Y, Wu Y *et al.* Prognostic role of microRNA-21 in various carcinomas: a systematic review and meta-analysis. *Eur. J. Clin. Invest.* 41(11), 1245–1253 (2011).

65 Ferdin J, Kunej T, Calin GA. Non-coding RNAs: identification of cancer-associated microRNAs by gene profiling. *Technol. Cancer Res. Treat.* 9(2), 123–138 (2010).

66 Vosa U, Vooder T, Kolde R, Vilo J, Metspalu A, Annilo T. Meta-analysis of microRNA expression in lung cancer. *Int. J. Cancer* 132(12), 2884–2893 (2013).

67 Song C, Xu Z, Jin Y, Zhu M, Wang K, Wang N. The network of microRNAs, transcription factors, target genes and host genes in human renal cell carcinoma. *Oncol. Lett.* 9(1), 498–506 (2015).

68 Ghai V, Wang K. Recent progress toward the use of circulating microRNAs as clinical biomarkers. *Arch. Toxicol.* 90(12), 2959–2978 (2016).

69 Wang Z, Chen W, Zhang Y, Xu L, Wang Q. Diagnostic value of circulating microRNAs for nasopharyngeal cancer: a systematic review and meta-analysis. *J. Cancer Res. Ther.* 10(Suppl.), C173–C178 (2014).

70 Pritchard CC, Kroh E, Wood B *et al.* Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. *Cancer Prev. Res.* 5(3), 492–497 (2012).

71 Li JY, Zheng LL, Wang TT, Hu M. Functional annotation of metastasis-associated microRNAs of melanoma: a meta-analysis of expression profiles. *Chin. Med. J. (Engl.)* 129(20), 2484–2490 (2016).

72 Liu L, Wang S, Cao X, Liu J. Analysis of circulating microRNA biomarkers for breast cancer detection: a meta-analysis. *Tumour Biol.* 35(12), 12245–12253 (2014).

73 Cui Z, Lin D, Song W, Chen M, Li D. Diagnostic value of circulating microRNAs as biomarkers for breast cancer: a meta-analysis study. *Tumour Biol.* 36(2), 829–839 (2015).

74 Cheng Y, Deng X, Yang X *et al.* Urine microRNAs as biomarkers for bladder cancer: a diagnostic meta-analysis. *OncoTargets Ther.* 8, 2089–2096 (2015).

75 Ma MZ, Kong X, Weng MZ *et al.* Candidate microRNA biomarkers of pancreatic ductal adenocarcinoma: meta-analysis, experimental validation and clinical significance. *J. Exp. Clin. Cancer Res.* 32, 71 (2013).

76 Enerly E, Steinfeld I, Kleivi K *et al.* miRNA-mRNA integrated analysis reveals roles for miRNAs in primary breast tumors. *PLoS ONE* 6(2), e16915 (2011).

77 Zawada AM, Rogacev KS, Muller S *et al.* Massive analysis of cDNA ends (MACE) and miRNA expression profiling identifies proatherogenic pathways in chronic kidney disease. *Epigenetics* 9(1), 161–172 (2014).

78 Rupaimoole R, Ivan C, Yang D *et al.* Hypoxia-upregulated microRNA-630 targets Dicer, leading to increased tumor progression. *Oncogene* 35(33), 4312–4320 (2016).

79 Mattioli C, Pianigiani G, Pagani F. Cross talk between spliceosome and microprocessor defines the fate of pre-mRNA. *Wiley Interdiscipl. Rev. RNA* 5(5), 647–658 (2014).

80 Butkyte S, Ciupas L, Jakubauskiene E *et al.* Splicing-dependent expression of microRNAs of miRtron origin in human digestive and excretory system cancer cells. *Clin. Epigenetics* 8, 33 (2016).

81 Aslan D, Garde C, Nygaard MK *et al.* Tumor suppressor microRNAs are downregulated in myelodysplastic syndrome with spliceosome mutations. *Oncotarget* 7(9), 9951–9963 (2016).

82 Rakheja D, Chen KS, Liu Y *et al.* Somatic mutations in DROSHA and DICER1 impair microRNA biogenesis through distinct mechanisms in Wilms tumours. *Nat. Commun.* 2, 4802 (2014).

83 Anglesio MS, Wang Y, Yang W *et al.* Cancer-associated somatic DICER1 hotspot mutations cause defective miRNA processing and reverse-strand expression bias to predominantly mature 3p strands through loss of 5p strand cleavage. *J. Pathol.* 229(3), 400–409 (2013).

84 Gorucu Yilmaz S, Erdal ME, Avci Ozge A, Sungur MA. SNP variation in microRNA biogenesis pathway genes as a new innovation strategy for Alzheimer disease diagnostics: a study of 10 candidate genes in an understudied population from the eastern Mediterranean. *Alzheimer Dis. Assoc. Dis.* 30(3), 203–209 (2016).

85 Puppin C, Durante C, Sponziello M *et al.* Overexpression of genes involved in miRNA biogenesis in medullary thyroid carcinomas with RET mutation. *Endocrine* 47(2), 528–536 (2014).

86 Afonso-Grunz F, Muller S. Principles of miRNA–mRNA interactions: beyond sequence complementarity. *Cell. Mol. Life Sci.* 72(16), 3127–3141 (2015).

• **Discussion of laboratory and *in silico* approaches for miRNA analysis.**

87 Nicoloso MS, Sun H, Spizzo R *et al.* Single-nucleotide polymorphisms inside microRNA target sites influence tumor susceptibility. *Cancer Res.* 70(7), 2789–2798 (2010).

88 Clop A, Marcq F, Takeda H *et al.* A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nat. Genet.* 38(7), 813–818 (2006).

