miRCarta: a central repository for collecting miRNA candidates

Christina Backes[1,*], Tobias Fehlmann[1], Fabian Kern[1], Tim Kehl[2], Hans-Peter Lenhof[2], Eckart Meese[3] and Andreas Keller[1]

[1]Chair for Clinical Bioinformatics, Saarland Informatics Campus, Saarland University, Germany, [2]Center for Bioinformatics, Saarland Informatics Campus, Saarland University, Germany and [3]Institute for Human Genetics, Medical School, Saarland University, Germany

ABSTRACT

The continuous increase of available biological data as consequence of modern high-throughput technologies poses new challenges for analysis techniques and database applications. Especially for miRNAs, one class of small non-coding RNAs, many algorithms have been developed to predict new candidates from next-generation sequencing data. While the amount of publications describing novel miRNA candidates keeps steadily increasing, the current gold standard database for miRNAs - miRBase - has not been updated since June 2014. As a result, publications describing new miRNA candidates in the last three to five years might have a substantial overlap of candidates without noticing. With miRCarta we implemented a database to collect novel miRNA candidates and augment the information provided by miRBase. In the first stage, miRCarta is thought to be a highly sensitive collection of potential miRNA candidates with a high degree of analysis functionality, annotations and details on each miRNA. We added—besides the full content of the miRBase—12,857 human miRNA precursors to miRCarta. Users can match their own predictions to the entries of miRCarta to reduce potential redundancies in their studies. miRCarta provides the most comprehensive collection of human miRNAs and miRNA candidates to form a basis for further refinement and validation studies. The database is freely accessible at https://mircarta.cs.uni-saarland.de/.

INTRODUCTION

MicroRNAs (miRNAs) play a central role in post-transcriptional gene regulation. This class of short noncoding RNAs with an average length of 17–23 nucleotides can bind to their complementary target mRNAs and repress their translation or mediate their degradation (1–3). Since one miRNA potentially regulates many genes and may therefore severely influence the overall regulation network, their expression changes have been the focus of many publications describing various diseases (4–8) and are discussed as potential biomarkers (9–13).

The central repository for miRNAs is the miRBase database (14), which is currently at its 21st version (released 06/14). The last update of miRBase has been >3 years ago. This is problematic for several reasons. Firstly, many miRNA prediction algorithms have been developed and applied to next-generation sequencing (NGS) data in recent years. The published results of these predictions often claim to have found hundreds or thousands of new miRNA candidates (15–18). Since these candidates were so far not integrated in miRBase, different studies contain substantial redundancies. Secondly, several independent groups have found that the current version of miRBase seems to already contain artifacts, wrongly annotated and false positive miRNAs, probably due to the integration of predicted candidates that were not experimentally validated (16,19–21). To overcome this, the miRBase provides a high-confidence miRNA set defined by lower thresholds of reads that must be mapping to the mature sequences besides other rules. In their latest publication, they collated 305 deep sequencing data sets from 38 species to annotate these high-confidence miRNA sets (14). In the meantime, the publicly available small RNA sequencing data has increased exponentially and should be used to further refine a current high-confidence miRNA set. In addition, even some validated miRNAs have not yet made their way into miRBase.

With miRCarta, we aimed to develop a database to bridge the gap between the already available annotations in miRBase and the more recent miRNA predictions from publications or our tool miRMaster (22). To this end, we initially integrated the content of miRBase releases 1.1-21 and enhanced our database with new analysis tools, annota-

