

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

miRviewer: A multispecies microRNA homologous viewer

BMC Research Notes 2012, **5**:92 doi:10.1186/1756-0500-5-92

Adam Kiezun (akiezun@rics.bwh.harvard.edu)
Shay Artzi (artzi@csail.mit.edu)
Shira Modai (shiram@gmail.com)
Naama Volk (naamavolk@gmail.com)
Ofer Isakov (oferisak@post.tau.ac.il)
Noam Shomron (nshomron@post.tau.ac.il)

ISSN 1756-0500

Article type Technical Note

Submission date 6 June 2011

Acceptance date 13 February 2012

Publication date 13 February 2012

Article URL <http://www.biomedcentral.com/1756-0500/5/92>

This peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in *BMC Research Notes* are listed in PubMed and archived at PubMed Central.

For information about publishing your research in *BMC Research Notes* or any BioMed Central journal, go to

<http://www.biomedcentral.com/info/instructions/>

miRviewer: A multispecies microRNA homologous viewer

ArticleCategory	:	Technical Note
ArticleHistory	:	Received: 06-Jun-2011; Accepted: 17-Jan-2012
ArticleCopyright	:	© 2012 Kiezun et al; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Adam Kiezun,^{Aff1†}

Email: akiezun@rics.bwh.harvard.edu

Shay Artzi,^{Aff2†}

Email: artzi@csail.mit.edu

Shira Modai,^{Aff3}

Email: shiram@gmail.com

Naama Volk,^{Aff3}

Email: naamavolk@gmail.com

Ofer Isakov,^{Aff3}

Email: oferisak@post.tau.ac.il

Noam Shomron,^{Aff3}

Corresponding Affiliation: Aff3

Phone: +972-3-6406594

Fax: +972-3-6407432

Email: nshomron@post.tau.ac.il

Aff1 Division of Genetics, Department of Medicine, Brigham and Woman's Hospital, Harvard Medical School, Boston MA 02115, USA

Aff2 Program Analysis and Transformation Group, IBM, Watson Research Center, Hawthorne, NY 10532, USA

Aff3 Department of Cell and Developmental Biology, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

†Contributed equally to this work

Abstract

Background

MicroRNAs (miRNAs) are short non-coding RNAs that regulate gene expression via binding to the 3' ends of mRNAs. MiRNAs have been associated with many cellular events ascertaining their central role in gene regulation. In order to better understand miRNAs of interest it is of utmost importance to learn about the genomic conservation of these genes.

Findings

The miRviewer web-server, presented here, encompasses all known miRNAs of currently fully annotated animal genomes in a visual 'birds-eye' view representation. miRviewer provides a graphical outlook of the current miRNA world together with sequence alignments and secondary structures of each miRNA. As a test case we experimentally examined the expression of several miRNAs in various animals.

Conclusions

miRviewer completes the homologous miRNA space with hundreds of unreported miRNAs and is available at: <http://people.csail.mit.edu/akiezun/miRviewer>

Findings

The gene expression pathway is very often regulated at the post-transcriptional level by small non-coding RNAs termed microRNAs (miRNAs) [1]. MiRNAs are 22 nucleotides long molecules that bind the 3' untranslated (3' UTR) region of the mRNA leading to facilitated mRNA degradation or inhibition of translation [2]. MiRNAs are pivotal regulators in many cellular processes such as development and differentiation (for example see [3]). Each miRNA has on average 200 conserved targets identified by miRNA:target sequence complementarity [4]. The miRNA genes are transcribed mostly by RNA polymerase II and are either expressed from intergenic regions or as part of a coding transcript when located in introns [4]. Most miRNAs were generated via genomic duplication events, though other events, such as through repetitive genomic sequences, were also observed. The unique evolutionary features of miRNAs are their high gain versus loss ratio and their resistance to changes or mutations particularly in the mature and functional targeting "seed" region [5].

