

RESEARCH ARTICLE

# The Sequence and Structure Determine the Function of Mature Human miRNAs

Katarzyna Rolle<sup>\*☯</sup>, Monika Piwecka<sup>☯</sup>, Agnieszka Belter<sup>☯</sup>, Dariusz Wawrzyniak, Jaroslaw Jeleniewicz, Mirosława Z. Barciszewska, Jan Barciszewski<sup>\*</sup>

Institute of Bioorganic Chemistry of the Polish Academy of Sciences, Noskowskiego 12/14, 61–704 Poznan, Poland

☯ These authors contributed equally to this work.

\* [Jan.Barciszewski@ibch.poznan.pl](mailto:Jan.Barciszewski@ibch.poznan.pl) (JB); [Katarzyna.Rolle@ibch.poznan.pl](mailto:Katarzyna.Rolle@ibch.poznan.pl) (KR)



CrossMark  
click for updates

## OPEN ACCESS

**Citation:** Rolle K, Piwecka M, Belter A, Wawrzyniak D, Jeleniewicz J, Barciszewska MZ, et al. (2016) The Sequence and Structure Determine the Function of Mature Human miRNAs. PLoS ONE 11(3): e0151246. doi:10.1371/journal.pone.0151246

**Editor:** M. Lucrecia Alvarez, Mayo Clinic Arizona, UNITED STATES

**Received:** October 9, 2015

**Accepted:** February 15, 2016

**Published:** March 31, 2016

**Copyright:** © 2016 Rolle et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was supported by the European Regional Development Fund within the Innovative Economy Programme, action 1.3.1. (POIG.01.03.01-30-050/09), <http://www.poig.2007-2013.gov.pl/Strony/default.aspx>, the National Science Centre, grant 5955/B/P01/2010/38, <https://www.ncn.gov.pl/?language=en>, and the Ministry of Science and Higher Education of the Republic of Poland by KNOW program. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Micro RNAs (miRNAs) (19–25 nucleotides in length) belong to the group of non-coding RNAs are the most abundant group of posttranscriptional regulators in multicellular organisms. They affect a gene expression by binding of fully or partially complementary sequences to the 3'-UTR of target mRNA. Furthermore, miRNAs present a mechanism by which genes with diverse functions on multiple pathways can be simultaneously regulated at the post-transcriptional level. However, little is known about the specific pathways through which miRNAs with specific sequence or structural motifs regulate the cellular processes. In this paper we showed the broad and deep characteristics of mature miRNAs according to their sequence and structural motifs. We investigated a distinct group of miRNAs characterized by the presence of specific sequence motifs, such as UGUGU, GU-repeats and purine/pyrimidine contents. Using computational function and pathway analysis of their targeted genes, we were able to observe the relevance of sequence and the type of targeted mRNAs. As the consequence of the sequence analysis we finally provide the comprehensive description of pathways, biological processes and proteins associated with the distinct group of characterized miRNAs. Here, we found that the specific group of miRNAs with UGUGU can activate the targets associated to the interferon induction pathway or pathways prominently observed during carcinogenesis. GU-rich miRNAs are prone to regulate mostly processes in neurogenesis, whereas purine/pyrimidine rich miRNAs could be involved rather in transport and/or degradation of RNAs. Additionally, we have also analyzed the simple sequence repeats (SSRs). Their variation within mature miRNAs might be critical for normal miRNA regular activity. Expansion or contraction of SSRs in mature miRNA might directly affect its mRNA interaction or even change the function of that distinct miRNA. Our results prove that due to the specific sequence features, these molecules can also be involved in well-defined cellular processes depending on their sequence contents. The pathway mapping and theoretical gene target identification allowed us to create a biological framework to show the relevance of the specific miRNAs in regulation the distinct type of targets.

**Competing Interests:** The authors have declared that no competing interests exist.

## Introduction

Among all biological macromolecules, RNAs are one of the most structurally and functionally diverse molecular players within the cell. They fulfill many different functions such as coding and transferring genetic information, controlling various cellular processes, as well as catalyzing cleavage reactions, synthesis and ligations [1]. RNAs execute these roles through their variable and dynamic tertiary structures, which enable the RNA to interact specifically with themselves, with other RNAs, ligands and RNA-binding proteins [1]. miRNAs are a class of endogenous, non-coding small RNA, 19–25 nucleotides in length, that mediate the post-transcriptional regulation of mRNAs expression. They are evolutionary conserved, and control diverse biological processes as developmental timing, differentiation, cell proliferation, cancer or neurodegenerative disorders and apoptosis [2–9]. In the canonical pathway, the regulatory function of miRNAs is achieved primarily by binding target mRNAs at sites frequently located in the 3' untranslated region (UTR). However, miRNA binding sequences have been identified also in the coding region and in 5'UTR [10–11]. miRNA activity requires its incorporation in a RNA-induced silencing complex (RISC). Here, a target recognition relies mainly on near-perfect complementarity of the mRNA with the miRNA “seed region”, 6–8 nucleotide-long sequence at the 5' end of miRNA. Upon target binding, repression of gene expression is accomplished by suppressing translation, but some recent reports suggest that the major role of miRNAs is the reduction of target mRNA stability [2, 12–19]. The specificity of miRNA targeting in mammals is not restricted to the 5' and 3' ends of the seed region and Watson-Crick binding of the seed is not always sufficient for the repression of targets in mammals [20–21]. It has already been reported that the residues in positions 13–16 from the 5' end of the miRNAs are involved in stabilization of miRNA-mRNA interactions, especially when base-pairing in the 5' seed region is suboptimal [8, 20–22]. Unlike G and U residues, which can form wobble pairs (C/U or A/G), A residue can only pair with U residue of target mRNAs. It is known that sequence at 5' end of miRNA determines the preferential strand selection and/or rejection in maturation process, between two strands of the same precursor duplex [21]. Therefore, pronounced occurrence of U at the 5' end enhances the preference of the strand selection.

Thus, mature miRNA can control the expression of thousands of target mRNAs, and on the other hand, it is believed that a single mRNA is targeted by multiple miRNAs [20–23]. However, due to the current, incomplete understanding of the mechanism of miRNAs action beyond their biogenesis and target inhibition, the last process is somewhat unpredictable and often not as efficient as expected.

Recently, some non-canonical pathways of miRNA biogenesis have been proposed [24–28]. The analyses of data from next generation sequencing or microarrays, have revealed subclasses of miRNA species, derived from alternative biogenesis pathways. They only partially meet the classical definition, such as mirtrons—miRNAs derived from snoRNAs, from endogenous short-hairpin RNAs (endo-shRNAs) or tRNA precursors [27–28].

The broad analysis of the role of miRNA-target base-pairing in regulation by miRNAs has been shown that this interaction is essential for miRNA specificity in regulatory processes [22, 28]. Although the strength of the interaction is one of the essential factors that can alter the stability of a dsRNA duplex, the contribution of the base-pairing potency to miRNA-target duplex stability remains still unclear. There was also already shown that the sequence of the particular class of miRNA is important for mRNA repression, e.g. in *Arabidopsis* stress-responsive miRNAs were found to be G/C rich [29]. This nucleotide content increases the miRNA-target duplex stability, which accelerates the stress response.

miRNA-mediated gene regulation and mRNA expression are undoubtedly dependent on both nucleic acid conformation and target recognition. We assumed then, if base-pairing

potency could affect the functionality of the miRNA-mediated gene expression regulation, the full understanding of various biological functions of miRNA needs the knowledge about the sequence and structure of mRNA targets, miRNAs precursors but also of mature miRNAs is needed.

In this paper, we consistently studied across the mature human miRNAs: (i) the length distribution; (ii) characteristics of nucleotides content, their occupancy position, sequence-based motifs, like: UGUGU, GU-repeats and purine/pyrimidine contents; (iii) simple sequence repeats (SSRs). Additionally, we have also examined the structural tetranucleotide motifs (UUCG, GAAA, GCAA, GAGA, GUGA, GGAA, CUUG, UUUG) involved in stable hairpin formation and their propensity for regulation of specific group of targets. Using ModeRNA software, for the first time we have modeled the tertiary structure of mature miRNAs, suggesting that the miRNAs-target recognition can be not only sequence-, but also structure-related. Thus, in this paper we investigate the sequence features of mature miRNAs, convincing that the nature of these analysis can reveal the complexity of the miRNA molecular structures, what can give a hint for a genetic, including cancer, research. To predict the biological functions behind our sequence analysis results, we search for the corresponding pathways of the distinct group of miRNAs using DNA intelligent (DIANA) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses. Accordingly, with DIANA-miRPath software we performed a computational target prediction combined with pathway analysis. Apart from the pathways analysis, the gene ontology including biological processes was also taken into account. The validity of the KEGG pathway clustering approach is further supported by the additional analysis with PANTHER classification system, that provided also the list of different protein classes being involved in the distinct processes and pathways. Based on these analysis we have found a link between the sequence content and pathways, biological processes and proteins associated with the distinct group of characterized miRNAs.

## Materials and Methods

### Library construction and clustering

All metazoan and human miRNAs sequences, available on 17<sup>th</sup> October 2012, were downloaded from miRBASE version 19 in FASTA format (<ftp://ftp.sanger.ac.uk/pub/mirbase/19.0/>). The data were scanned for unique sequences of mature miRNAs. The molecules with identical sequences, but annotated under different names in the miRBASE, were removed. The cured library contained 2042 human miRNA sequences.

### Analysis of miRNA nucleotide sequences and SSR analysis

The analysis was done with the use of Phyton language, version 2.7.3. For each of the analyses, different scripts were prepared ([S1 file](#)). Whole sequence analysis was conducted on a specific number of sequences, to a single one accuracy.

For calculation of the monorepeats in a single miRNA sequences, the regular expressions language was used.

Di-, tri-, tetra- and pentanucleotide repeats were identified and localized by the software SSRIT (<http://www.gramene.org/db/searches/ssrtool>).

### Calculation SSR relative count

SSR relative count is the total repeats per miRNA on average. The calculated number of repeats were divided by the number of miRNAs is on each analyzed group (e.g. relative count for mononucleotide repeats  $2326/2042 = 1.14$ ).

## Functional analysis with DIANA miRPath

For the identification of the networks and pathways of the selected miRNAs we used DIANA miRPath (v2.0) software (<http://diana.imis.athena-innovation.gr>) [30]. The prediction algorithm DNA Intelligent Analysis (DIANA) DIANA-microT-CDS (v5.0) was used for the identification of potential target RNAs from each cluster. A core analysis was performed to recognition of the most relevant miRNA targets, canonical pathways, biological functions and physiological processes from the interactions provided by the DIANA database. The identification of all the significantly targeted by the selected miRNA pathways was performed in the mode “Union of pathways”. The enrichment analysis and calculation of the significance levels (p-values) for each selected miRNA individually was performed. Fisher’s meta-analysis method was applying for calculation a merged p-value for each pathway. For the analysis a *posteriori* approach was used, where the statistical results showed the probability that the surveyed pathway is significantly enriched with gene targets of at least one selected miRNA. The significance for all the miRNA-mRNA pairs in a pathway were performed and calculated, followed by the combination into a merged P-value for each pathway. The results are reported and visualized as heat maps, and the pathways are clustered based on significance levels.

## Classification and pathway analysis with PANTHER System

The list of miRNAs with distinct sequence motif were subjected to TargetScanHuman program, release 7.0 (August 2015) analysis prior to obtain the full list of predicted targets for the specific group of miRNAs. The lists of targets were then uploaded to the PANTHER (protein annotation through evolutionary relationship) Classification System (<http://pantherdb.org>). By overrepresentation test the comparison of provided data set with a reference list in PANTHER database was performed. Ontology categories for over- and underrepresentation were determined statistically with Binominal distribution test.