*To whom correspondence should be addressed. Tel: +49 681 302 68607; Email: c.backes@mx.uni-saarland.de
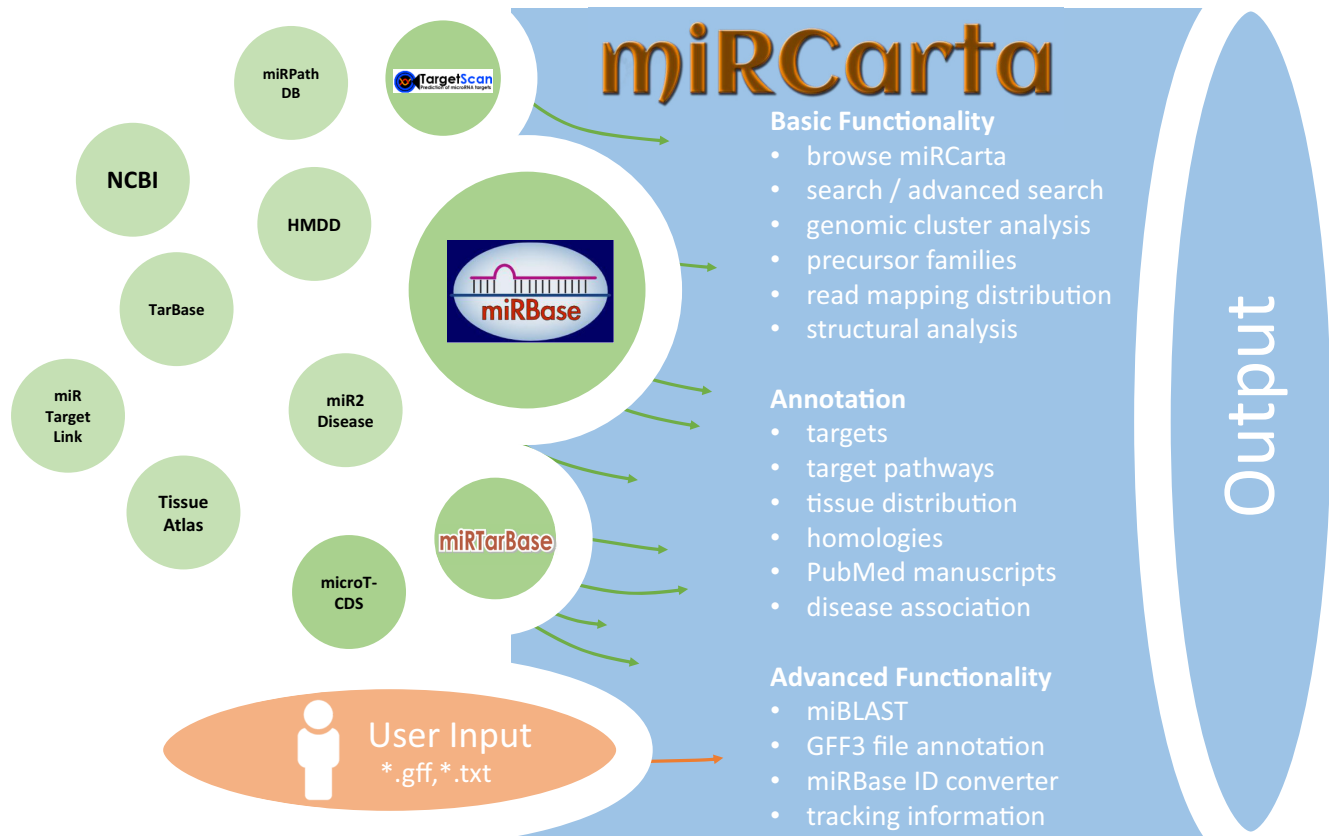
Figure 1. Overview of the integrated or linked data sources and the functionality of miRCarta.

tions, and background information on miRNAs. This part of miRCarta can be used as if querying miRBase and is independent of the remaining updated database content. In a next step, we retrieved updated genomes for 148 organisms that had miRNA annotations in miRBase and remapped the miRNAs to get up-to-date locations for these organisms. We put our focus on the most frequently studied organism in biomedical research, namely *Homo sapiens*. For human, we collected over 18 000 small RNA sequencing data sets from the Sequence Read Archive (SRA) (23), The Cancer Genome Atlas (TCGA) (24) and in-house data sets. These data were processed with our tool miR-Master to predict novel miRNA candidates. To these predictions, we added miRNA candidates from publications (15,16) and miRBase resulting in a total of 24 148 human mature miRNA candidates. To facilitate the decision if the integrated candidates are potentially real miRNAs, we visualize the expression profiles along the corresponding precursors using the mapping results of the 18 035 samples against the stem loop sequences. This way, researchers are able to select conveniently promising candidates for further experimental validation. In addition, we provide a batch query for researchers to match their own miRNA predictions to the entries of miRCarta to reduce potential redundancies in their studies.

This first release of miRCarta provides the most comprehensive collection of human miRNA candidates to date and can serve as an entry point for researchers working in this field searching for current miRNA annotations and predictions. miRCarta is freely accessible at: https://mircarta.cs.uni-saarland.de/.