In order to better understand a miRNA of interest, whether to learn about its genomic evolution, its potential target genes in various organisms, or in order to select a model animal to work on, it is important to comprehend its evolutionary conservation. Here we present the miRviewer web-server, which provides a comprehensive view of all known miRNAs in annotated animal genomes. The easily assessable data represents the miRNA gene repertoire in 50 animals. The web interface provides the user a simple and graphic view of the miRNA world. Additional layers of information, such as conservation, multiple alignments of the sequences, indication of mutated miRNAs, genomic loci and secondary structure, are also available.

Implementation

miRviewer presents homologous miRNAs that are either included in miRbase v.16 [6] or were identified by a full cross-search using miRNAmirer [7]. Specifically, for each miRNA in miRbase v.16 a search was performed in all animal genomes in Ensembl 57 [8], using default miRNAmirer settings. The most significant match identified was used in creating the current miRviewer web-server (Figures 1 and 2).

Figure 1 A partial screenshot of miRviewer's main entry page. Species are indicated on the X-axis while miRNAs are on the Y-axis. Level of conservation is indicated in shades of green and numerical values are indicated on top. Current sorting is human centric and is ordered by most to least conserved miRNAs. A dot at the center of each box indicates the presence of the miRNA in miRbase [6]. Lack of such dot indicates the identification of the miRNA gene through miRNAmirer [7]

Figure 2 A screenshot of miR-98a multiple alignment in different organisms. Though this magnification does not allow reading the exact sequences, it does allow viewing each row representing miR-98a from a different animal. Mature (blue), mutation (red) and all other sequences (gray) are indicated (also see text). On the right are links to miRbase, genomic locations on UCSC Genome Browser [9] and Ensembl [8]. Below the alignments, conservation of the pre-miRNA sequence is plotted

The candidate miRNAs identified by our miRNAmirer method [7] were recognized using the following parameters: (i) Use BLAST [<http://www.ncbi.nlm.nih.gov/BLAST>] to find matches in target genomes (the whole precursor miRNA from the query is used); (ii) Filter with e-value threshold (default 0.05 per chromosome); (iii) Extend the match by adding flanking nucleotides (default 50) up- and down-stream from the match (Ensembl genome database; [<http://www.ensembl.org>]). Examine all possible extensions of the match within threshold length (default min 70 nt, max 180 nt); (iv) Filter with RNA secondary folding energy threshold (default -25 kcal/mole; RNAfold with options "-p -d2 -noLP" [<http://www.tbi.univie.ac.at/RNA/>]); (v) Filter with minimal base-pairing threshold (default 55% pairing; with 20 gap penalty and 0.5 extension penalty); (vi) Filter with requirement for hairpin-shape secondary structure; (vii) Filter with alignment of precursor sequences (default 56% identity); (viii) Filter with alignment of mature miRNA sequences (default 80% identity); (ix) Filter with maximum number of mismatches in mature miRNA sequences (default 3 nt); (x) Filter with conservation of seed (2–8 nt, required 100% conservation [10]); (xi) Filter with position of mature miRNA on the hairpin (max 4 nt overlap of mature sequence and hairpin loop). In this study we looked at homolog genes which are genes related to each other by descent from a common ancestral DNA sequence. We do not segregate between orthologs, genes in different species that evolved from a common ancestral gene by speciation, and paralogs, genes separated by the event of genetic duplication. miRviewer presents novel miRNAs that are evolutionary conserved based on previously defined features [11]. The parameters used to find these miRNAs are considered stringent as they identify homolog matches after a miRNA has been experimentally confirmed and the default search filters such that it selects 95% of known miRNAs in training genomes (criteria was also based on [11]).

