Bioinformatics, YYYY, 0–0
doi: 10.1093/bioinformatics/xxxxx
Advance Access Publication Date: DD Month YYYY
Manuscript Category

Subject Section

sRNAPrimerDB: Comprehensive primer design and search web service for small non-coding RNAs

Shengsong Xie[1, 4*], Qin Zhu[2*], Wubin Qu[3*], Zhong Xu[1], Xiangdong Liu[1, 4], Xinyun Li[1, 4], Shijun Li[1], Wubin Ma[5], Yiliang Miao[1, 4], Lisheng Zhang[1], Xiaoyong Du[2], Wuzi Dong[6], Haiwei Li[3], Changzhi Zhao[1], Yunlong Wang[2], Yaping Fang[2§] and Shuhong Zhao[1, 4§]

[1]Key Laboratory of Agricultural Animal Genetics, Breeding and Reproduction of Ministry of Education & Key Lab of Swine Genetics and Breeding of Ministry of Agriculture and Rural Affairs, Huazhong Agricultural University, Wuhan 430070, P. R. China; [2]Agricultural Bioinformatics Key Laboratory of Hubei Province, College of Informatics, Huazhong Agricultural University, Wuhan 430070, P. R. China; [3]iGeneTech Bioscience Co., Ltd., Beijing 102206, P. R. China; [4]The Cooperative Innovation Center for Sustainable Pig Production, Huazhong Agricultural University, Wuhan 430070, P. R. China; [5]Howard Hughes Medical Institute, Department of Medicine, School of Medicine, University of California, San Diego, La Jolla, CA 92093, USA; [6]College of Animal Science and Technology, Northwest A & F University, Yangling, 712100, P. R. China.

[§]To whom correspondences should be addressed.

[*]These authors contributed equally to this work.

Associate Editor: XXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Motivation:

Small non-coding RNAs (ncRNAs), especially microRNAs (miRNAs) and piwi-interacting RNAs (piRNAs), play key roles in many biological processes. However, only a few tools can be used to develop the optimal primer or probe design for the expression profile of small ncRNAs. Here, we developed sRNAPrimerDB, the first automated primer designing and query web service for small ncRNAs.

Results:

The primer online designing module of sRNAPrimerDB is composed of primer design algorithms and quality evaluation of the polymerase chain reaction (PCR) primer. Five types of primers, namely, generic or specific reverse transcription primers, specific PCR primers pairs, TaqMan probe, double hairpin probe, and hybridization probe for different small ncRNA detection methods, can be designed and searched using this service. The quality of PCR primers is further evaluated using melting temperature, primer dimer, hairpin structure, and specificity. Moreover, the sequence and size of each amplicon are also provided for the subsequent experiment verification. At present, 531,306 and 2,941,669 primer pairs exist across 223 species for miRNAs and piRNAs, respectively, according to sRNAPrimerDB. Several primers designed by sRNAPrimerDB are further successfully validated by subsequent experiments.

Availability: sRNAPrimerDB is a valuable platform that can be used to detect small ncRNAs. This module can be publicly accessible at http://www.srnaprimerdb.com or http://123.57.239.141.

Contact: ypfang@mail.hzau.edu.cn or shzhao@mail.hzau.edu.cn

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

More than 2,687 small non-coding RNA (ncRNA) families have been reported in the Rfam (13.0) database in the last decade (Kalvari *et al*., 2018). Among these families, microRNAs (miRNAs) and piwi-interacting RNAs (piRNAs) are the most extensively studied small ncRNAs that are important to gene expression, mRNA stability, protein translation, and chromosome modification (Lagos-Quintana *et al*., 2001,

Bartel, 2018, Shen *et al*., 2018, Zhang *et al*., 2018). Small ncRNAs are large in quantity. For instance, 35,828 mature miRNAs are deposited in miRBase (v21), and 100,639 piRNAs are deposited in the piRNABank database (Kozomara and Griffiths-Jones, 2014, Backes *et al*., 2018, Sai Lakshmi and Agrawal, 2008). Primer designing is a crucial factor to detect and quantify small ncRNAs (Rodriguez *et al*., 2015). To date, several strategies have been developed for small ncRNA detection; however, each of them has several limitations. For instance, the two widely used strategies for the reverse transcription (RT) of small ncRNAs involve either the addition of a poly(A) tail and priming with an anchored oligo(dT) RT primer or the usage of a specific stem-loop RT primer (Shi and Chiang, 2005, Chen *et al*., 2005). S-Poly(T) miRNA assay uses a specific oligo(dT) RT primer instead of universal RT adapter based on these previous approaches and to improve the sensitivity and specificity of small ncRNA detection assays (Kang *et al*., 2012). Poly(U) tailed and poly(A) stem-loop RT primers were used to perform RT reaction (Mei *et al*., 2012). In addition, one-step real time reverse transcription polymerase chain reaction (RT-PCR) was conducted to detect miRNAs (Yan *et al*., 2013). A chemical-based method that uses a two-hairpin probe to increase the specificity for miRNA detection was developed (Li *et al*., 2014). Two improved RT-qPCR-based methods, which were modified from S-Poly(T) miRNA assays, were developed to satisfy different experimental needs and increase the diversity of small ncRNA detection methods (Kang *et al*., 2012). These two methods adopted a poly(T) tailed stem-loop RT primer with or without specific nucleotides binding to the 3′-end of small ncRNAs instead of using a specific linear RT primer in this study.