DATA SOURCES

miRCarta was conceived to provide the information of miRBase (14) as well as more recent data stemming from miRNA predictions. To this end, we integrated the content of miRBase releases 1.1-21, including all naming and sequence changes that the entries of miRBase underwent, as well as the location information for miRNAs in the latest miRBase release comprising 108 organisms. We enhanced this basic information by adding additional data sources and links to external databases. To be able to also filter for miRNA targets, we integrated miRNA target predictions from microT-CDS v5.0 (25) and TargetScan v7.1 (26) and the experimentally validated targets from miRTarBase v6.1 (27). For miRBase precursors, we added links to the Human microRNA Disease Database (HMDD) (28) and to NCBI Gene if official gene symbols were available. miR-Base miRNAs are linked to the miRNA pathway dictionary (miRPathDB) (29), miRTargetLink Human (30), Tissue Atlas (31), miR2Disease (32) and TarBase (33). Since
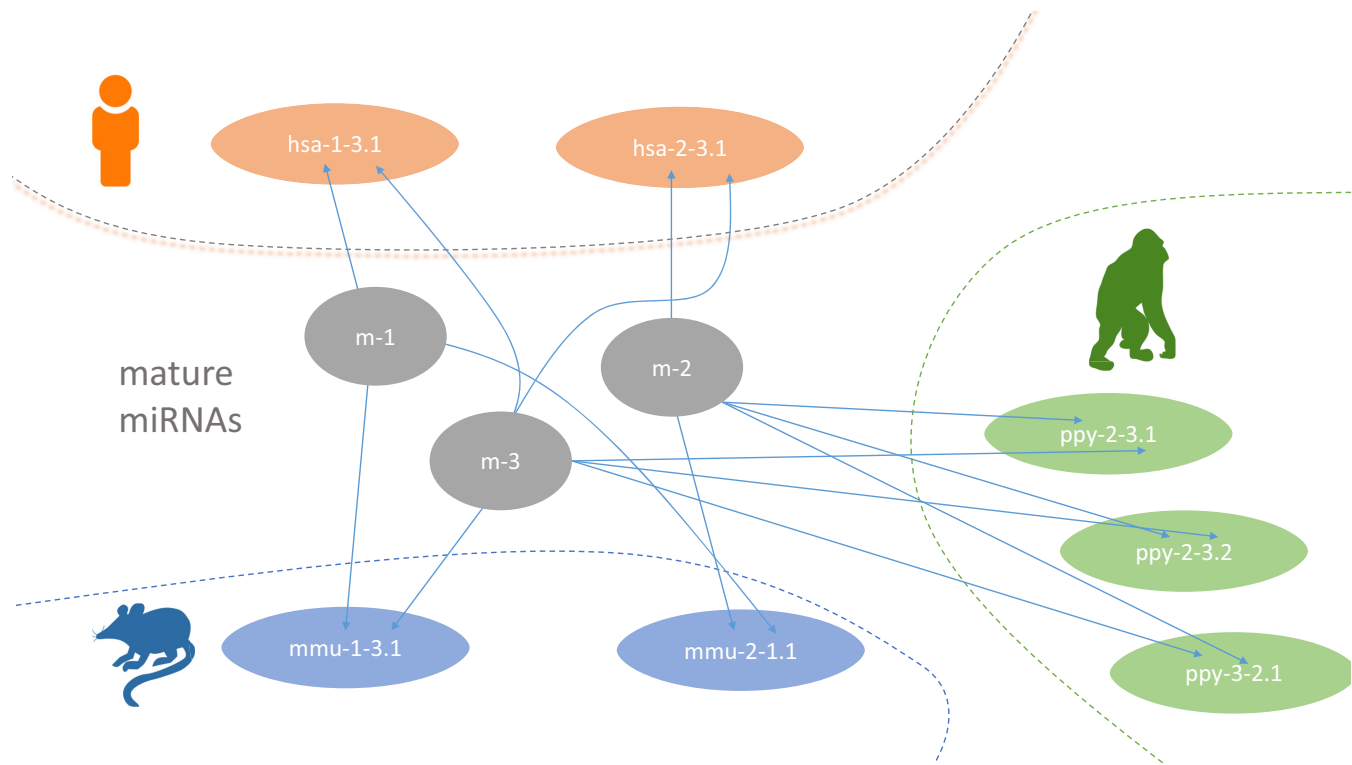
Figure 2. Figurative example for the new naming scheme in miRCarta. MiRNAs are named with m-[number] and are organism unspecific. Precursors are named [organism_abbreviation]-[5p miRNA]-[3p miRNA].[location ID]. In this example, we have a human precursor hsa-1-3.1, consisting of miRNAs m-1 and m-3. If this precursor has another location in the genome it gets another location ID as exemplified for ppy-2-3.1 and ppy-2-3.2.