## miRNA sequence alignment

ClustalW command line version of the multiplatform sequence alignment was used for comparison of distinct miRNAs sequences and other RNAs with previously, experimentally confirmed hairpin structures.

## Secondary structure analysis

Mfold program, version 3.5, was used to calculate the folding free energy in the conditions given by the program for all the sequences towards formation of hairpin structures (<http://mfold.bioinfo.rpi.edu>).

## miRNAs structure modeling

ModeRNA program (<http://iimcb.genesilico.pl/moderna/>) for miRNAs 3D modeling based on the experimentally confirmed structures (templates) was used [31]. The structure for RNAs were downloaded from PDB database. At least 80% of similarity was taken as a threshold with the sequences alignment. The PyMOL program was applied as a tool for figures generation.

## Results

### The nucleotide composition in mature human miRNAs

The analysis of a length heterogeneity of 2042 sequences of human mature miRNAs resulted that number of nucleotides is a discrete random variable, which ranges from 16–27. miRNAs

sequences with 21 (12%) and 23 (14%) nucleotides are most abundant, nevertheless miRNAs with 22 nucleotides in length are predominant (47%) (S1 Fig). These results corroborate with the recently published observations [32]. The sequence length analysis of human mature miRNAs showed that the length heterogeneity is related to some biological factors such as evolution conservations or miRNA's regulatory mechanism [32]. It has been also established that miRNAs that regulate cancer-associated targets (oncogenes/tumor suppressors) show stable lengths of 22 nucleotides, whereas longer miRNAs tend to regulate more genes [32].

Furthermore we analyzed the nucleotides content in human mature miRNAs, what showed inequality of base presence. It showed a prevalence of guanosine and uridine at 29 and 26%, and lower levels of adenosine (23%) and cytidine (22%), respectively. This observations is also consistent with the previously shown results [32]. Although the most abundant nucleotide in miRNAs is guanosine, the first position at 5' end is more frequently occupied by the uridine (35%) (Fig 1a and 1b). Furthermore, the second and third position are usually occupied by adenosine, making the 5' end (nucleotides position- 1–3) mostly UAA-rich. The 5' end for the most represented lengths of mature miRNAs: 21 and 22 nt long is UAA-rich, what is shown on Fig 1a. miRNAs with 23 nt in length at the 5' end show mostly UGG sequence. The seed sequence which is considered as position 2–8 is basically G-rich for 21 and 23 nt long miRNAs, whereas for 22 nt long- U/A rich.

However, one can notice, that A nucleotide at 2<sup>nd</sup> and 3<sup>rd</sup> positions is only slightly overrepresented than G. (Fig 1a and 1b). It has been also observed that the presence of A residue in the targeted mRNA, from position 11 of the 5' end of a miRNA, has a great impact on the improvement of the miRNAs binding to the specific target [20, 33]. This strengthen of the interaction is observed, although it was shown, that the exact base-pairing between the A residue and the nucleotide at position 1 of the 5' end of the miRNA is not necessary [20, 22]. The strongest representation of the U residue in the first and the last position of miRNAs sequence was observed not only for human, but also five other mammalian species, such as *Pongo pygmaeus*, *Pan troglodytes*, *Macaca mulatta*, *Mus musculus* and *Rattus norvegicus*. This observation suggests that both ends, more likely than just the 5' end, are also involved in target recognition [34].

Based on a single position analysis of the whole sequences, we have also looked at the purines or pyrimidines tracts in mature miRNAs. We assumed the threshold for purine/pyrimidine rich sequences over 70% for G, A and C,U, respectively. Based on that criterion we found 141 CU-rich and 187 AG-rich sequences. We found: miR-4290, miR-1281, miR-4716-5p, miR-483-3p and miR-877-3p containing almost only UC nucleotides and miR-6124, miR-483-5p, miR-4271, miR-4644, miR-4716-3p and miR-1234-5p purine-rich. In the Table 1 we showed only highly purine/pyrimidine rich sequences, where the level of the investigated nucleotides is over 90%.

The functional analysis of the cellular pathways analysis showed that the purine/pyrimidine-rich sequences show the involvement mostly in lysine degradation, RNA transport and spliceosome (Fig 2a, S1 Table). The search generated a rank-ordered list of KEGG and PANTHER pathways, issuing statistical significance based on p-values. Examining the pathways affected by the distinct group of miRNAs we highlighted pathways associated with different physiological processes connected to the specified group of miRNAs. The heat map indicates the high enrichment of the e.g. miR-483 in RNA degradation, miR-4271 in lysine degradation or miR-4716 in spliceosome (Fig 2b). The list of proteins involved in the pathways and processes associated with the purine/pyrimidine-rich miRNAs also corroborate with these results (S1 Table). The membrane trafficking regulatory, membrane traffic or nucleic acid binding proteins are the most common proteins being a main players in binding and RNA transport processes (S1b Table). Apart from the targeting the distinct group of mRNAs, polypyrimidine-tracts could allow possibly for the interaction of miRNAs with other proteins, such as PTB (polypyrimidine tract-binding protein), preventing the typical regulatory action of miRNAs





**Table 1. Pyrimidines- and purines-rich sequences of miRNAs.** The table shows the content of the pyrimidines or purines nucleotides according to the sequence of miRNAs. The total contents of purines or pyrimidines of the miRNAs presented in the table is above 90%.

miRNA	Sequence	Pyrimidine/purine content [%]	Sequence length
<b>Pyrimidines-rich sequences</b>			
hsa-miR-4290	UGCCCUCCUUUCUUCUCCUC	95	19
hsa-miR-1281	UCGCCUCCUCCUCCUC	94	17
hsa-miR-4716-5p	UCCAUGUUUCCUUCUCCUUCU	91	22
hsa-miR-483-3p	UCACUCCUCCUCCUCCGUCU	90	21
hsa-miR-877-3p	UCCUCUUCUCCUCCUCCAG	90	21
<b>Purines-rich sequences</b>			
hsa-miR-6124	GGGAAAAGGAAGGGGAGGA	100	20
hsa-miR-4271	GGGGGAAGAAAAGGUGGG	95	22
hsa-miR-483-5p	AAGACGGGAGGAAAGAAGGGAG	95	22
hsa-miR-4644	UGGAGAGAGAAAAGAGACAGAAG	91	23
hsa-miR-4716-3p	AAGGGGAAGGAAACAUGGAGA	91	22
hsa-miR-1234-5p	GGGGGGGGGGGGGGGGCCG	90	21

doi:10.1371/journal.pone.0151246.t001

[35]. On the other hand, such miRNAs can interact with Kozak sequence- purine-rich motif, similar to prokaryotic Shine-Dalgarno sequence [36–37]. Recently, the novel mechanism for post-transcriptional control of gene expression was shown [38]. It involves the formation of an intermolecular G-quadruplex between small RNA and mRNA, where a four stranded RNA structure is formed through direct RNA-RNA dimerization. This is hint to suppose, that the G-rich miRNAs could be also regulate the gene expression via this kind of mechanism.

### Simple sequence repeats (SSR) analysis

Regardless the extensive research, there are only limited data about the nucleotide composition, especially in terms of the nucleotide repeats in mature miRNAs. This analysis could be beneficial, since the SSRs can be potentially used for genetic maps construction, as a genetic markers, but also in the linkage association research or phylogenetics and population genetics studies [39–44]. It was also shown, that SSRs could serve for a target genes involved in diseases, for evolution history research or finally, in the execution of the paternity tests [45–48].

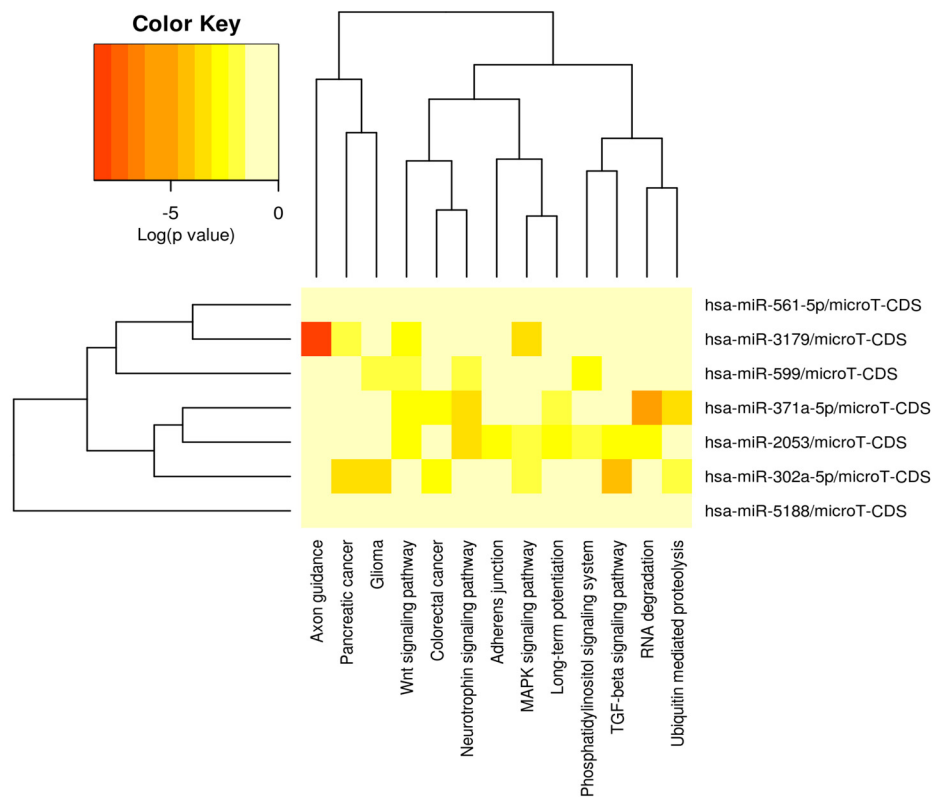
Based on the nucleotide sequence analysis, we performed SSR monitoring. We searched in human miRNAs for the occurrence and nature of SSRs. The most important standard for the identification of SSRs in a given sequence is the definition of the minimum repeat number. The minimum repeat units in track is annotated as the valid SSR tract. In order to detect various repeats in mature miRNAs, and taking into account the length of the molecules, we considered minimum three repeats. Consequently, three minimum repeats were used for the survey of SSRs in pre-miRNA studies [49]. Among the analyzed mature miRNA sequences we were able to define: mono-, di-, tri-, tetra- and pentanucleotide repeats.