On the main web-server page the conservation degree was calculated as the proportion of identical bases. In the alignment interface the mature miRNA sequences are designated by blue highlighting. Outlier nucleotides are colored red. However, we note that in some cases the BLAST searches identify a mature miRNA without the first nucleotide even when the first nucleotide is identical. Thus, some miRNAs might be ‘missing’ a (colored) first nucleotide. The reason for this perhaps occurs due to the inherent features of the search and should be cautiously taken into consideration. Given that many miRNAs are currently reported to have several isoforms [12–14] we suggest using our web-server in conjugation with the ‘deep sequencing view’ of miRbase which allows the evaluation of the small changes in mature miRNA nucleotide ends. The secondary structure of each miRNA, presented to the right of each sequence (Figure 3), allows the appreciation of the folding of each miRNA sequence. Close observation of the secondary structure of each miRNA between the various species carries relevant information such as the development of tight or relaxed hairpin structures.

Figure 3 A partial screenshot of known and novel miR-21 secondary structures. This view, presented alongside (to the right) of the sequence alignment (presented in Figure 2) allows viewing the predicted secondary structure (based on RNAfold, <http://www.tbi.univie.ac.at/RNA>) of a specific miRNA among the various organisms. The three letter abbreviations represent the animal and the negative numbers indicate the folding energy (ΔG)

Methods

Preparing RNA samples and testing miRNA expression

Brain marmoset RNA was received from Dr. Yona Goldshmit. Rabbit heart sample were received from Dr. Michael Har-Lev. Dolphin epithelium sample was received from Dr. Nadav Shashar. HeLa cell line was used as a positive control for known miRNAs (data not shown). Tissue samples were homogenized prior to RNA extraction. Total RNA was extracted from all samples using TRIzol reagent (Invitrogen) according to the manufacturer’s instructions. One microgram of total RNA was used to generate cDNA using the High Capacity Reverse Transcription Kit with random primers (Applied Biosystems) according to manufacturer’s instruction and in a final volume of 20 μ L. PCR amplification was done using C1000 Thermal Cycler (Bio-Rad) in the following conditions: enzymatic activation at 94°C for 3 min, 30–35 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 30 s and elongation at 72°C for 30 s and then extra 5 min of elongation at 72°C. miRNA-16-1 and miRNA-101-1 species specific primers were obtained from the miRViewer latest version (pre-miR-16-1F: 5'-TCA GCA GTG CCT TAG CAG C-3', pre-miR-16-1R: 5'-CAA CCT TAC TTC AGC AGC AC-3', pre-miR-101-1F: 5'-TGG CTC AGT TAT CAC AGT GC-3', pre-miR-101-1R: 5'-TGC CAT CCT TCA GTT ATC ACA-3'). Total RNA only, of each sample, was used as a negative control for the PCR reaction. Samples were stained with Ethidium Bromide and run in a 2% Agarose gel. Samples were then extracted from the gel using Wizard SV Gel (Promega) according to the manufacturer’s instructions and sequenced primed by the PCR primers.

Sequencing and analysis

Human SupT1 and Hela cell lines were used in the deep sequencing portion. Total RNA was extracted using TRIzol reagent (Invitrogen) and each sample was prepared for deep sequencing following Illumina's Small RNA sample preparation protocol v1.5. Briefly, samples were ligated with 3' and 5' adapters, reverse-transcribed and then PCR amplified. cDNA library was prepared from 93–100bp PCR products and sequenced in separate lanes on an Illumina Genome Analyzer IIx instrument at the Tel Aviv University Genome High-Throughput Sequencing Laboratory. Adapters were removed from the output files using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit), discarding all reads shorter than 16 nucleotides after clipping. Alignment was performed using Burrows-Wheeler Aligner (BWA) ([15]), against a reference of all the predicted miRviewer mature miRNAs sequences with additional three nucleotides flank. Hits were counted when a read aligned perfectly to a predicted miRNA (alignment against the transcript sense strand without any mismatches). Target conservation was calculated by applying TargetScan [10] on the seed region of each predicted miRNA.