Dimers and off-target amplicon of primers can cause signal dampening, false-negatives, and false-positives (Dellett and Simpson, 2016). Thus, the qualities of PCR primers, such as primer specificity, optimal annealing temperatures, dimers, and hairpin, should be evaluated. However, the current primer design options for amplification of small ncRNAs are severely limited because of the small sizes of the primers. Although three tools are available for miRNA primer design, only a specific type of primer is designed and available for these tools. For instance, "miRNA Primer Design Tool" (Czimmerer *et al*., 2013) can be only used to design primers for stem-loop RT-qPCR. miRprimer (Busk, 2014) and miPrimer (Kang *et al*., 2017) can only be used to design primers for "miR-specific RT-qPCR". Thus, a general tool to design various primers for detection of small ncRNAs must be developed. This software should design primers for small ncRNAs by considering the primer design and their quality evaluation with published primer designing methods.

In this study, a platform named sRNAPrimerDB with primer designing programs and calculated primers for small ncRNAs was developed. sRNAPrimerDB can be applied to design primers and probes for up to nine small ncRNA detection assays. In particular, sRNAPrimerDB can be freely accessed and is easily used to search and design primers for small ncRNAs, such as miRNAs and piRNAs. Several primers across designing methods in sRNAPrimerDB were selected and further confirmed by subsequent experiments. The detailed protocols confirmed by our experiments are freely available in sRNAPrimerDB. In summary, sRNAPrimerDB provides users a comprehensive platform to quickly design and search primers and probes, which can be helpful to accurately detect the expression of small ncRNAs.

2 Materials and Methods

2.1 Data sources

sRNAPrimerDB has collected 35,828 mature miRNAs from 223 organisms (including 115 metazoa, 73 viridiplantae, 29 viruses, 5 chromalveolata, and 1 mycetozoa). The sequences of all these miRNAs were downloaded from miRBase database (http://www.mirbase.org/) (Kozomara and Griffiths-Jones, 2014). The corresponding cDNA sequences of human, mouse, rat, pig, rabbit, cow, sheep, cat, zebra fish, Drosophila melanogaster, and chicken were downloaded from NCBI GenBank database (https://www.ncbi.nlm.nih.gov/genbank/). A total of 100,639 piRNAs were retrieved from piRNABank (http://pirnabank.ibab.ac.in/) (Sai Lakshmi and Agrawal, 2008). Pre-designed primers and probes of miRNA for RT-qPCR and Northern blotting assays are available for all the miRNAs of invertebrates, vertebrates, and plants. Pre-designed primers and probes of piRNAs were also designed.

2.2 Development of sRNAPrimerDB

Various detection methods can be used to assess the quantitative expression of small ncRNAs. Multiple small ncRNA detection methods, such as Linear poly(A) tailed RT-PCR (Shi and Chiang, 2005), Linear S-poly(A) tailed RT-PCR (Kang *et al*., 2012), Stem-loop RT-PCR (Chen *et al*., 2005), Stem-loop poly(U) tailed RT-PCR (Mei *et al*., 2012), One-step real time RT-PCR (Yan *et al*., 2013), Double-stem-loop LH-PCR (Li *et al*., 2014), and Northern blot or in situ hybridization (ISH) (Valoczi *et al*., 2004, Yin, 2018), are integrated into sRNAPrimerDB. By comprehensively analyzing different experimental methods, we improved the two optimized designs of RT primers. For example, we found that linear RT primers can be replaced with specific or universal stem-loop RT primers to further increase the specificity of Linear S-poly(A) tailed RT-PCR (Kang *et al*., 2012). Based on the characteristics of RT primers, these two methods were named Stem-loop poly(A) tailed RT-PCR and Stem-loop S-poly(A) tailed RT-PCR. The schematic of the principles for these small ncRNA detection assays is shown in Supplementary Figure S1. A detailed description of the experimental protocol of these methods can be found in the sRNAPrimerDB website. The characteristics of primer pairs and probe for each method are described and summarized in Supplementary Table S1. Primer design and quality assessment were developed for each method using the Perl language. The quality of designed PCR primers was assessed by using three third-party tools, namely, Ntthal (Untergasser *et al*., 2007), RNAstructure (Bellaousov *et al*., 2013), and MEFprimer-2.0 (Qu *et al*., 2012). sRNAPrimerDB provides the optimal candidate primer pairs and probes for each small ncRNA detection assay based on the results of the primer quality assessment.