Mononucleotide repeats poly (U) and poly (G) were the predominant repeats in all of the analyzed mature miRNAs (Table 2). The longest mononucleotide repeats observed, was the 18 nucleotides SSR poly (G). For poly (U), (A) and (C) repeats we were able to find maximum 10 nucleotides long, both for “U” and “A” SSR and 7 for “C”, respectively. Dinucleotides were the second most common repeats in mature miRNAs (Table 2). We have found that GU/UG were also predominant for human miRNAs (Table 2). To get comprehensive view on specified miRNAs, we established the list GU/UG-rich miRNAs (Table 3). The less frequent, with almost the same relative count, were the AG/GA and AU/UA repeats (Table 2). CG/GC repeats in mature

(a)

Purine/pyrimidine-rich sequences					
miRNAs	KEGG Pathway	Pathway ID	# of Genes (union)	p-value (union)	# miRNAs involved in pathway
hsa-miR-4290	Lysine degradation	hsa00310	15	3.74e-10	5
hsa-miR-1281	RNA transport	hsa03013	17	6.30e-05	3
hsa-miR-4716-5p	Spliceosome	hsa03040	18	9.12e-05	2
hsa-miR-483-3p	PI3K-Akt-signaling pathway	hsa04151	25	0.000253	3
hsa-miR-877-3p	Transcriptional misregulation in cancer	hsa05202	19	0.000362	3
hsa-miR-6124	N-Glycan biosynthesis	hsa00510	2	0.000727	2
hsa-miR-4271	Other types of O-glycan biosynthesis	hsa00514	2	0.005192	2
hsa-miR-483-5p	p53 signaling pathway	hsa04115	12	0.0057423	3
hsa-miR-4644	RNA degradation	hsa03018	12	0.007102	2
hsa-miR-4716-3p	Prostate cancer	hsa05215	14	0.014886	4
hsa-miR-1234-5p					

(b)



**Fig 2. The KEGG pathway analysis for pyrimidine/purine (P/P)-rich miRNAs.** A. The table illustrating the: list of P/P-rich miRNAs (first column); the IDs and KEGG pathways names (second and third column); the number of genes and P/P-rich miRNAs involved associated with the pathways (fourth and sixth column). P-value was given in fifth column as a result of statistical analysis. P-value threshold is considered 0.05. B. P/P-rich miRNAs in predicted pathway heat map. Significant miRNA-pathway interaction  $p < 0.001$ .

doi:10.1371/journal.pone.0151246.g002



**Table 2. Occurrence and relative count of SSRs in mature miRNAs.** Occurrence and relative count of all SSRs in mature miRNAs: mononucleotides, dinucleotides and the list of the longest tetra- and pentanucleotides.

SSRs	Mono	Di	Tri	Tetra	Penta	Total
Total numbers of repeat	2326 (1.14)	324 (0.16)	43 (0.02)	3 (0.001)	4 (0.002)	2700 (1.32)
Repeat type and number	A: 462 (0.23) U: 569 (0.28) G: 802 (0.40) C: 493 (0.24)	AG/GA: 66 (0.03) GU/UG: 83 (0.04) AC/CA: 45 (0.02) CU/UC: 67 (0.03) AU/UA: 47 (0.02) CG/GC: 16 (0.008)	See details in <a href="#">S2 Table</a>	(AGGA) <sub>3</sub>	(UGGGC) <sub>4</sub>	2700 (1.32)

doi:10.1371/journal.pone.0151246.t002

miRNAs are relatively rare ([Table 2](#)). The most frequent dinucleotide repeat we have found within the sequence of has-miR-1277-5p-1, was (AU) motif, repeated 7 times.

Interestingly, most of the trinucleotide repeats contain base “U” and “G”, although it can be noticed that this type of SSR is less represented. AGU, GGU and UGU SSR were not observed in humans mature miRNAs sequences, whereas GUG motif was observed 6 times. The detailed data about trinucleotide repeats are given in [S2 Table](#). Mature human miRNAs with the longest trinucleotide repeat- (AGG)<sub>4</sub> is hsa-miR-4298-1.

Although, in the mature miRNAs tetranucleotide and pentanucleotide repeats were less expected to detect, however we were still able to find such SSRs type ([Table 2](#)). Interesting example of mature miRNA, with pentarepeats, is hsa-miR-3620-5p-1, in which we identified (UGGGC)<sub>4</sub> ([Table 2](#)).

**Table 3. The list of GU-rich miRNAs.** GU-tracts are marked in bold among the sequence.

miRNA	Sequence
miR-574-5p	UGAG <b>GU</b> GUGUGUGUGUGUGAG <b>GU</b> GUGUGU
miR-941	CACCCGGCUG <b>GU</b> UGGCACAUG <b>GU</b> GC
miR-3149	UUU <b>GU</b> AUGGAUUAUG <b>GU</b> GUGUGUAU
miR-1238-5p	<b>GU</b> GAGUGGGAGCCCCAG <b>GU</b> GUGUG
miR-545-3p	UCAGCAAACAUAUUUAU <b>UG</b> UGUGGC
miR-2278	GAGAGCA <b>GU</b> UGUGUGUGUGCCUGG
miR-3148	UGGAAAAAACUG <b>GU</b> GUGUGGCUU
let-7b-5p	UGAG <b>GU</b> AGUAG <b>GU</b> UGUGUGG <b>GU</b> U
miR-493-3p	UGAAG <b>GU</b> CUACUG <b>GU</b> UGGCCAGG
miR-1180	UUUCCGGCUCGC <b>GU</b> GG <b>GU</b> GUGU
miR-539-5p	GGAGAAUUUAUCCUUG <b>GU</b> GUGUGU
miR-32-3p	CAAUUUA <b>GU</b> GUGUGUGUGAUUUU
miR-206	UGGAAUG <b>GU</b> AAGGAAG <b>GU</b> GUGUGG
miR-1299	UUCUGGAAUUC <b>UG</b> UGUGAGGGGA
miR-3911	UG <b>GU</b> UGGAUCCUGGAGGAGGCA
miR-297	AUGUAUG <b>GU</b> UGCAUGUGCAUG
miR-610	UGAGCUAAAUG <b>GU</b> UGGCUGGGGA
miR-1228-5p	<b>GU</b> GGGGGGGGGCAG <b>GU</b> GUGUG
miR-595	GAA <b>GU</b> UGGCCUG <b>GU</b> GUGUGUCU
miR-4455	AGG <b>GU</b> GUGUGUGUGUUUU
miR-3650	AG <b>GU</b> GUGUGUCUGAGAGUCC
miR-147a	<b>GU</b> GUGUGGAAAUUCUUCUGC
miR-660-3p	ACCUCUG <b>GU</b> UGCAUGGAUUA

doi:10.1371/journal.pone.0151246.t003

Until now, there is only a single study focused on SSRs in short sequences, such as pre-miRNAs [49]. However, since miRNAs are functional molecules, we decided to monitor the SSRs in mature miRNAs. For this group of molecules, we have noticed the same tendency that has been already observed for pre-miRNAs. Among the majority of the analyzed miRNAs, poly (A/U) repeats were more frequent than poly (G/C) repeats. Having converted this data into genome level, this observation, both among the previous pre-miRNAs analysis and our mature miRNAs studies, is similar to primate genome [50]. Mononucleotide and dinucleotide repeats were significantly predominant, which was similar to pre-miRNAs and to that of introns, in which majority of SSRs were also mono- and dinucleotides, whereas tri-, tetra-, penta- and hexanucleotide repeats were relatively rare. It has been reported that (GT)<sub>n</sub> is the most predominant dinucleotide repeat motif in animal and invertebrates. Additionally, similarly to pre-miRNAs sequences, the most abundant repeats were (GU)<sub>n</sub>/(UG)<sub>n</sub> as well. Thus, one could assume that SSRs, both in pre-miRNAs and miRNAs, might have a functional meaning. They can be considered for providing a molecular basis for organization of pre-miRNAs *in vivo* or rapid mature miRNAs maturation. SSRs changes within the mature miRNAs sequences, similarly to pre-miRNAs might have a critical impact on the normal miRNA regulation activity and the variations of SSRs in mature miRNAs can influence their direct mRNA target interaction or even alter the function of that distinct miRNA.

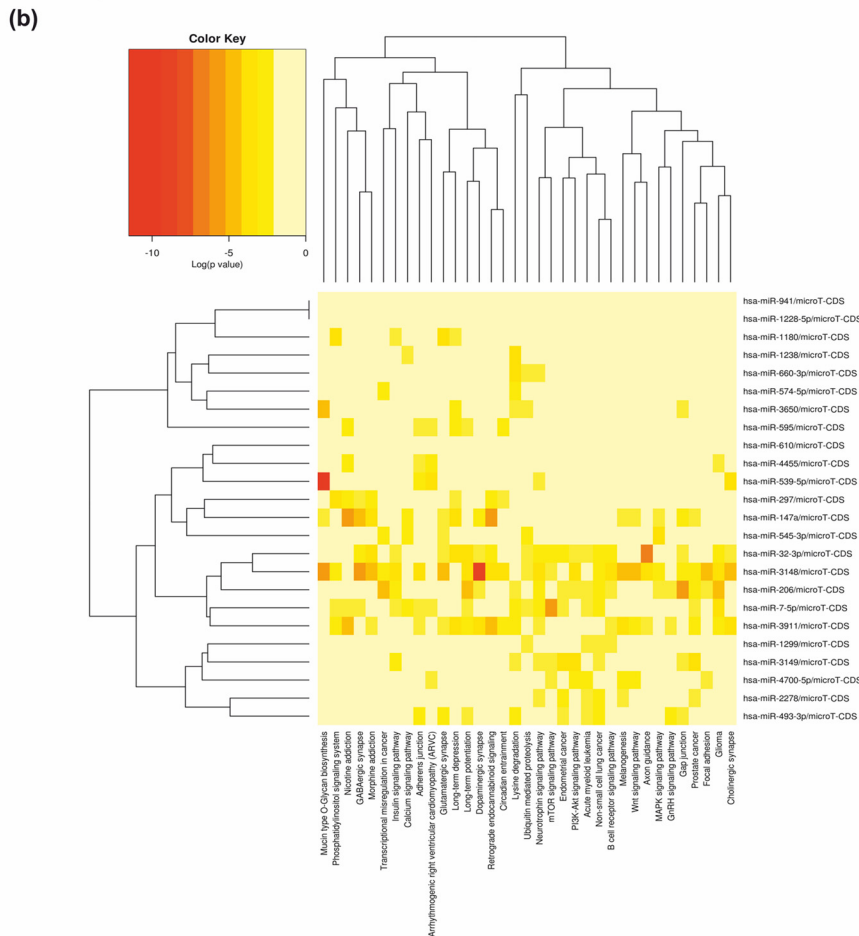
Additionally, our analysis showed that miRNAs enriched with the particular sequence motif are predicted to target different pathways. The pathways associated with GU-rich miRNAs as the most significant indicate a “dopaminergic synapse”, “lysine degradation” and “long-term potentiation”. The more intense red color in a heat map indicates higher probability that a specific pathway is significantly enriched with target genes for a certain miRNA (Fig 3). This analysis also revealed that the GU-rich miRNAs more likely are involved in the neurological processes such as “Opioid proopiomelanocortin pathway” or “Axon guidance mediated by netrin” (Fig 3, S3 Table). The list of the most important proteins involved in the processes with the participation of GU-rich miRNAs covers among others proteins of voltage-gated sodium channel, sodium channel, but also SNARE proteins, which the primary role is to mediate lysosome formation, particularly in the presynaptic membrane in neurons (S3 Table). Although AU-rich elements (AREs) are very abundant in the 3'UTRs of many different mammalian mRNAs with unstable structure, the presence and function of GU-rich elements (GREs) are still poorly understood. There was found that through genome-wide analysis at least 5% of human genes contain GREs in their 3'UTR. The functional over-representation of it is assigned for the genes involved in transcription, nucleic acid metabolism, developmental processes and neurogenesis [51]. Until now, there is no report showing the importance of the GU-rich miRNAs in specific processes. However, in our global sequence analysis of miRNAs, we demonstrate, that this type of miRNAs target also mRNAs involved in biological processes such as mRNA 3'-end processing, mRNA transcription or nervous system development (S3 Table). On the other hand, it was shown that GREs are the targets for at least one RNA-binding protein: CUG-binding protein 1 (CUG-BP1)—the member of the highly conserved CELF family of RNA-binding proteins that are post-transcriptional regulators of deadenylation, mRNA decay, translation and pre-mRNA processing [52–57]. It is possible CUG-BP1 could also bind directly GU-rich miRNAs and affect the protein expression. This needs, however, further investigation.