Findings and discussion

miRviewer is a web-server which allows the global overview of miRNAs in the animal kingdom. Our web-server is different and has unique features over other miRNA gene catalogues online. miRortho [16] for example, which contains valuable information on animal miRNAs, does not allow a one screen view of all miRNAs across the animal phyla. Another useful web-server, MapMi [17], does not exhibit an alignment of the miRNA matched sequences. Importantly, miRviewer is based on additional novel miRNA datasets which were identified using our miRNAmminer algorithm [7]. Of particular interest, for 27 organisms we report, for the first time, their entire set of annotated homologous miRNAs. Thus, our web-server has both known and predicted miRNAs, clearly differentiated on miRviewer webpage.

In order to confirm a small representation of predicated miRNAs we chose to amplify precursor miR-16 and miR-101 from monkey, rabbit and dolphin total RNA (see Methods and Figure 4). All three mammals showed expression of these miRNAs in the tested tissues. Similarly, when parsing deep sequencing data of two human cell lines for novel miRviewer human mature miRNAs 12 of them were identified to be expressed at more than 10 reads per miRNA (Table 1). Some of these miRNAs have many conserved mRNA target sites suggesting a possible functional role during cellular processes. These two sets of experiments indicate that miRNAmminer, and hence miRviewer, represent genuine novel miRNAs and strengthen our comprehensive view of the miRNA world. We caution that by no means have these experiments intended to represent a complete analysis of all predicted miRviewer miRNAs but rather to indicate that future confirmation is feasible. We also note that these results validate not only our predictive algorithm but also the utility of our web-server for species specific primer design.

Figure 4 Confirmation of the expression of two novel miRNAs, miR-101 and miR-16, in monkey (marmoset), rabbit and dolphin. PCR was used to amplify expressed miRNAs based on presence in miRviewer. Amplified fragments were stained and separated in Agarose gel then excised and sequenced for confirmation (Additional file 1: Figure S1). These miRNAs were

identified as part of a very large set of homologous miRNAs using the miRNAminer tool [7] all of which are presented in miRviewer

Table 1 Novel miRviewer identified human mature miRNAs discovered using deep sequencing

miRNA Name	Hela	SupT1	Total	Mature Sequence	Seed	Conserved Targets	Conserved Sites
hsa-miR-365-1*	407	901	1308	ACCGAUUUUCAA AUGGUGCUA	CCGAUUU	9	9
hsa-miR-652*	180	135	315	GCCCUUUUAACA UUGCACUG	CCCUUUU	486	521
hsa-miR-664b	131	98	229	AACACCAUUGUCACACUCCA	ACACCAU	195	204
hsa-miR-301a*	79	27	106	CAACCCUAGGAGAGGGUGCCAUUC	AACCCUA	70	72
hsa-miR-513a	0	65	65	AGGGGUUCACCGAGCAACA UUCG	GGGGUUC	88	110
hsa-miR-128-1*	21	37	58	ACGGGUUAUUCUUGGGUGGAUAA	CGGGUUAU	2	2
hsa-miR-152*	23	0	23	GGAAACA UUCUGCACAAACUAG	GAAACA U	360	375
hsa-miR-1249*	1	20	21	GCUGGUAAAAUGGAACCAAA	CUGGUAA	146	151
hsa-miR-103-1*	13	4	17	GGCUUCUUUACAGUGCUGCCUUG	GCUUCUU	392	412
hsa-miR-2145	12	4	16	GCGACGAGCCCCUCGCACAAACC	CGACGAG	1	1
hsa-miR-107*	1	14	15	AGCUUCUUUACAGUGUUGCCUUG	GCUUCUU	392	412
hsa-miR-203*	11	0	11	GUGGAUAUCCUUCUAUGUUUA	UGGAUAU	116	119

* refers to the minor miRNA isoform (based on miRbase)

Future developments of miRviewer include incorporation of additional animal genomes as they are sequenced and annotated. Adding further layers of information such as performing statistical analysis of each miRNA family and allowing the user to choose an animal to display as a reference set (the current reference set is based on human). We would like to note that the main webpage would remain the same for simplicity and ease of use. The additional layers of information would be added as hyperlink.