the downloaded and integrated data of miRTarBase and microT-CDS are slightly different from their online versions, links to their web sites were also added. Obviously, all miRBase entries integrated in miRCarta are also connected to their original source in miRBase. An overview of miRCarta's data sources, external links and functionality is illustrated in Figure 1.

RE-ANNOTATING miRNAs/PRECURSORS AND NEW NAMING SCHEME

We collected the newest genome releases from NCBI RefSeq/Genbank (34) for 148 organisms that had miRNA/precursor annotations in miRBase. In brief, the precursors from miRBase for these 148 organisms were mapped with Bowtie 1.1.2 (35) against their respective genomes. Since a central aim of miRCarta was to include new potential candidates, a new naming scheme for miR-NAs and precursors had to be created (Figure 2). Mature miRNAs in miRCarta are named with m-[number] and are organism unspecific. Precursors are however organism specific, starting with a three letter code for an organism, followed by the number of the 5′ miRNA and the number of the 3′ miRNA and ending with a locus identifier, e.g. hsa-1-52.1 consists of mature miRNAs m-1 and m-52. To improve the miRNA annotations in miRBase, we collected for human 18 035 small RNA sequencing samples from SRA (23), TCGA (24) and in-house data sets. We mapped the reads against the human precursor sequences and derived the sequence of the canonical forms from the

expression profiles. Thereby, we processed the miRNAs in their median RPMMM (reads per million mapped to miRNAs) expression order across all our samples, resulting in the most expressed miRNA as m-1 (corresponding to hsa-miR-21-5p) and so on. Using these sequences as basis, we added predictions from publications (15,16) and our tool miRMaster (22) for the 18 035 samples to this pool, as well as the re-mapped sequences for the remaining 147 organisms to complete the information added to miRCarta. More details on this integration and naming process can be found in the Supplemental Material.

A side effect of the re-annotation is that a miRNA in miRBase might not be identical to a miRNA in miRCarta anymore, e.g. it can be shifted to the left or right or have a different length. Still, we deemed this re-annotation necessary since our analyses showed that the currently annotated canonical form represented only in 42% of cases the actually most expressed form across all of our samples. In the web interface we provide links between these miRCarta precursors/miRNAs and miRBase precursors/miRNAs to allow for an easier comparison.

DATABASE IMPLEMENTATION AND FUNCTIONALITY

Implementation

miRCarta consists of a MySQL database and a MongoDB NoSQL database. The MySQL part contains organisms, sequences, locations, miRNAs, precursors and targets, while the NoSQL database stores the expression data as matrix

Figure 3. Example of a precursor view for a predicted candidate in miRCarta. First, we list several basic facts about the precursor like its sequence, location, links to miRNAs, etc. In addition, we visualize the stem loop structure with the FornaContainer plugin (36) and color the miRNAs in the same way as in the sequence of the precursor. Below the structure, we show the pileup plots for the normalized or raw read counts with plotly.js. The user can easily switch here between log and linear scale or even visualize only counts with zero or one mismatches. The button 'Show details' opens a new HTML page, where more information can be found on how many samples had reads for this precursor and graphics showing if we found these reads rather continuously in several experiments or only a few. The last part shows the genomic context of the current precursor in a window of ±10 kb. This way it can be easily assessed if there are more precursors in this range or if the precursor lies in a gene or close to a gene for example. The genomic context is also interactive and allows for zooming in and out, and shows more information when clicking on a gene or miRNA etc.

Figure 4. Excerpt of the results of uploading a GFF3 file for the predictions of Friedländer *et al.* (40). We find overlaps for 1461 of 4934 uploaded precursors/miRNAs in miRCarta. The first four rows in the table show examples for entries we did not find in miRCarta and the genomic context view shows that there are also no other miRNAs in a window of ±10 kb around the annotated location. The fifth row shows an entry where we have overlaps in miRCarta and the genomic context view illustrates that there are many other miRNA annotations in range.

format for a more efficient access. The MySQL database schema is illustrated in Supplemental Figure S1. The server-side backend of the web application uses Django 1.11 and is written in Python 3. The web interface is implemented in Django's HTML template language and is enhanced with several JavaScript libraries for a more interactive user experience. The tables we visualize are created with the jQuery plugin DataTables, the genomic context visualization is done using TnT Genome, and the expression profile plots are rendered with plotly.js, the structure visualization with FornaContainer (36). For styling we use Bootstrap 3 and custom CSS files.