### The identification of interferon induction motif (IIM)

Going deeper into the sequence analysis within the sequences of mature miRNAs, we identified 5'-UGUGU-3' motif, which is known for stimulation of innate immune response. We have found 50 of mature miRNAs within that motif (Table 4). That feature of the UGUGU sequence

(a)

GU-rich sequences					
miRNAs	KEGG Pathway	Pathway ID	# of Genes (union)	p-value (union)	# miRNAs involved in pathway
miR-4455	Dopaminergic synapse	hsa04728	51	1.17e-08	5
miR-3650	Lysine degradation	hsa00310	19	2.16e-08	9
miR-147a	Long-term potentiation	hsa04720	32	1.84e-07	7
miR-660-3p	Mucin type O-Glycan biosynthesis	hsa00512	14	2.23e-07	4
miR-1228-5p					
miR-610	Retrograde endocannabinoid signaling	hsa04723	38	5.74e-07	5
miR-297					
miR-3911	Long-term depression	hsa04730	21	6.71e-07	7
miR-1299					
miR-4700-5p	Prostate cancer	hsa05215	41	1.51e-06	7
miR-206	Gap junction	hsa04540	36	4.79e-06	7
miR-32-3p	Glioma	hsa05214	31	5.87e-06	6
miR-539-5p					
miR-1180	Glutamate synaptic	hsa04724	39	7.07e-06	7
miR-493-3p	Nicotine addiction	hsa05033	12	1.00e-05	6
miR-7b-5p					
miR-3148	Neurotrophin signaling pathway	hsa04722	55	1.15e-05	9
miR-2278					
miR-545-3p	GABAergic synapse	hsa04727	30	0.00023	5
miR-1238-5p	Cholinergic synapse	hsa04725	40	0.00041	4
miR-3149					
miR-941	Endometrial cancer	hsa05213	27	0.00051	6
miR-574-5p					



**Fig 3. The KEGG pathway analysis for GU-rich miRNAs.** A. The table illustrating the: list of GU-rich miRNAs (first column); the IDs and KEGG pathways names (second and third column); the number of genes and GU-rich miRNAs involved associated with the pathways (fourth and sixth column). P-value was given in fifth column as a result of statistical analysis. P-value threshold is considered 0.05. B. GU-rich miRNAs in predicted pathway heat map. Significant miRNA-pathway interaction  $p < 0.001$ .

doi:10.1371/journal.pone.0151246.g003

**Table 4. Selected human miRNAs contained immunostimulatory motifs.** The table presents the miRNA sequences with at least one UGUGU motif (highlighted in bold).

miRNA	Sequence (5'-3')	miRNA	Sequence (5'-3')
miR-3682-5p	CUACUUCUAC <b>CGUGU</b> UAUCAU	miR-539-5p	GGAGAAAUUAUCCU <b>UGUGU</b>
miR-2278	GAGAGCAG <b>UGUGU</b> GUUGCCUGG	miR-1180-3p	UUUCCGGCUCGCG <b>UGUGU</b>
miR-1238-5p	GUGAGUGGGAGCCCCAG <b>UGUGU</b>	miR-450a-5p	UUUUGCGA <b>UGUGU</b> UCCUAUAU
miR-3149	UUGUAUGGAUA <b>UGUGU</b> GUGUAU	miR-493-3p	UGAAGGUCUAC <b>UGUGU</b> GCCAGG
let-7b-5p	UGAGGUAGUAGGU <b>UGUGU</b> GGUU	miR-4669	<b>UGUGU</b> CCGGGAAGUGGAGGAGG
miR-4455	AGGG <b>UGUGU</b> GUGUUUUU	miR-4753-3p	UUCUCUUUCUUUAGCCU <b>UGUGU</b>
miR-4717-5p	UAGGCCACAGCCACCCA <b>UGUGU</b>	miR-624-5p	UAGUACCAGUACCU <b>UGUGU</b> UCA
miR-573	CUGAAGUGA <b>UGUGU</b> AACUGAUCAG	miR-3657	<b>UGUGU</b> CCCAUUAUUGGUGAUU
miR-660-3p	ACCUC <b>UGUGU</b> GCAUGGAUUA	miR-3148	UGGAAAAACUGG <b>UGUGU</b> GCUU
miR-378a-5p	CUCUGACUCCAGGUCC <b>UGUGU</b>	miR-3177-5p	<b>UGUGU</b> ACACACGUGCCAGGCGCU
miR-599	GU <b>UGUGU</b> CAGUUUAUCAAC	miR-581	UCU <b>UGUGU</b> UCUCUAGAUCAU
miR-1226-3p	UCACCAGCCC <b>UGUGU</b> UCCCUAG	miR-642b-5p	GGUCCCUCCUCCAAA <b>UGUGU</b> CU
miR-1270	CUGGAGAUUUGGAAGAGC <b>UGUGU</b>	miR-147a	<b>UGUGU</b> GGAAAUGCUUCUGC
miR-610	UGAGCUAAA <b>UGUGU</b> GCUGGGA	miR-592	<b>UGUGU</b> CAAUAUGCGAUGAUGU
miR-3911	<b>UGUGU</b> GGAUCCUGGAGGAGGCA	miR-545-3p	UCAGCAAACAUAU <b>UGUGU</b> GC
miR-3650	AGG <b>UGUGU</b> CUGUAGAGUCC	miR-1228-5p	GUGGGCGGGGCGAG <b>UGUGU</b> G
miR-4789-5p	GUAUACACCUGAUUA <b>UGUGU</b> AUG	iR-642a-5p	GUCCUCUCCAAA <b>UGUGU</b> CUUG
miR-330-5p	UCUCUGGGCC <b>UGUGU</b> CUUAGGC	miR-3152-3p	<b>UGUGU</b> UAGAAUAGGGGCAAUAA
miR-941	CACCCGGC <b>UGUGU</b> GCACAUGUGC	miR-32-3p	CAAUUUAG <b>UGUGU</b> GUGAUUUU
miR-597	<b>UGUGU</b> CACUCGAUGACCACUGU	miR-1299	UUCUGGAAU <b>UGUGU</b> GAGGGA
miR-5580-5p	UGCUGGCUAUUUCAUA <b>UGUGU</b>	miR-297	AUGUA <b>UGUGU</b> GCAUGUGCAUG
miR-4700-5p	UCUGGGGAUGAGGACAG <b>UGUGU</b>	miR-187-3p	UCGUGUCU <b>UGUGU</b> UGCAGCCGG
miR-892a	CAC <b>UGUGU</b> CCUUUCUGCGUAG	miR-5010-3p	UUU <b>UGUGU</b> CUCCCAUCCCCAG
miR-206	UGGAAUGUAAGGAAG <b>UGUGU</b> GG	miR-574-5p	UGAG <b>UGUGU</b> UGUGUGUGAG <b>UGUGU</b>
miR-595	GAAGUGUGCCUGG <b>UGUGU</b> C	miR-649	AAACC <b>UGUGU</b> UGUUCAAGAGUC

doi:10.1371/journal.pone.0151246.t004

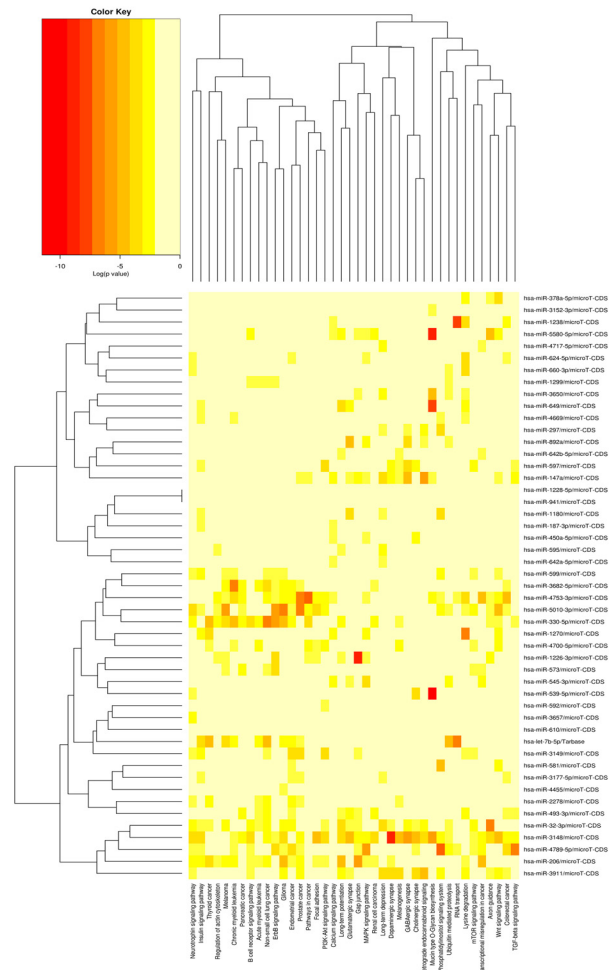
was first shown for siRNAs as extremely important, as it may cause variety of non-specific side effects, including stimulation of interferon and cytokine production, global shutdown of protein synthesis or nonspecific degradation of mRNAs [58–59]. The miRNA, containing interferon stimulation motifs, can induce itself one or more undesirable effects, such as proliferation blockage, differentiation or apoptosis of cancer cells, or it can even serve as potent immunomodulatory agent [58–59].

We have found, that apart from the potential of these group of miRNAs to possible direct activation of interferon response, they also target particular mRNAs involved in specific pathways and cellular processes. Analysis for UGUGU-rich miRNAs ranked at the top of the list following pathways: pathways in cancer, PI3K-Akt signaling, MAP-signaling and hedgehog signaling pathways (Fig 4, S4 Table). Our analysis showed that these type of miRNAs is involved among others in processes such as induction of apoptosis, immune system process, immune response or macrophage activation (S4 Table). Due to the KEGG and PANTHER analyses of miRNAs involvement in this type of biological processes, it appears, that also for IIM-rich miRNAs we could expect the direct recognition of the mRNAs involved in PI3K-Akt signaling and MAP signaling pathways (Fig 4a, S4 Table). The potential involvement of IIM-contained miRNAs in these pathways are also supported by the list of protein class provided by the PANTHER analysis (S4 Table). It was shown that positioning of this 5'-UGUGU-3' motif especially at the 5'- end of the sense strand of siRNAs results in a rapid and enhanced induction

(a)