Conclusions

The miRviewer web-server, presented here, encompasses all known and predicted miRNAs of fully annotated animal genomes. miRviewer's interface provides the user with a graphic outlook of the current miRNA world together with additional information such as sequence alignments and secondary structure. miRviewer is available at:
<http://people.csail.mit.edu/akiezun/miRviewer>

Availability and requirements

miRviewer is available online at: <http://people.csail.mit.edu/akiezun/miRviewer>

Competing interests

None declared.

Authors' contributions

NS, AK and SA conceived and designed the web-server. AK and SA wrote and ran the algorithms to search for homologous miRNAs and constructed the web-server. SM and NV performed the experimental validation. OI performed the deep sequencing data analysis. All authors read, contributed to and approved the final manuscript.

Acknowledgments

Work in the Shomron laboratory is supported by the Chief Scientist Office, Ministry of Health, Israel; The Kurz-Lion Foundation; Israel Cancer Association; Ori Levy Foundation for Mitochondrial Research; The Wolfson Family Charitable Fund. This research was also supported by the I-CORE Program of the Planning and Budgeting Committee and The Israel Science Foundation (grant number 41/11). OI is supported by the Safra Bioinformatics Fellowship.

References

1. Fabian MR, Sonenberg N, Filipowicz W: **Regulation of mRNA translation and stability by microRNAs**. *Annu Rev Biochem* 2010, **79**:351–379.
2. Bushati N, Cohen SM: **microRNA functions**. *Annu Rev Cell Dev Biol* 2007, **23**:175–205.
3. Boyerinas B, Park SM, Shomron N, Hedegaard MM, Vinther J, Andersen JS, Feig C, Xu J, Burge CB, Peter ME: **Identification of let-7-regulated oncofetal genes**. *Cancer Res* 2008, **68**(8):2587–2591.
4. Shomron N, Golan D, Hornstein E: **An evolutionary perspective of animal microRNAs and their targets**. *J Biomed Biotechnol* 2009, **2009**:594738.
5. Bartel DP: **MicroRNAs: target recognition and regulatory functions**. *Cell* 2009, **136**(2):215–233.
6. Griffiths-Jones S: **miRBase: microRNA sequences and annotation**. *Curr Protoc Bioinformatics* 2010, Chapter 12:Unit 12.9.1–10.
7. Artzi S, Kiezun A, Shomron N: **miRNAMiner: a tool for homologous microRNA gene search**. *BMC Bioinforma* 2008, **9**:39.
8. Flicek P, Amode MR, Barrell D, Beal K, Brent S, Chen Y, Clapham P, Coates G, Fairley S, Fitzgerald S, Gordon L, Hendrix M, Hourlier T, Johnson N, Kähäri A, Keefe D, Keenan S, Kinsella R, Kokocinski F, Kulesha E, Larsson P, Longden I, McLaren W, Overduin B, Pritchard B, Riat HS, Rios D, Ritchie GR, Ruffier M, Schuster M, Sobral D, Spudich G, Tang YA, Trevanion S, Vandrovcova J, Vilella AJ, White S, Wilder SP, Zadissa A, Zamora J, Aken BL, Birney E, Cunningham F, Dunham I, Durbin R, Fernández-Suarez XM, Herrero J, Hubbard TJ,

Parker A, Proctor G, Vogel J, Searle SM: **Ensembl 2011**. *Nucleic Acids Res* 2011, **39**(Database issue):D800–D806.

9. Sanborn JZ, Benz SC, Craft B, Szeto C, Kober KM, Meyer L, Vaske CJ, Goldman M, Smith KE, Kuhn RM, Karolchik D, Kent WJ, Stuart JM, Haussler D, Zhu J: **The UCSC Cancer Genomics Browser: update 2011**. *Nucleic Acids Res* 2011, **39**(Database issue):D951–D959.

10. Lewis BP, Burge CB, Bartel DP: **Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets**. *Cell* 2005, **120**(1):15–20.

11. Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen X, Dreyfuss G, Eddy SR, Griffiths-Jones S, Marshall M, Matzke M, Ruvkun G, Tuschl T: **A uniform system for microRNA annotation**. *RNA* 2003, **9**(3):277–279.

12. Fernandez-Valverde SL, Taft RJ, Mattick JS: **Dynamic isomiR regulation in Drosophila development**. *RNA* 2010, **16**(10):1881–1888.

13. Lee LW, Zhang S, Etheridge A, Ma L, Martin D, Galas D, Wang K: **Complexity of the microRNA repertoire revealed by next-generation sequencing**. *RNA* 2010, **16**(11):2170–2180.

14. Pantano L, Estivill X, Martí E: **SeqBuster, a bioinformatic tool for the processing and analysis of small RNAs datasets, reveals ubiquitous miRNA modifications in human embryonic cells**. *Nucleic Acids Res* 2010, **38**(5):e34.

15. Li H, Durbin R: **Fast and accurate short read alignment with Burrows-Wheeler transform**. *Bioinformatics (Oxford, England)* 2009, **25**(14):1754–1760.

16. Gerlach D, Kriventseva EV, Rahman N, Vejnar CE, Zdobnov EM: **miROrtho: computational survey of microRNA genes**. *Nucleic Acids Res* 2009, **37**(Database issue):D111–D117.

17. Guerra-Assunção JA, Enright AJ: **MapMi: automated mapping of microRNA loci**. *BMC Bioinforma* 2010, **11**:133.

Additional file

Additional_file_1 as DOC

Additional file 1: Figure S1. Sequencing results of pre-miRNA-16-1 and pre-miRNA-101-1 of (A) Rabbit and (B) Marmoset PCR samples, compared to miRViewer (and miRNAMiner) predictions.

conservation	1.0	0.97	0.94	0.91	0.88	0.85	0.82	0.79	0.76	0.73	0.7	0.67	0.64
--------------	-----	------	------	------	------	------	------	------	------	------	-----	------	------

miRNA	hsa	ptr	ggo	ppv	mml	cja	tsy	mmr	oga	tbe	cpo	dor	mmu	rno	str
Known/new	792/48	426/15	68/508	442/163	350/132	0/381	0/194	0/248	0/250	0/205	0/272	0/224	471/40	324/57	0/228
let-7 (28)	*	*		*	*								*	*	
mir-34 (10)	*	*	*	*	*								*	*	
mir-133 (10)	*	*	*	*	*								*	*	
mir-92 (10)	*	*	*	*	*								*	*	
mir-216 (8)	*	*	*	*	*								*	*	
mir-7 (9)	*	*	*	*	*								*	*	
mir-124 (12)	*	*	*	*	*								*	*	
mir-101 (7)	*	*	*	*	*								*	*	
mir-135 (11)	*	*	*	*	*								*	*	
mir-15 (6)	*	*	*	*	*								*	*	
mir-181 (9)	*	*	*	*	*								*	*	
mir-196 (10)	*	*	*	*	*								*	*	
mir-29 (13)	*	*	*	*	*								*	*	
mir-125 (11)	*	*	*	*	*								*	*	
mir-1 (9)	*	*		*	*								*	*	

pre-miR-101

pre-miR-16

Monkey

Rabbit

Dolphin

Monkey

Rabbit

Dolphin

+

-

+

-

+

-

+

-

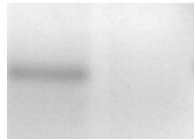
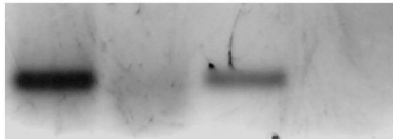
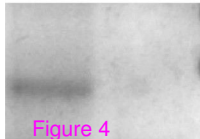
+

-

+

-

cDNA



Additional files provided with this submission:

Additional file 1: Additional file 1.doc, 41K

<http://www.biomedcentral.com/imedia/5279070796735162/supp1.doc>