Functionality

miRCarta integrates the miRBase database and additional new miRNA candidates, expression data, updated organisms, and genomes as entry point for miRNA researchers. As illustrated in Figure 1, miRCarta provides different levels of functionality.

Basic functionality

Browse. The classical entry point 'Browse' lists all miR-NAs and precursors for a selected organism. For precursors, this view also visualizes the normalized read counts of the mapped NGS data without and with mismatches. This way a user can assess if the expression profile over a (putative or known) precursor seems likely for miRNA expression and

more rapidly identify real precursors/miRNAs from false positive annotations.

Advanced search. Using 'Advanced Search', users can restrict their query results to certain miRNAs and/or precursors of certain organisms that might have been validated with a certain experiment and so on. The results are visualized as HTML table, unless one of the download options is checked.

Precursor families. For the miRBase content, we also integrated the precursor families. A user can search for precursor names or miRBase accession numbers or select an organism and get all precursor families of the input as result. If nothing is selected all precursor families will be listed.

Genomic clusters. 'Genomic Clusters' visualizes clusters of precursors within a selectable window size in a tabular format and as stacked bar plots along the chromosomes. In Supplemental Figure S2, we queried the miRBase part for clusters in *Homo sapiens*. The stacked bar plot visualization directly shows that the largest clusters can be found on chromosomes 14 and 19.

Read mapping distribution. For human precursors, we visualize the mappings of the sequencing reads of the collected 18 035 samples with and without mismatch. The pileup plots can be found in the single precursor views (Figure 3) as well as in the tabular overviews of 'Browse' and 'Advanced Search' for *H. sapiens*. In the precursor view the pileup can be switched between normalized and raw read counts, log and linear scale, and also visualize perfect matching reads and reads with one mismatch separately or combined. More information about the number of mapped reads and the corresponding number of different samples for the precursor can be found on a separate HTML page by clicking on 'Show details' below the plot.

Structural analysis. The secondary structures for the precursor sequences are computed with RNAfold (37) and visualized with FornaContainer (36). This illustration is available in the precursor specific views.

Annotation

Targets. Since we integrated miRTarBase, microT-CDS, and TargetScan, miRCarta can provide a combined search of miRNAs and targets using experimentally validated or predicted targets, respectively. If all three databases are selected, the resulting table will contain for each database a column with either 0 or 1 as entry, which can be used for sorting and filtering for results that have e.g. only hits in all three target databases.

Target pathways. For potential target pathways, we linked the tools MiRTargetLink (30) and miRPathDB (29) for miRNAs. The links can be found on the right hand side of the miRNA views if they are available.

Tissue distribution. For miRNAs we also provide links to the tool TissueAtlas (31), which shows the miRNA abundance in 61 tissue biopsies of two individuals.

Homologies. We mapped the miRCarta miRNAs with their respective flanks (see Supplemental Material) against all 148 organisms without mismatch. If such a miRNA sequence is found in an organism where it has not been annotated so far, we list these findings under 'miRNA homologies' on the right hand side of the miRNA view.

PubMed manuscripts. We provide links to the manuscripts describing miRNAs/precursors in miRBase, as well as validation experiments for the targets in miRTarbase.

Disease association. For disease association, we provide links for miRNAs to miR2Disease (32) and to HMDD (28) for precursors.

Advanced functionality

miBLAST. Under 'miBLAST', users can enter a miRNA sequence and get the BLAST (38) results for miRBase and miRCarta miRNAs.

GFF3 file annotation. Using the analysis tool 'GFF3 upload' users can upload their own standard GFF3 files containing e.g. the locations of predictions of miRNAs and precursors for a certain organism. The data is matched against the available miRCarta entries and the result is visualized as a table, which shows how many findings are new or have overlaps with entries in miRCarta.

miRBase ID converter. The naming of miRNAs changed during different releases of miRBase, which can cause problems when comparing findings to older manuscripts where a different miRBase release was used. With the 'Identifier Conversion' tool researchers can convert their miRBase names into the latest available version.