UGUGU-rich sequences					
miRNAs	KEGG Pathway	Pathway ID	# of Genes (union)	p-value (union)	# miRNAs involved in pathway
miR-3682-5p, miR-2278, miR-1238-5p, miR-3149, let-7b-5p, miR-4455, miR-4717-5p, miR-573, miR-660-3p, miR-378a-5p, miR-599, miR-1226-3p, miR-1270, miR-610, miR-3911, miR-3650, miR-4789-5p, miR-330-5p, miR-941, miR-597, miR-5580-5p, miR-4700-5p, miR-892a, miR-206, miR-595, miR-539-5p, miR-1180-3p, miR-450a-5p, miR-493-3p, miR-4669, miR-4753-3p, miR-624-5p, miR-3657, miR-3148, miR-3177-5p, miR-581, miR-642b-5p, miR-147a, miR-592, miR-545-3p, miR-1228-5p, miR-642a-5p, miR-3152-3p, miR-32-3p, miR-1299, miR-297, miR-187-3p, miR-5010-3p, miR-574-5p, miR-649	Pathways in cancer	hsa05200	186	2.79e-27	43
	PI3K-Akt signaling pathway	hsa04151	183	3.43e-25	43
miR-3682-5p, miR-2278, miR-1238-5p, miR-3149, let-7b-5p, miR-4455, miR-4717-5p, miR-573, miR-660-3p, miR-378a-5p, miR-599, miR-1226-3p, miR-1270, miR-610, miR-3911, miR-3650, miR-4789-5p, miR-330-5p, miR-941, miR-597, miR-5580-5p, miR-4700-5p, miR-892a, miR-206, miR-595, miR-539-5p, miR-1180-3p, miR-450a-5p, miR-493-3p, miR-4669, miR-4753-3p, miR-624-5p, miR-3657, miR-3148, miR-3177-5p, miR-581, miR-642b-5p, miR-147a, miR-592, miR-545-3p, miR-1228-5p, miR-642a-5p, miR-3152-3p, miR-32-3p, miR-1299, miR-297, miR-187-3p, miR-5010-3p, miR-574-5p, miR-649	MAP signaling pathway	hsa04010	138	1.06e-19	44
	HTLV-I infection	hsa05166	137	1.25e-19	40
miR-3682-5p, miR-2278, miR-1238-5p, miR-3149, let-7b-5p, miR-4455, miR-4717-5p, miR-573, miR-660-3p, miR-378a-5p, miR-599, miR-1226-3p, miR-1270, miR-610, miR-3911, miR-3650, miR-4789-5p, miR-330-5p, miR-941, miR-597, miR-5580-5p, miR-4700-5p, miR-892a, miR-206, miR-595, miR-539-5p, miR-1180-3p, miR-450a-5p, miR-493-3p, miR-4669, miR-4753-3p, miR-624-5p, miR-3657, miR-3148, miR-3177-5p, miR-581, miR-642b-5p, miR-147a, miR-592, miR-545-3p, miR-1228-5p, miR-642a-5p, miR-3152-3p, miR-32-3p, miR-1299, miR-297, miR-187-3p, miR-5010-3p, miR-574-5p, miR-649	Regulation of actin cytoskeleton	hsa04810	124	1.77e-18	42
	Focal adhesion	hsa04510	123	2.01e-17	43
miR-3682-5p, miR-2278, miR-1238-5p, miR-3149, let-7b-5p, miR-4455, miR-4717-5p, miR-573, miR-660-3p, miR-378a-5p, miR-599, miR-1226-3p, miR-1270, miR-610, miR-3911, miR-3650, miR-4789-5p, miR-330-5p, miR-941, miR-597, miR-5580-5p, miR-4700-5p, miR-892a, miR-206, miR-595, miR-539-5p, miR-1180-3p, miR-450a-5p, miR-493-3p, miR-4669, miR-4753-3p, miR-624-5p, miR-3657, miR-3148, miR-3177-5p, miR-581, miR-642b-5p, miR-147a, miR-592, miR-545-3p, miR-1228-5p, miR-642a-5p, miR-3152-3p, miR-32-3p, miR-1299, miR-297, miR-187-3p, miR-5010-3p, miR-574-5p, miR-649	Endocytosis	hsa04144	107	5.99e-17	41
	Wnt signaling pathway	hsa04310	99	2.31e-15	40
miR-3682-5p, miR-2278, miR-1238-5p, miR-3149, let-7b-5p, miR-4455, miR-4717-5p, miR-573, miR-660-3p, miR-378a-5p, miR-599, miR-1226-3p, miR-1270, miR-610, miR-3911, miR-3650, miR-4789-5p, miR-330-5p, miR-941, miR-597, miR-5580-5p, miR-4700-5p, miR-892a, miR-206, miR-595, miR-539-5p, miR-1180-3p, miR-450a-5p, miR-493-3p, miR-4669, miR-4753-3p, miR-624-5p, miR-3657, miR-3148, miR-3177-5p, miR-581, miR-642b-5p, miR-147a, miR-592, miR-545-3p, miR-1228-5p, miR-642a-5p, miR-3152-3p, miR-32-3p, miR-1299, miR-297, miR-187-3p, miR-5010-3p, miR-574-5p, miR-649	Chemokine signaling pathway	hsa04062	100	1.94e-14	39
	Axon guidance	hsa04360	80	1.43e-13	40
miR-3682-5p, miR-2278, miR-1238-5p, miR-3149, let-7b-5p, miR-4455, miR-4717-5p, miR-573, miR-660-3p, miR-378a-5p, miR-599, miR-1226-3p, miR-1270, miR-610, miR-3911, miR-3650, miR-4789-5p, miR-330-5p, miR-941, miR-597, miR-5580-5p, miR-4700-5p, miR-892a, miR-206, miR-595, miR-539-5p, miR-1180-3p, miR-450a-5p, miR-493-3p, miR-4669, miR-4753-3p, miR-624-5p, miR-3657, miR-3148, miR-3177-5p, miR-581, miR-642b-5p, miR-147a, miR-592, miR-545-3p, miR-1228-5p, miR-642a-5p, miR-3152-3p, miR-32-3p, miR-1299, miR-297, miR-187-3p, miR-5010-3p, miR-574-5p, miR-649	RNA transport	hsa03013	89	4.09e-13	41
	Ubiquitin mediated proteolysis	hsa04120	85	1.19e-12	40
miR-3682-5p, miR-2278, miR-1238-5p, miR-3149, let-7b-5p, miR-4455, miR-4717-5p, miR-573, miR-660-3p, miR-378a-5p, miR-599, miR-1226-3p, miR-1270, miR-610, miR-3911, miR-3650, miR-4789-5p, miR-330-5p, miR-941, miR-597, miR-5580-5p, miR-4700-5p, miR-892a, miR-206, miR-595, miR-539-5p, miR-1180-3p, miR-450a-5p, miR-493-3p, miR-4669, miR-4753-3p, miR-624-5p, miR-3657, miR-3148, miR-3177-5p, miR-581, miR-642b-5p, miR-147a, miR-592, miR-545-3p, miR-1228-5p, miR-642a-5p, miR-3152-3p, miR-32-3p, miR-1299, miR-297, miR-187-3p, miR-5010-3p, miR-574-5p, miR-649	Dopaminergic synapse	hsa04728	82	1.80e-12	43
	Glutamatergic synapse	hsa04724	72	1.86e-12	43
miR-3682-5p, miR-2278, miR-1238-5p, miR-3149, let-7b-5p, miR-4455, miR-4717-5p, miR-573, miR-660-3p, miR-378a-5p, miR-599, miR-1226-3p, miR-1270, miR-610, miR-3911, miR-3650, miR-4789-5p, miR-330-5p, miR-941, miR-597, miR-5580-5p, miR-4700-5p, miR-892a, miR-206, miR-595, miR-539-5p, miR-1180-3p, miR-450a-5p, miR-493-3p, miR-4669, miR-4753-3p, miR-624-5p, miR-3657, miR-3148, miR-3177-5p, miR-581, miR-642b-5p, miR-147a, miR-592, miR-545-3p, miR-1228-5p, miR-642a-5p, miR-3152-3p, miR-32-3p, miR-1299, miR-297, miR-187-3p, miR-5010-3p, miR-574-5p, miR-649	Calcium signaling pathway	hsa04020	93	1.86e-12	46

(b)



**Fig 4. The KEGG pathway analysis for interferon induction motif contained-(IIM) miRNAs.** A. The table illustrating the: list of IIM-contained miRNAs (first column); the IDs and KEGG pathways names (second and third column); the number of genes and IIM-rich miRNAs involved associated with the pathways (fourth and sixth column). P-value was given in fifth column as a result of statistical analysis. P-value threshold is considered 0.05. B. IIM-rich miRNAs in predicted pathway heat map. Significant miRNA-pathway interaction  $p < 0.001$ .

doi:10.1371/journal.pone.0151246.g004

of type I IFN. Rapid production of IFN  $\beta$  involves thus activation of signaling cascades governed by effectors that are intermediates in the JAK/STAT, mitogen-activated kinase (MAPK) and phosphatidylinositol 3-kinase (PI3K) pathways [60–61].

This results indicates that IIM-rich miRNAs, similarly to the cancer-associated miRNAs (CA-miRNAs) are closely connected to cancer [62]. The one of the most prominently observed pathway for these class of IIM- miRNAs- pathways in cancer is highly represented by the gene targets for IIM-miRNAs. This kind of pathway clearly demonstrated the involvement into biological capabilities such as evading of apoptosis, block of differentiation, unlimited replication, increased angiogenesis, sustained ability for invasion and metastasis in malignant transformation [62]. Mitogen –activated protein kinase (MAPK) pathway functions also as integrating signals that affect proliferation, differentiation, survival and migration. The results indicates, that the IIM-rich miRNAs could also be a promising clinical targets for cancer through the MAPK pathway. The phosphatidylinositol 3-kinase(PI3K)-Akt signaling pathway is activated by many types of cellular stimuli or toxic agents and it is involved in regulation of basic physiological cellular functions such as transcription, translation, proliferation, growth or survival [63]. There was shown that serine/threonine kinase Akt/PKB plays significant role in this pathway. Moreover, a impaired activation of the PI3K-Akt pathway has been associated with the development of different types of diseases such as diabetes, mellitus, autoimmunity, and finally- cancer [64].

### The characteristic of stem-loop hairpin structure of mature human miRNAs

It is already known that over 50% of the hairpins predicted for the human 16S and 23S rRNA have tetranucleotides loops [65]. The tetraloops are thought to fulfill a variety of functions, including recognition elements for interactions with proteins and other RNAs. They can regulate the activity of a biological system by shifting the equilibrium between alternate structures [66]. It was observed, with the use of the NMR methods that small stem-loops can exist, in solution, in equilibrium with duplex forms [67]. About 70% of these tetraloops have the consensus loop sequences (GNRA) or UNCG (where N = A,C,G or U; and R = A or G. It has been shown that RNA hairpins within these sequences form unusually stable loop conformation [65–67].

**Table 5. Occurrence of tetranucleotides motifs connected with hairpin formation in human miRNA sequences.** The all miRNAs with the ability to form secondary structure were subjected to the analysis. The searching was performed according to: first—motifs present in loop of the predicted hairpins, second: motifs present in the loop where at least 3 base pair were predicted in the stem and finally-where at least 3 base pair and additionally C-G as a closing pair in the stem were predicted.

Motif	Frequency		
	Sequences		
	Hairpin-formation	3 base-pairs in stem	3 base-pairs and C-G as a closing pair in stem
GGAA	25	21	8
CUUG	10	8	4
GCAA	13	10	2
GUGA	23	17	0
GAAA	19	17	3
GAGA	32	18	1
UUCG	5	5	2
UUUU	28	21	12
UUUG	24	22	6
<b>Total</b>	<b>179</b>	<b>139</b>	<b>38</b>

doi:10.1371/journal.pone.0151246.t005



With Mfold algorithm, we found 1431 (70%) sequences of human miRNAs, able to fold into the hairpin structure, whereas 588 sequences did not show such a propensity. The minimal free energy of these structures ( $\Delta G$ ) falls in the range from -0.1 to -11.1 kcal/mol, however, the most widely represented structures indicate the energy levels from -0.4 to -3.3 kcal/mol [68]. In order to gain better insight into the potential secondary structure of mature miRNAs, we have also analyzed a possibility of the hairpin stem-loop formation. Our calculation revealed that the most frequently motif occurred in miRNAs, is the four nucleotides motif (tetraloop), found in 433 sequences of mature RNAs. Other hairpin-loop of miRNA sequences are: 413, 283 and 114 for three-, five- and six-nucleotides loops, respectively.