2.3 Database implementation

sRNAPrimerDB was implemented using Golang (https://golang.org), mongodb (https://www.mongodb.com/), bootstrap3 (http://getbootstrap.com/), jquery (http://jquery.com/) technologies on Ubuntu 4.8.2 operating system. This module is suitable on Chromium, Internet Explorer versions, and Firefox browsers. The database is supported by the integration of mongodb and beego (https://github.com/astaxie/beego). sRNAPrimerDB is composed of three layers (Supplementary Figure S2). In the innermost layer, mongodb was used as the back-end database because of its flexibility and scalability. The pre-designed primers and probes are deposited in mongodb. The middle layer is the core layer, which links the database and interface layers. Open-source Go web framework beego was used in this layer. Beego implements the model-view-controller pattern that makes application exhibit loose coupling and high maintainability. Package gopkg.in/mgo.v2 was used to connect the mongo databases and sRNAPrimer stand-alone program for small ncRNA primer designing function. Bootstrap 3.3.7 was used to provide a friendly interface from both personal computer desktop and mobile devices. Jquery 2.1.1

(http://jquery.com/) was used to page the search results by retrieving primer information for query small ncRNAs and organisms. In addition, DataTables 1.10.15 (https://cdn.datatables.net/1.10.15/) were used to organize and present the layout of nine kinds of small ncRNA detection methods and plenty of specials that are ready to order primers and probe datasets in the download page.

2.4 miRNA mimics and transient transfection

HEK293T cell were purchased from the Cell Bank of Chinese Academy of Sciences (Shanghai, China) and were cultured in DMEM (GIBCO, USA) plus 10 % FBS. All cells were incubated at 37 °C with 5 % CO_2 atmosphere. miRNA mimics were transfected into HEK293T cell by using Lipofectamine RNAi MAX Transfection Reagent (Invitrogen, USA) according to the manufacturer's protocol, the concentration of miRNA mimic is 25, 50, 100 nmol/mL, respectively. The mimics of mmu-miR-10a-5p, mmu-miR-10b-5p and appropriate negative control molecules (miR-negative control, miR-NC), were synthesized by GenePharma (GenePharma, Shanghai, China). All sequences of miRNA mimics are listed in Supplementary Table S2.

2.5 Total RNA isolation and RT-PCR

Six different RT-PCR-based assays were performed to validate the primers designed by sRNAPrimerDB. The following tissues were collected: heart, liver, spleen, lung, kidney, skeletal muscle, and stomach from a 4-month-old Duroc boar. Total RNA was isolated from porcine tissues using the Trizol reagent (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions. The samples were then pooled, and the quality of total RNA was checked using gel electrophoresis with Umibio® 6 × RNA loading buffer (UR30402, Shanghai). Total RNA was then reverse transcribed to cDNA using RT primer (i.e., Linear poly(T) RT primers, Linear S-poly(T) RT primers, stem-loop RT primers, stem-loop poly(A) RT primers, stem-loop poly(T) RT primers, Stem-loop S-poly(T) RT primers) and RNase Inhibitor, ThermoScript™ Reverse Transcriptase (Invitrogen). The reverse transcriptase reaction for six different RT-PCR assays using the corresponding primers was performed according to the protocol of previous studies (Shi and Chiang, 2005; Kang *et al*., 2012; Chen *et al*., 2005; Mei *et al*., 2012). PCR was performed using *Taq* PCR Master Mix (Takara, Otsu, Japan) on the Applied Biosystems 2720 Thermal Cycler according to the manufacturer's instructions. The PCR product was subcloned into the pMD18-T vector by the TA-cloning and sequencing. All primers were listed in the Supplementary Table S2. Details on the experimental procedures are described in the sRNAPrimerDB website (http://www.srnaprimerdb.com).

2.6 Real-time PCR

The miRNA primer pairs for detection of mmu-miR-10a-5p were designed by sRNAPrimerDB, with 5s *rRNA* used as reference gene. All primers are listed in Supplementary Table S2. The reverse transcription reaction was performed using stem-loop RT primers according to Chen et al (2005). Real-time PCR was performed using a standard SYBR Green PCR kit (Toyobo, Japan) in the BioRad iQ5 Real-Time PCR Detection System according to the manufacturer's instructions.