Tracking Information. Inspired by the tool miRBase Tracker (39), we provide tracking information for each identifier in miRBase, which allows in a straightforward way to illustrate the changes a miRBase name or sequence underwent during different miRBase releases.

Application examples

To demonstrate the 'GFF3 Upload' functionality, we collected the predictions from Friedländer *et al.* (40) as independent test set, converted the locations with liftOver into GRCh38 coordinates and created a gff3 file. This file was uploaded in miRCarta using the default parameters. In Figure 4, we can see the first five entries of which four have not been found in miRCarta. The fifth has an overlap with two entries in miRCarta. The genomic context visualization is especially helpful if no overlap has been found to assess whether other miRNA precursors might be in range. Altogether, 1461 of 4934 entries from the Friedländer dataset have already been annotated in miRCarta.

miRCarta's precursor view enables users to grasp the structure and expression profiles more easily than it was possible using miRBase. In Supplemental Figure 3 we visualize the precursor hsa-mir-5739 in miRBase on the left-hand side and in miRCarta on the right-hand side. The

structure in miRBase is visualized as ASCII code and it is hard to assess for this precursor if this is a good structure or not. In the miRCarta view, it is directly clear that this is not a valid precursor by looking at the folding structure and the expression profile below.

Since we visualize the expression profiles also in 'Browse' and 'Advanced Search' for precursors in *H. sapiens*, these plots can also be easily used to scroll through longer lists of precursors and select interesting candidates for further validations.

FUTURE WORK

While still undiscovered miRNAs may exist, the current collection of miRCarta represents a substantial part of the human miRNome. This set—tailored to be a very sensitive collection of miRNAs—contains certainly a large number of false positive predictions. Nevertheless, this set will be useful for other researchers to match their own predictions against it to reduce redundancies in their studies. This current 'high-sensitivity' set will form the basis for our further developments. In a next step, we will reduce the 'high-sensitivity' set by merging similar findings, e.g. overlapping precursors from different predictions. This will result in a collapsed set with slightly lesser sensitivity than the original set. At last, we will create a 'high-specificity' set which will consist of experimentally validated miRNAs relying on cloning the precursor and providing evidence for the mature miRNAs using Northern Blots. In addition, we will also annotate miRNA isoforms and include more information about other organisms.

CONCLUSION

miRCarta bridges the gap between established annotations in miRBase and more recent miRNA predictions from publications and our software miRMaster. In our proof-of-concept study for human we succeeded to demonstrate that these candidates—which certainly contain false positive hits—contain interesting candidates for further validation. With this approach, we aim to create a high-resolution map of potential human miRNAs, which will be refined in further releases to create finally a set of experimentally validated real miRNAs.

DATA AVAILABILITY

miRCarta is publicly accessible at https://mircarta.cs.uni-saarland.de/.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