Among tetraloops, we specified sequences: UUCG, GAAA, GCAA, GAGA, GUGA, GGAA, CUUG, UUUG—involved in the hairpin loop formation. Many sequences of mature miRNAs can fold exactly into the hairpin loop, with the nucleotides of interest located in the loop, containing minimum 3 base pairs in the stem additionally (Table 5).

We found that 179 miRNAs can form stable hairpin structure, with the 4-nt motifs (Table 5). The most represented motif in human miRNA sequences is GAGA, although GGAA, GAAA or UUUU motifs are also abundant. UUUU motif in the loop, together with the CG as a closing pair in the stem, is most widely represented among the mature human miRNAs (Table 5).

The observation that miRNA can fold into the secondary structure is consistent also with previous reports, that plant miRNA can form secondary structure with free energy of: -9.3 to +1.5 kcal/mol [69]. It directly supports the observation that given miRNAs differ dramatically in terms of the half-life. Thus, the knotty secondary structures of miRNAs makes them nucleases resistant, implicating their long surveillance in the cell. The higher order structure of miRNA can play a crucial role in conformational changes during miRNA-mRNA interactions. This could modulate the pairing and also explain the different degree of genetic regulation for the specific miRNA. The secondary structure could be also extremely important in the mechanism by which sequences for some miRNAs are selected, what can modulate its affinity with their mRNA targets. It can also provide the specificity of miRNAs-mRNA interaction what could be achieved by the presence of miRNA secondary structure, by precluding the possibility of binding of other miRNAs and genomic RNAs with complementary sequences.

Although the mature miRNAs are generally considered as single stranded, there are very few reports suggesting a self-complementarity in mature miRNAs, that has already been observed in more than 50% of mature miRNAs. They are prone to form hairpins and/or homo-duplexes in solution [38, 57–58]. NMR studies have shown that hsa-miR-520h mature strand can fold into hairpin structure or into self-complementary homo-duplex, in higher concentration [70].




Next, we used the ModeRNA program for modelling of miRNA 3D structure, based on templates of related molecule. To find tertiary structures, we searched the PDB database for experimentally confirmed (X-ray, NMR or CD), about 24 nucleotides long RNAs. We have found 168 sequences, which were aligned to human mature miRNA sequences. For further analysis, we only took into account sequences with at least 80% similarity, since only that level provides reliable, homology modeling. The alignment showed 35 miRNAs that were further processed with ModeRNA server (Fig 5a). Positions with identical nucleotides were fixed, whereas remaining positions were modeled by the program. The best matched results we obtained include miR-381-3p with 1R4H RNA, miR-4649-5p with 2KYE and miR-3677-5p with 1Q8N RNA (Fig 5b). Due to the homology modeling, ModeRNA analysis makes our calculation and secondary structure prediction more accurate and give strong support for miRNAs hairpin loop formation hypothesis.

Our CD, NMR and enzymatic probing data for some miRNAs also prove that miRNAs have the intrinsic potential to form secondary structure and that hairpin possibly is a prevailing form of miRNA in the cell [68].

(a)

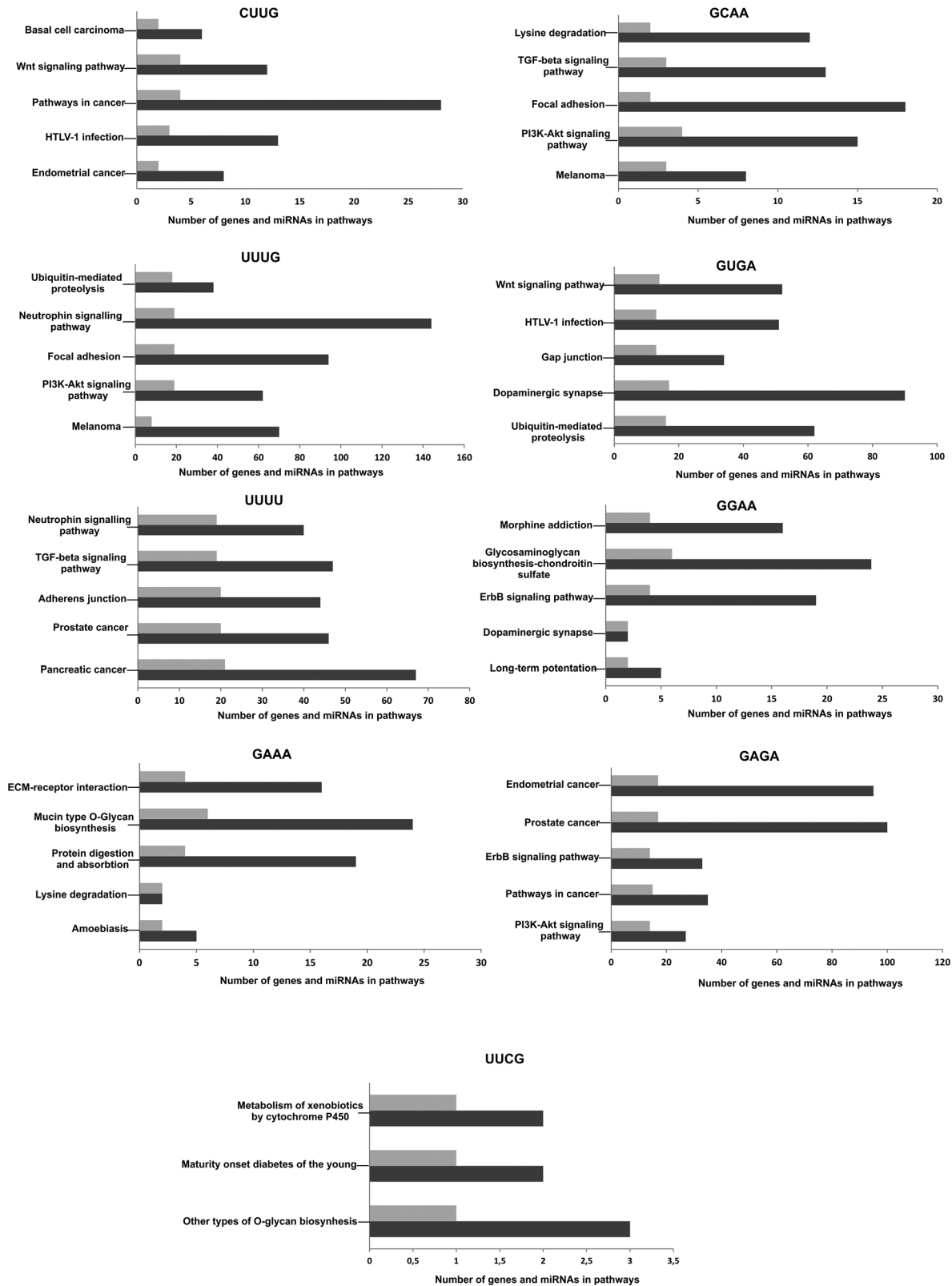
miRNA/RNA structure (PDB)	Alignment	miRNA/RNA structure (PDB)	Alignment	miRNA/RNA structure (PDB)	Alignment
hsa-miR-4649-5p 2KYE:A	UGGGCGAGGGUGGGCUCUCAGAG -----GAGAGUUGGGCUCUC----- *** * *****	hsa-miR-4725-3p 2K65:A	UGGGGAAGGGCUCAGUCUCGGG -----CAGUCUC----- *****	hsa-miR-4656 1R4H:A	UGGGCUGAGGGCAGGAGCCUGU -----GGGCA--AGCCC-- ***** ** *
hsa-miR-34b-5p 2K65:B	UAGGCAGUGUCAUUAGCUGAUUG -----CAGUCUC----- *****	hsa-miR-4720-3p 2OJ7:A	UGC UUAAGUUGUACCAAGUAU -GCU--GUUGU----- *** *****	hsa-miR-4799-3p 2K65:A	ACUGGCAUGCUGCAUUUAUAUA ---GGCA--CUG----- **** **
hsa-miR-381-3p 1R4H:A	UAUACAAGGGCAAGCUCUCUGU -----GGGCAAGCCC----- ***** *	hsa-miR-4654 2K65:A	UGUGGGAUCUGGAGGCAUCUGG -----GGCA-CUG-- *****	hsa-miR-93-5p 2OJ7:A	CAAAGUCGUGUUCUGGUCAGGUAG -----GCUGUU-GU----- ***** **
hsa-miR-648 2K65:A	AAGUGUCAGGGCAGCUGGU -----GGCAGCAG-- *****	hsa-miR-635 2K65:A	ACUUGGGCACUGAAACAUGUCC -----GGCAGUC----- *****	hsa-miR-216a-5p 2K65:A	UAAUCUCAGCUGGGCAACUGUGA -----GGCA-CUG-- **** **
hsa-miR-3677-5p 1R4H:A	CAGUGGCCAGAGCCUCUGCAGUG -----GGGCA-AGCCC----- ** * *****	hsa-miR-5195-5p 2K65:A	AACCCCUAAGGCAACUGGAUGG -----GGCA-CUG-- *****	hsa-miR-3064-5p 2OJ7:A	UCUGGCUGUUGUGGUGGCAA ----GCUGUUGU----- *****
hsa-miR-27a-5p 2OJ7:A	AGGGCUUAGCUGCUUGUGAGCA -----GCUG-UUGU----- **** **	hsa-miR-597-5p 1KOS:A	UGUGUCACUCGAUGACCACUGU CUGUGU---UCGAU--CCACAG-- ***** ** **	hsa-miR-365b-5p 2K65:A	AGGGACUUUCAGGGGCAGCUGU -----GGCA-CUG-- **** **
hsa-miR-4726-5p 1ZIG:A	AGGGCCAGAGGAGCCUGGAGUGG -GGGCAG--AGCCU----- **** ** *****	hsa-miR-33b-3p 1Q75:A	CAGUGCCUCGGCAGUCAGCCC -----GGCUCU-CAGU-AGCC-- * ** * *****	hsa-miR-183-5p 2K65:A	UAUGGCACUGGUAUUAUCACU ---GGCAGUC----- *****
hsa-miR-550b-3p 2K65:A	UCUUAUCUCCUCAGGCACUG -----GGCAGUC----- *****	hsa-miR-4447 2OJ7:A	GGUGGGGUCUGUUGUUU -----GCUUUGU-- *****	hsa-miR-135b-3p 1R4H:A	AUGUAGGGCUAAAAGCCAUAGG ----GGGC---AAGCCC-- **** *****
hsa-miR-4260 1R4H:A	CUUGGGCAUGGAGUCCCA -----GGGCA--AGCCC-- **** ** *	hsa-miR-199a-3p 2K65:B	ACAGUAGUCGACAUUGGUUA -CAGU-GUC----- **** **	hsa-miR-5095 1AFX:A	UUAAGGCGUGAACACCAGCG ----GGUGAGAC-ACC-- ** *****
hsa-miR-23b-5p 2K65:B	UGGGUUCUGGCAUGCUGAUUU -----GGCA-CUG-- **** **	hsa-miR-34a-5p 2K65:B	UGGCAGUGUCUAGCUGGUUGU ---CAGUCUC----- *****	hsa-miR-296-3p 2KYE:A	GAGGUUGGUGGAGGUCUCUC GAGAGUUGG-----CUCUC-- *** *****
hsa-miR-3670 2K65:A	AGAGCUCACAGCUGCCUUCUCUA -----CAG-UGUC----- *** **	hsa-miR-34a-5p 2OJ7:A	UGGCAGUGUCUAGCUGGUUGU -----GCUUUGU-- **** **	hsa-miR-3177-3p 2K65:A	UGCACGGCACUGGGGACACGU ----GGCAGUC----- *****
hsa-miR-4487 2F87:A	AGAGCUGGCGAAGGGCAG -----GGCUGAAGGGCC-- *****	hsa-miR-6131 1ATW:A	GGCUGGUCAGUAGGAGUG -GCU--CCAGAUG-AGCG *** ***** ** *		