•• **Effects of a miRNA target mutation on animal phenotype and breed.**

89 Iuliano R, Vismara MF, Dattilo V, Trapasso F, Baudi F, Perrotti N. The role of microRNAs in cancer susceptibility. *Biomed. Res. Int.* 2013, 9 (2013).

90 Fan Y, Qian X, Zhang C. U/G SNP rs111904020 in 3′UTR of STAT3 regulated by miR-214 promotes hepatocellular carcinoma development in Chinese population. *Tumour Biol.* 37(11), 14629–14635 (2016).

91 Landi D, Gemignani F, Barale R, Landi S. A catalog of polymorphisms falling in microRNA-binding regions of cancer genes. *DNA Cell Biol.* 27(1), 35–43 (2008).

92 Bhattacharya A, Ziebarth JD, Cui Y. Systematic analysis of microRNA targeting impacted by small insertions and deletions in human genome. *PLoS ONE* 7(9), e46176 (2012).

93 Silveyra P, Diangelo SL, Floros J. An 11-nt sequence polymorphism at the 3′UTR of human SFTPA1 and SFTPA2 gene variants differentially affect gene expression levels and miRNA regulation in cell culture. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 307(1), L106–L119 (2014).

94 Piletic K, Kunej T. Minimal standards for reporting microRNA: target interactions. *Omics* 21(4), 197–206 (2017).

95 Dweep H, Georgiou GD, Gretz N, Deltas C, Voskarides K, Felekkis K. CNVs-microRNAs interactions demonstrate unique characteristics in the human genome. An interspecies *in silico* analysis. *PLoS ONE* 8(12), e81204 (2013).

96 Spengler RM, Oakley CK, Davidson BL. Functional microRNAs and target sites are created by lineage-specific transposition. *Hum. Mol. Genet.* 23(7), 1783–1793 (2014).

97 Godnic I, Zorc M, Jevsinek Skok D *et al.* Genome-wide and species-wide *in silico* screening for intragenic microRNAs in human, mouse and chicken. *PLoS ONE* 8(6), e65165 (2013).

98 Zorc M, Kunej T. *In silico* screening of the chicken genome for overlaps between genomic regions: microRNA genes, coding and non-coding transcriptional units, QTL, and genetic variations. *Chromosome Res.* 24(2), 225–230 (2016).

99 Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A. Identification of mammalian microRNA host genes and transcription units. *Genome Res.* 14(10A), 1902–1910 (2004).

100 Paczynska P, Grzemski A, Szydlowski M. Distribution of miRNA genes in the pig genome. *BMC Genetics* 16, 6 (2015).

101 Lukasik A, Wojcikowski M, Zielenkiewicz P. Tools4miRs – one place to gather all the tools for miRNA analysis. *Bioinformatics* 32(17), 2722–2724 (2016).

102 Henry VJ, Bandrowski AE, Pepin AS, Gonzalez BJ, Desfeux A. OMICtools: an informative directory for multi-omic data analysis. *Database (Oxford)* 2014, pii:bau069 (2014).

103 Rasal KD, Nandanpawar PC, Swain P, Badhe MR, Sundaray JK, Jayasankar P. MicroRNA in aquaculture fishes: a way forward with high-throughput sequencing and a computational approach. *Rev. Fish Biol. Fisheries* 26(2), 199–212 (2016).

104 Riffo-Campos AL, Riquelme I, Brebi-Mieville P. Tools for sequence-based miRNA target prediction: what to choose? *Int. J. Mol. Sci.* 17(12), 1987 (2016).

105 Shinde S, Bhadra U. A complex genome-microRNA interplay in human mitochondria. *Biomed. Res. Int.* 2015, 206382 (2015).

•• **Presentation of bioinformatic tools for analysis of miRNA regulome.**

106 Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* 42, D68–D73 (2014).

107 Kiezun A, Artzi S, Modai S, Volk N, Isakov O, Shomron N. miRviewer: a multispecies microRNA homologous viewer. *BMC Res. Notes* 5, 92 (2012).

108 Lee YJ, Kim V, Muth DC, Witwer KW. Validated microRNA target databases: an evaluation. *Drug Dev. Res.* 76(7), 389–396 (2015).

• **Evaluation of miRNA target databases and their shortcomings.**

109 Cammaerts S, Strazisar M, Dierckx J, Del Favero J, De Rijk P. miRVaS: a tool to predict the impact of genetic variants on miRNAs. *Nucleic Acids Res.* 44(3), e23 (2016).

110 Schmeier S, Schaefer U, MacPherson CR, Bajic VB. dPORE-miRNA: polymorphic regulation of microRNA genes. *PLoS ONE* 6(2), e16657 (2011).

111 Bhattacharya A, Ziebarth JD, Cui Y. PolymiRTS Database 3.0: linking polymorphisms in microRNAs and their target sites with human diseases and biological pathways. *Nucleic Acids Res.* 42, D86–D91 (2014).

112 Ryan BC, Werner TS, Howard PL, Chow RL. ImiRP: a computational approach to microRNA target site mutation. *BMC Bioinformatics* 17, 190 (2016).

113 Muller S, Rycak L, Winter P, Kahl G, Koch I, Rotter B. omiRas: a web server for differential expression analysis of miRNAs derived from small RNA-Seq data. *Bioinformatics* 29(20), 2651–2652 (2013).

114 Chae H, Rhee S, Nephew KP, Kim S. BioVLAB-MMIA-NGS: microRNA-mRNA integrated analysis using high-throughput sequencing data. *Bioinformatics* 31(2), 265–267 (2015).

115 Eilbeck K, Lewis SE, Mungall CJ *et al.* The sequence ontology: a tool for the unification of genome annotations. *Genome Biol.* 6(5), R44 (2005).