REFERENCES

1. Bartel,D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.
2. Sontheimer,E.J. and Carthew,R.W. (2005) Silence from within: endogenous siRNAs and miRNAs. *Cell*, **122**, 9–12.
3. Filipowicz,W., Jaskiewicz,L., Kolb,F.A. and Pillai,R.S. (2005) Post-transcriptional gene silencing by siRNAs and miRNAs. *Curr. Opin. Struct. Biol.*, **15**, 331–341.
4. Li,Y.J., Ping,C., Tang,J. and Zhang,W. (2016) MicroRNA-455 suppresses non-small cell lung cancer through targeting ZEB1. *Cell Biol. Int.*, **40**, 621–628.
5. Keller,A., Leidinger,P., Steinmeyer,F., Stahler,C., Franke,A., Hemmrich-Stanisak,G., Kappel,A., Wright,I., Dorr,J., Paul,F. *et al.* (2014) Comprehensive analysis of microRNA profiles in multiple sclerosis including next-generation sequencing. *Mult. Scler.*, **20**, 295–303.
6. Keller,A., Leidinger,P., Vogel,B., Backes,C., ElSharawy,A., Galata,V., Müller,S., Marquart,S., Schrauder,M., Strick,R. *et al.* (2014) miRNAs can be generally associated with human pathologies as exemplified for miR-144*. *BMC Med.*, **12**, 224.
7. Fenoglio,C., Ridolfi,E., Cantoni,C., De Riz,M., Bonsi,R., Serpente,M., Villa,C., Pietroboni,A.M., Naismith,R.T., Alvarez,E. *et al.* (2013) Decreased circulating miRNA levels in patients with primary progressive multiple sclerosis. *Mult. Scler.*, **19**, 1938–1942.
8. Keller,A., Leidinger,P., Bauer,A., Elsharawy,A., Haas,J., Backes,C., Wendschlag,A., Giese,N., Tjaden,C., Ott,K. *et al.* (2011) Toward the blood-borne miRNome of human diseases. *Nat. Methods*, **8**, 841–843.
9. Roth,P., Keller,A., Hoheisel,J.D., Codo,P., Bauer,A.S., Backes,C., Leidinger,P., Meese,E., Thiel,E., Korfel,A. *et al.* (2015) Differentially regulated miRNAs as prognostic biomarkers in the blood of primary CNS lymphoma patients. *Eur. J. Cancer*, **51**, 382–390.
10. Ma,J., Lin,Y., Zhan,M., Mann,D.L., Stass,S.A. and Jiang,F. (2015) Differential miRNA expressions in peripheral blood mononuclear cells for diagnosis of lung cancer. *Lab. Invest.*, **95**, 1197–1206.
11. Sayed,A.S., Xia,K., Yang,T.L. and Peng,J. (2013) Circulating microRNAs: a potential role in diagnosis and prognosis of acute myocardial infarction. *Dis. Markers*, **35**, 561–566.
12. Keller,A., Backes,C., Haas,J., Leidinger,P., Maetzler,W., Deuschle,C., Berg,D., Ruschil,C., Galata,V., Ruprecht,K. *et al.* (2016) Validating Alzheimer's disease micro RNAs using next-generation sequencing. *Alzheimers Dement.*, **12**, 565–576.
13. Margue,C., Reinsbach,S., Philippidou,D., Beaume,N., Walters,C., Schneider,J.G., Nashan,D., Behrmann,I. and Kreis,S. (2015) Comparison of a healthy miRNome with melanoma patient miRNomes: are microRNAs suitable serum biomarkers for cancer? *Oncotarget*, **6**, 12110–12127.
14. Kozomara,A. and Griffiths-Jones,S. (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.*, **42**, D68–D73.
15. Londin,E., Loher,P., Telonis,A.G., Quann,K., Clark,P., Jing,Y., Hatzimichael,E., Kirino,Y., Honda,S., Lally,M. *et al.* (2015) Analysis of 13 cell types reveals evidence for the expression of numerous novel primate- and tissue-specific microRNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E1106–E1115.
16. Backes,C., Meder,B., Hart,M., Ludwig,N., Leidinger,P., Vogel,B., Galata,V., Roth,P., Menegatti,J., Grasser,F. *et al.* (2016) Prioritizing and selecting likely novel miRNAs from NGS data. *Nucleic Acids Res.*, **44**, e53.
17. Friedländer,M.R., Mackowiak,S.D., Li,N., Chen,W. and Rajewsky,N. (2011) miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.*, **40**, 37–52.
18. Jha,A., Panzade,G., Pandey,R. and Shankar,R. (2015) A legion of potential regulatory sRNAs exists beyond the typical microRNAs microcosm. *Nucleic Acids Res.*, **43**, 8713–8724.
19. Vitsios,D.M., Davis,M.P., van Dongen,S. and Enright,A.J. (2016) Large-scale analysis of microRNA expression, epi-transcriptomic features and biogenesis. *Nucleic Acids Res.*, **45**, 1079–1090.
20. Fromm,B., Billipp,T., Peck,L.E., Johansen,M., Tarver,J.E., King,B.L., Newcomb,J.M., Sempere,L.F., Flatmark,K., Hovig,E. *et al.* (2015) A uniform system for the annotation of vertebrate microRNA genes and the evolution of the human microRNAome. *Annu. Rev. Genet.*, **49**, 213–242.