(b)

miRNA	miR-3677-5p	miR-381-3p	miR-4649-5p
<b>PRIMARY</b>	CAGUGGCCAGAGCCUCGAGUG	UAUACAAGGGCAAGCUCUCUGU	UGGGCGAGGGUGGGCUCUCAGAG
<b>SECONDARY</b>	<pre>       A G       C C       G C C       U G C C       A G C       C U C       5' C-GCAGUG 3'           </pre>	<pre>       A A       C C G       G U       G C       A U C       A U C       C G       5' UAUU-U 3'           </pre>	<pre>       U G G       G C G       G C       G U       G C       A U       5' UGGGCG-CAGAG 3'           </pre>
<b>TERTIARY</b>			

**Fig 5. Tertiary structure of miRNAs as a result of ModeRNA modeling.** The homological comparison of miRNAs to the known, experimentally confirmed short RNA sequences was performed in the three main steps: sequence analysis (primary structure), secondary structure folding and tertiary structure modeling. A. Sequence alignment of miRNAs to the experimentally confirmed short RNAs. B. Three selected miRNAs modeled in ModeRNA program with high similarity to the short RNA, which structures are deposited in PDB.

doi:10.1371/journal.pone.0151246.g005



**Fig 6. The major class of potential pathways associated with the miRNAs with the stable secondary structure and defined tetraloop.** The number of genes connected to the pathways were marked with dark grey, with light grey- the number of miRNAs.

doi:10.1371/journal.pone.0151246.g006

## The association of structured miRNAs and the cellular pathways

The pathway analysis revealed that miRNAs contained different sequences in the tetraloop also can cluster into specific group. This clustering approach revealed that miRNAs enriched in CUUG or GUGA sequence in the tetraloop are a regulators of the targets from Wnt signaling pathways. Neutrophin signaling pathway appears to be more associated with miRNA contained UUUG or UUUU motifs, whereas dopaminergic synapse pathways are more targeted by miRNAs with GUGA and GGAA motifs and PI3K-Akt signaling pathway by UUUG and GAGA-enriched miRNAs, respectively (Fig 6, S5 Table).

## Conclusions

For a very long time, RNA was considered to be exclusively the carrier of genetic information, but now, the group of RNAs with regulatory function have evolved into the miRNAs, long-noncoding RNAs (lncRNAs), circular RNAs or miRNA-sponges [71].

Although, one miRNA can potentially regulate hundreds of different mRNAs, the majority of transcripts are still actively expressed and translated, which supports the existence of other mechanisms, that counteract miRNA regulation, to achieve homeostasis.

Expression of the gene families or several components of a particular signaling pathway are frequently regulated by miRNAs. The pathway analysis of potential targets based on the sequence, thus the structure of miRNAs could then enhance the probability of identification and verification of the relevant miRNA-target interactions.

Our observation strongly suggest that miRNA persistence is related to biological function, thus better characterization of miRNA structure, stability and associated regulatory mechanism should provide new avenues for the characterization of their biological function. We postulate that, due to the specific sequence features, these molecules can also be involved in very well defined cellular processes depending on their sequence contents. Moreover, the unique features encoded in the sequence and in the structure of mature miRNA can be a key to understand the mRNA target recognition.

## Supporting Information

### S1 Fig. The length heterogeneity of human mature miRNAs.

(DOC)

### S1 File. ("mirna\_an.py"). Python scripts used for miRNA nucleotide sequences and SSR analysis.

(PY)

**S1 Table. The results from the analysis pyrimidines/purines-rich miRNAs using PANTHER classification system.** The table presents: top 10 biological processes related to input miRNAs; most significant pathways derived from overrepresentation test **and** top 10 protein classes related to input miRNAs. +/- shows over— or underrepresentation. Second and third columns contain the number of genes in reference and input list, respectively. P-value threshold is considered 0.05.

(DOC)

### S2 Table. Occurrence and relative count of trinucleotide repeats in mature miRNAs.

(DOC)

**S3 Table. The results from the analysis of GU-rich miRNAs using PANTHER classification system.** The table presents: top 10 biological processes related to GU-rich miRNAs; most significant pathways derived from overrepresentation test and top 10 protein classes related to

GU-rich miRNAs. +/- shows over—or underrepresentation. Second and third columns contain the number of genes in reference and input list, respectively. P-value threshold is considered 0.05.

(DOC)

**S4 Table. The results from the analysis of interferon induction motif (IIM)- contained miRNAs using PANTHER classification system.** The table presents: top 10 biological processes related to IIM motif-contained miRNAs; most significant pathways derived from overrepresentation test and top 10 protein classes related to IIM motif-contained miRNAs. +/- shows over—or underrepresentation. Second and third columns contain the number of genes in reference and input list, respectively. P-value threshold is considered 0.05.

(DOC)

**S5 Table. Molecular KEGG pathways analysis for miRNAs with defined tetraloops in the secondary structure.**

(DOC)

## Acknowledgments

We would like to thank to dr Patrick P. Perrigue for a critical reading of the manuscript.

This work was supported by the European Regional Development Fund within the Innovative Economy Programme, action 1.3.1. (POIG.01.03.01-30-050/09), National Science Centre, grant 5955/B/P01/2010/38 and Ministry of Science and Higher Education of the Republic of Poland by KNOW program.

## Author Contributions

Conceived and designed the experiments: KR MP AB DW JJ MZB JB. Performed the experiments: KR MP AB DW JJ. Analyzed the data: KR MP AB DW JJ MZB JB. Contributed reagents/materials/analysis tools: KR MP AB DW JJ MZB JB. Wrote the paper: KR MP AB JB.

## References

1. Wan Y, Kertesz M, Spitale RC, Segal E and Chang HY (2011) Understanding the transcriptome through RNA structure. *Nat Rev Genet* 12: 641–655. doi: [10.1038/nrg3049](https://doi.org/10.1038/nrg3049) PMID: [21850044](https://pubmed.ncbi.nlm.nih.gov/21850044/)
2. Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, et al. (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425: 415–419. PMID: [14508493](https://pubmed.ncbi.nlm.nih.gov/14508493/)
3. Ambros V (2004) The functions of animal microRNAs. *Nature* 431: 350–355. PMID: [15372042](https://pubmed.ncbi.nlm.nih.gov/15372042/)
4. Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116: 281–297. PMID: [14744438](https://pubmed.ncbi.nlm.nih.gov/14744438/)
5. Chen CZ, Li L, Lodish HF and Bartel DP (2004) MicroRNAs modulate hematopoietic lineage differentiation. *Science* 303: 83–86. PMID: [14657504](https://pubmed.ncbi.nlm.nih.gov/14657504/)
6. Esau C, Kang X, Peralta E, Hanson E, Marcusson EG, Ravichandran LV, et al. (2004) MicroRNA-143 regulates adipocyte differentiation. *J Biol Chem* 279: 52361–52365. PMID: [15504739](https://pubmed.ncbi.nlm.nih.gov/15504739/)
7. He L and Hannon GJ (2004) MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 5: 522–531. PMID: [15211354](https://pubmed.ncbi.nlm.nih.gov/15211354/)
8. Brennecke J, Stark A, Russell RB and Cohen SM (2005) Principles of microRNA-target recognition. *PLoS Biol* 3:e85. PMID: [15723116](https://pubmed.ncbi.nlm.nih.gov/15723116/)
9. Gregory RI and Shiekhattar R (2005) MicroRNA biogenesis and cancer. *Cancer Res* 65: 3509–3512. PMID: [15867338](https://pubmed.ncbi.nlm.nih.gov/15867338/)
10. Fabian MR and Sonenberg N (2012) The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. *Nat Struct Mol Biol*. 19:586–593 doi: [10.1038/nsmb.2296](https://doi.org/10.1038/nsmb.2296) PMID: [22664986](https://pubmed.ncbi.nlm.nih.gov/22664986/)
11. Lund E, Güttinger S, Calado A, Dahlberg JE and Kutay U (2004) Nuclear export of microRNA precursors. *Science* 303: 95–98. PMID: [14631048](https://pubmed.ncbi.nlm.nih.gov/14631048/)

12. Bartel DP (2013) MicroRNAs: target recognition and regulatory functions. *Cell* 136: 215–233.
13. Ha I, Wightman B and Ruvkun G (1996) A bulged lin-4/lin-14 RNA duplex is sufficient for *Caenorhabditis elegans* lin-14 temporal gradient formation. *Genes Dev* 10: 3041–3050. PMID: [8957004](#)
14. Vella MC, Choi EY, Lin SY, Reinert K and Slack FJ (2004) The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. *Genes Dev* 18: 132–137. PMID: [14729570](#)
15. Rana TM (2007) Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Biol* 8: 23–36. PMID: [17183358](#)
16. Guo H, Ingolia NT, Weissman JS and Bartel DP (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 466: 835–840. doi: [10.1038/nature09267](#) PMID: [20703300](#)
17. Bazzini AA, Lee MT and Giraldez AJ (2012) Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. *Science* 336: 233–237. doi: [10.1126/science.1215704](#) PMID: [22422859](#)
18. Janas MM and Novina CD (2012) Not lost in translation: stepwise regulation of microRNA targets. *EMBO J* 31: 2446–2447. doi: [10.1038/emboj.2012.119](#) PMID: [22522704](#)
19. Hutvagner G and Zamore PD (2002). A microRNA in a multiple-turnover RNAi enzyme complex. *Science* 297: 2056–2060. PMID: [12154197](#)
20. Grimson A, Farh KK, Johnston WK, Garrett-Engle P, Lim LP and Bartel DP. (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell* 27: 91–105. PMID: [17612493](#)
21. Lee I, Ajay SS, Yook JI, Kim HS, Hong SH, Kim NH, et al. (2009) New class of microRNA targets containing simultaneous 5'-UTR and 3'-UTR interaction sites. *Genome Res* 19: 1175–1183. doi: [10.1101/gr.089367.108](#) PMID: [19336450](#)
22. Doench JG and Sharp PA (2004) Specificity of microRNA target selection in translational repression. *Genes Dev* 18: 504–511. PMID: [15014042](#)
23. Lewis BP, Burge CB and Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120: 15–20. PMID: [15652477](#)
24. Lim LP, Lau NC, Garrett-Engle P, Grimson A, Schelter JM, Castle J, et al. (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 433: 769–773. PMID: [15685193](#)
25. Meijer HA, Kong YW, Lu WT, Wilczynska A, Spriggs RV, Robinson SW, et al. (2013) Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation. *Science* 340: 82–85. doi: [10.1126/science.1231197](#) PMID: [23559250](#)
26. Miyoshi K, Miyoshi T and Siomi H (2010) Many ways to generate microRNA-like small RNAs: non-canonical pathways for microRNA production. *Mol Genet Genomics* 284: 95–103. doi: [10.1007/s00438-010-0556-1](#) PMID: [20596726](#)
27. Yang JS and Lai EC (2011) Alternative miRNA biogenesis pathways and the interpretation of core miRNA pathway mutants. *Mol Cell* 43: 892–903. doi: [10.1016/j.molcel.2011.07.024](#) PMID: [21925378](#)
28. Brennecke J, Stark A, Russell RB and Cohen SM (2005) Principles of microRNA-target recognition. *PLoS Biol* 3: e85. PMID: [15723116](#)
29. Mishra AK, Agarwal S, Jain CK and Rani V (2009) High GC content: critical parameter for predicting stress regulated miRNAs in *Arabidopsis thaliana*. *Bioinformatics* 4: 151–154. PMID: [20198191](#)
30. Paraskevopoulou MD, Georgakilas G, Kostoulas N, Vlachos IS, Vergoulis T, Reczko M, et al. (2013) DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. *Nucleic Acids Res.* W169–73. doi: [10.1093/nar/gkt393](#) PMID: [23680784](#)
31. Rother M, Rother K, Puton T and Bujnicki JM (2011) ModeRNA: a tool for comparative modeling of RNA 3D structure. *Nucleic Acids Res* 39: 4007–4022. doi: [10.1093/nar/gkq1320](#) PMID: [21300639](#)
32. Fang Z, Du R, Edwards A, Flemington EK and Zhang K (2013) The sequence structures of human microRNA molecules and their implications. *PLoS One* 8:e54215. doi: [10.1371/journal.pone.0054215](#) PMID: [23349828](#)
33. Filipowicz W, Bhattacharyya SN and Sonenberg N (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* 9: 102–114. doi: [10.1038/nrg2290](#) PMID: [18197166](#)
34. Wang B (2013) Base Composition Characteristics of Mammalian miRNAs. *J Nucleic Acids* 2013: 951570. doi: [10.1155/2013/951570](#) PMID: [23710337](#)
35. Sawicka K, Bushell M, Spriggs KA and Willis AE (2008) Polypyrimidine-tract-binding protein: a multi-functional RNA-binding protein. *Biochem Soc Trans* 36: 641–647. doi: [10.1042/BST0360641](#) PMID: [18631133](#)
36. Shine J and Dalgarno L (1975) Determinant of cistron specificity in bacterial ribosomes. *Nature* 254: 34–38. PMID: [803646](#)



37. Kozak M (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* 15: 8125–8148. PMID: [3313277](#)
38. Ito K, Go S, Komiyama M and Xu Y (2011) Inhibition of translation by small RNA-stabilized mRNA structures in human cells. *J Am Chem Soc* 133: 19153–19159. doi: [10.1021/ja206353c](#) PMID: [22007660](#)
39. Gkirtzou K, Tsamardinos I, Tsakalides P and Poirazi P (2010) MatureBayes: a probabilistic algorithm for identifying the mature miRNA within novel precursors. *PLoS One* 5:e11843. doi: [10.1371/journal.pone.0011843](#) PMID: [20700506](#)
40. Katti MV, Ranjekar PK and Gupta VS (2001) Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Mol Biol Evol* 18: 1161–1167. PMID: [11420357](#)
41. Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S and McCouch S (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. *Genome Res* 11:1441–52. PMID: [11483586](#)
42. Dereeper A, Argout X, Billot C, Rami JF and Ruiz M (2007) SAT, a flexible and optimized Web application for SSR marker development. *BMC Bioinformatics* 29; 8:465.
43. Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, et al.(2002). Genetic structure of human populations. *Science* 298: 2381–2385. PMID: [12493913](#)
44. Abdurakhmonov IY, Abdullaev AA, Saha S, Buriev ZT, Arslanov D, Kuryazov Z, et al. (2005) Simple sequence repeat marker associated with a natural leaf defoliation trait in tetraploid cotton. *J Hered* 96: 644–653. PMID: [16159909](#)
45. Flannery ML, Mitchell FJ, Coyne S, Kavanagh TA, Burke JI, Salamin N, et al. (2006) Plastid genome characterization in Brassica and Brassicaceae using a new set of nine SSRs. *Theor Appl Genet* 113: 1221–1231. PMID: [16909279](#)
46. Foster EA, Jobling MA, Taylor PG, Donnelly P, de Knijff P, Mieremet R, et al. (1998) Jefferson fathered slave's last child. *Nature* 396: 27–28. PMID: [9817200](#)
47. Mein CA, Esposito L, Dunn MG, Johnson GC, Timms AE, Goy JV, et al. (1998) A search for type 1 diabetes susceptibility genes in families from the United Kingdom. *Nat Genet* 19: 297–300. PMID: [9662409](#)
48. Pritchard JK, Seielstad MT, Perez-Lezaun A, Feldman MW (1999) Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Mol Biol Evol* 16: 1791–1798. PMID: [10605120](#)
49. Chen MI, Tan Z, Zeng G and Peng J (2010) Comprehensive analysis of simple sequence repeats in pre-miRNAs. *Mol Biol Evol* 27: 2227–2232. doi: [10.1093/molbev/msq100](#) PMID: [20395311](#)
50. Subramanian S, Madgula VM, George R, Mishra RK, Pandit MW, Kumar CS, et al. (2003). Triplet repeats in human genome: distribution and their association with genes and other genomic regions. *Bioinformatics* 19: 549–552. PMID: [12651711](#)
51. Halees AS, Hitti E, Al-Saif M, Mahmoud L, Vlasova-St Louis IA, Beisang DJ, et al. (2011) Global assessment of GU-rich regulatory content and function in the human transcriptome. *RNA Biol* 8:681–691. doi: [10.4161/ma.8.4.16283](#) PMID: [21691152](#)
52. Vlasova IA, Tahoe NM, Fan D, Larsson O, Rattenbacher B, Sternjohn JR, Vasdewani J, Karypis G, Reilly CS, Bitterman PB and Bohjanen PR (2008) Conserved GU-rich elements mediate mRNA decay by binding to CUG-binding protein. *Mol Cell* 29:263–270. doi: [10.1016/j.molcel.2007.11.024](#) PMID: [18243120](#)
53. Lee JE and Cooper TA. Pathogenic mechanisms of myotonic dystrophy (2009) *Biochem Soc Trans* 37:1281–1286. doi: [10.1042/BST0371281](#) PMID: [19909263](#)
54. Timchenko NA, Wang GL and Timchenko LT (2005). RNA CUG-binding protein 1 increases translation of 20-kDa isoform of CCAAT/enhancer-binding protein beta by interacting with the alpha and beta subunits of eukaryotic initiation translation factor 2. *J Biol Chem* 280:20549–20557. PMID: [15788409](#)
55. Masuda A, Andersen HS, Doktor TK, Okamoto T, Ito M, Andresen BS, et al. (2012) CUGBP1 and MBNL1 preferentially bind to 3' UTRs and facilitate mRNA decay. *Sci Rep* 2, 209. doi: [10.1038/srep00209](#) PMID: [22355723](#)
56. FuY Ramisetty SR, Hussain N and Baranger AM (2012) MBNL1-RNA recognition: contributions of MBNL1 sequence and RNA conformation. *ChemBiochem* 13: 112–119. doi: [10.1002/cbic.201100487](#) PMID: [22106026](#)
57. Kiliszek A, Kierzek R, Krzyzosiak WJ and Rypniewski W (2012) Crystallographic characterization of CCG repeats. *Nucleic Acids Res* 40: 8155–8162. doi: [10.1093/nar/gks557](#) PMID: [22718980](#)
58. Robbins M, Judge A and MacLachlan I (2009). siRNA and innate immunity. *Oligonucleotides* 19: 89–102. doi: [10.1089/oli.2009.0180](#) PMID: [19441890](#)

59. Stewart CR, Karpala AJ, Lowther S, Lowenthal JW and Bean AG (2011) Immunostimulatory motifs enhance antiviral siRNAs targeting highly pathogenic avian influenza H5N1. *PLoS One* 6:e21552. doi: [10.1371/journal.pone.0021552](https://doi.org/10.1371/journal.pone.0021552) PMID: [21747939](https://pubmed.ncbi.nlm.nih.gov/21747939/)
60. Hervas-Stubbs S, Perez-Gracia JL, Rouzaut A, Sanmamed MF, Le Bon A and Melero I (2011). Direct effects of type I interferons on cells of the immune system. *Clin. Cancer Res.* 17:2619–2627. doi: [10.1158/1078-0432.CCR-10-1114](https://doi.org/10.1158/1078-0432.CCR-10-1114) PMID: [21372217](https://pubmed.ncbi.nlm.nih.gov/21372217/)
61. Plataniias LC (2005). Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nat. Rev. Immunol.* 5:375–386. PMID: [15864272](https://pubmed.ncbi.nlm.nih.gov/15864272/)
62. Yang H, Zhang H, Zhu L, Wang J, Zhang C and Li D (2012) Pathway analysis of cancer-associated microRNA targets. *Int J Oncol.* 41:2213–2226. doi: [10.3892/ijo.2012.1658](https://doi.org/10.3892/ijo.2012.1658) PMID: [23064433](https://pubmed.ncbi.nlm.nih.gov/23064433/)
63. Wagner EF and Nebreda AR (2009) Signal integration by JNK and p38 MAPK pathways in cancer development. *Nat Rev Cancer.* 9:537–549. doi: [10.1038/nrc2694](https://doi.org/10.1038/nrc2694) PMID: [19629069](https://pubmed.ncbi.nlm.nih.gov/19629069/)
64. Osaki M, Oshimura M and Ito H (2004) PI3K-Akt pathway: its functions and alterations in human cancer. *Apoptosis* 9:667–676. PMID: [15505410](https://pubmed.ncbi.nlm.nih.gov/15505410/)
65. Carey J, Cameron V, de Haseth PL and Uhlenbeck OC (1983) Sequence-specific interaction of R17 coat protein with its ribonucleic acid binding site. *Biochemistry* 22: 2601–2610. PMID: [6347247](https://pubmed.ncbi.nlm.nih.gov/6347247/)
66. Jucker FM and Pardi A (1995) GNRA tetraloops make a U-turn. *RNA* 1: 219–222. PMID: [7585251](https://pubmed.ncbi.nlm.nih.gov/7585251/)
67. Cheong C, Varani G and Tinoco I Jr (1990) Solution structure of an unusually stable RNA hairpin, 5'GGAC(UUCG)GUCC. *Nature* 346: 680–682. PMID: [1696688](https://pubmed.ncbi.nlm.nih.gov/1696688/)
68. Belter A, Gudanis D, Rolle K, Piwecka M, Gdaniec Z, Naskret-Barciszewska MZ, et al. (2014) Mature miRNAs form secondary structure, which suggests their function beyond RISC. *PLoS One.* 9: e113848. doi: [10.1371/journal.pone.0113848](https://doi.org/10.1371/journal.pone.0113848) PMID: [25423301](https://pubmed.ncbi.nlm.nih.gov/25423301/)
69. Kumar Das A, Ganguli S, Gupta S and Datta A (2011). Secondary structural analysis of microRNA and their precursors in plants. *Int J Agriculture Sciences* 3: 62–64.
70. Maiti M, Nauwelaerts K, Lescrinier E, Schuit FC and Herdewijn P (2010) Self-complementary sequence context in mature miRNAs. *Biochem Biophys Res Commun* 392: 572–576. doi: [10.1016/j.bbrc.2010.01.072](https://doi.org/10.1016/j.bbrc.2010.01.072) PMID: [20097176](https://pubmed.ncbi.nlm.nih.gov/20097176/)
71. Taulli R, Loretelli C and Pandolfi PP (2013) From pseudo-ceRNAs to circ-ceRNAs: a tale of cross-talk and competition. *Nat Struct Mol Biol* 20: 541–543. doi: [10.1038/nsmb.2580](https://doi.org/10.1038/nsmb.2580) PMID: [23649362](https://pubmed.ncbi.nlm.nih.gov/23649362/)