21. Ludwig,N., Becker,M., Schumann,T., Speer,T., Fehlmann,T., Keller,A. and Meese,E. (2017) Bias in recent miRBase annotations potentially associated with RNA quality issues. *Sci. Rep.*, **7**, 5162.
22. Fehlmann,T., Backes,C., Kahraman,M., Haas,J., Ludwig,N., Posch,A., Würstle,M., Hübenthal,M., Franke,A., Meder,B. *et al.* (2017) Web-based NGS data analysis using miRMaster: a large-scale meta-analysis of human miRNAs. *Nucleic Acids Res.*, **45**, 8731–8744.
23. Kodama,Y., Shumway,M., Leinonen,R. and International Nucleotide Sequence Database, C. (2012) The Sequence Read Archive: explosive growth of sequencing data. *Nucleic Acids Res.*, **40**, D54–D56.
24. Cancer Genome Atlas Research Network (2008) Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature*, **455**, 1061–1068.
25. Paraskevopoulou,M.D., Georgakilas,G., Kostoulas,N., Vlachos,I.S., Vergoulis,T., Reczko,M., Filippidis,C., Dalamagas,T. and Hatzigeorgiou,A.G. (2013) DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. *Nucleic Acids Res.*, **41**, W169–W173.
26. Agarwal,V., Bell,G.W., Nam,J.W. and Bartel,D.P. (2015) Predicting effective microRNA target sites in mammalian mRNAs. *Elife*, **4**, doi:10.7554/eLife.05005.
27. Chou,C.H., Chang,N.W., Shrestha,S., Hsu,S.D., Lin,Y.L., Lee,W.H., Yang,C.D., Hong,H.C., Wei,T.Y., Tu,S.J. *et al.* (2016) miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Res.*, **44**, D239–D247.
28. Li,Y., Qiu,C., Tu,J., Geng,B., Yang,J., Jiang,T. and Cui,Q. (2014) HMDD v2.0: a database for experimentally supported human microRNA and disease associations. *Nucleic Acids Res.*, **42**, D1070–D1074.
29. Backes,C., Kehl,T., Stockel,D., Fehlmann,T., Schneider,L., Meese,E., Lenhof,H.P. and Keller,A. (2017) miRPathDB: a new dictionary on microRNAs and target pathways. *Nucleic Acids Res.*, **45**, D90–D96.
30. Hamberg,M., Backes,C., Fehlmann,T., Hart,M., Meder,B., Meese,E. and Keller,A. (2016) MiRTargetLink–miRNAs, genes and interaction networks. *Int. J. Mol. Sci.*, **17**, 564.
31. Ludwig,N., Leidinger,P., Becker,K., Backes,C., Fehlmann,T., Pallasch,C., Rheinheimer,S., Meder,B., Stahler,C., Meese,E. *et al.* (2016) Distribution of miRNA expression across human tissues. *Nucleic Acids Res.*, **44**, 3865–3877.
32. Jiang,Q., Wang,Y., Hao,Y., Juan,L., Teng,M., Zhang,X., Li,M., Wang,G. and Liu,Y. (2009) miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Res.*, **37**, D98–D104.
33. Vlachos,I.S., Paraskevopoulou,M.D., Karagkouni,D., Georgakilas,G., Vergoulis,T., Kanellos,I., Anastasopoulos,I.L., Maniou,S., Karathanou,K., Kalfakakou,D. *et al.* (2015) DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. *Nucleic Acids Res.*, **43**, D153–D159.
34. Coordinators,N.R. (2015) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **43**, D6–D17.
35. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
36. Kerpedjiev,P., Hammer,S. and Hofacker,I.L. (2015) Forna (force-directed RNA): Simple and effective online RNA secondary structure diagrams. *Bioinformatics*, **31**, 3377–3379.
37. Lorenz,R., Bernhart,S.H., Honer Zu Siederdissen,C., Tafer,H., Flamm,C., Stadler,P.F. and Hofacker,I.L. (2011) ViennaRNA Package 2.0. *Algorithms Mol. Biol.*, **6**, 26.
38. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
39. Van Peer,G., Lefever,S., Anckaert,J., Beckers,A., Rihani,A., Van Goethem,A., Volders,P.J., Zeka,F., Ongenaert,M., Mestdagh,P. *et al.* (2014) miRBase Tracker: keeping track of microRNA annotation changes. *Database (Oxford)*, **2014**, bau080.
40. Friedlander,M.R., Lizano,E., Houben,A.J., Bezdan,D., Banez-Coronel,M., Kudla,G., Mateu-Huertas,E., Kagerbauer,B., Gonzalez,J., Chen,K.C. *et al.* (2014) Evidence for the biogenesis of more than 1,000 novel human microRNAs. *Genome Biol.*, **15**, R57.