2.7 Northern blotting

Probes for mmu-miR-10a-5p, mmu-miR-10b-5p and mmu-miR-200a were designed by sRNAPrimerDB. All probes are listed in Supplementary Table S2. The following tissues were collected: the heart, liver, spleen, lung, kidney, muscle, stomach, small intestine, brain, fat, thymus, testis, and epididymis from 7-week-old C57BL/6 mice. Total RNA was isolated from HEK293T cells and mouse tissues using the Trizol reagent (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions. The quality of total RNA was checked using gel electrophoresis with Umibio® 6 × RNA loading buffer (UR30402). The Northern blotting reaction was performed according to the protocol of previous study (Ma *et al*., 2013). The blots were visualized in a FLA 9000 biomolecular imager and analyzed with the Multi Gauge software (Fujifilm, Tokyo, Japan). Ethidium bromide (EtBr) stain total RNA served as the loading control.

3 Results

3.1 sRNAPrimerDB algorithm and parameter design

sRNAPrimerDB is a comprehensive web service for a variety of small ncRNA detection assays. sRNAPrimerDB is mainly composed of two algorithms that design primers and probes. The first one is a primer design algorithm, and the second one is the quality evaluation of the PCR primers. In addition, sRNAPrimerDB involves searching, analyzing, uploading, and downloading of primers. The workflow of sRNAPrimerDB program is shown in Figure 1. Up to nine small ncRNA detection methods, which can be divided into RT-qPCR-based and hybridization-based strategies, were integrated into sRNAPrimerDB.

(1) Algorithms and parameters of primer and probe design for RT-qPCR-based methods

Linear poly(A) tailed RT-PCR, Linear S-poly(A) tailed RT-PCR, Stem-loop RT-PCR, Stem-loop poly(U) tailed RT-PCR, Stem-loop poly(A) tailed RT-PCR, and Stem-loop S-poly(A) tailed RT-PCR must design RT primers and PCR primer pairs to perform RT-qPCR-based assays (Supplementary Figure S3). The general design process in sRNAPrimerDB is as follows. First, RT primer with linear or Stem-loop secondary structure was designed for a given small ncRNA. Therefore, the "RT primer" was set as a parameter in sRNAPrimerDB, and users can use their own RT sequence or select the default sequence. RT primers can be divided into two categories: (1) generic or universal RT primer, such as linear poly(T) RT primer, stem-loop poly(A) RT primer, and stem-loop poly(T) RT primer (Supplementary Figure S4). These RT primers contain oligo(dT) or oligo(dA) and a universal reverse PCR primer anchor sequences, which can anneal to the poly-adenylated (poly A) or poly-uridyl (poly U) 3′-end tails RNA; (2) specific RT primer, such as linear S-poly(T) RT primer, stem-loop RT primer, and stem-loop S-poly(T) RT primer; the underlined letter "S" means specific (Supplementary Figure S4). The length of the poly(T) tailed RT primer, which contains specific nucleotides bound to the 3′-end of small ncRNAs, was set as a parameter. This parameter was named as "specific length", and the default value was set to 6 nts. Second, PCR primers were designed, and primer quality was assessed. Reverse PCR primer was calculated based on the complementary sequence of the RT primer (Supplementary Figure S5). The *T*m value for forward PCR primer was named as "Optimal *T*m", which aligned to the 5′ end of the target small ncRNA sequence. The default parameter value was set at 60 °C. In addition, to design the optimal primers, PCR primers below a certain *T*m were discarded. The default value of "*T*m cutoff" was set at 30 °C. sRNAPrimerDB extended the 5′ end of the forward PCR primer with 6-8 nucleotides (also named 5′ tag) to increase the *T*m value. These sequences were selected from predesigned 5′ tag libraries using a custom Perl script for generation of random DNA sequences (Supplementary Figure S5A). At least two nucleotide intervals (named "Gap Length") were found between the forward PCR and RT primers in a target to prevent the forward PCR primer from competing with the target of the RT primer (Supplementary Figure S5B). The *T*m values of the forward and reverse PCR primer for a

given small ncRNA should match. The size of the PCR primers was set as a parameter named "Primer size". The minimum and maximum lengths of the PCR primers were 18 and 25 nts, respectively. The GC% content of the primers was also set as a parameter, and the default value ranged from 20% to 80%.

A stretch of four or more continuous identical nucleotides is not allowed to avoid potential primer cross-reactivity due to low sequence complexity. Hairpin and cross-dimer were evaluated by ntthal and RNAstructure, respectively, to ensure the stable binding of primer/template. Ntthal provides sRNAPrimerDB alignment functionality based on nearest-neighbor thermodynamical approach. The specificity to a given small ncRNA is the most important property of a primer. A module called MEFprimer-2.0 was used to evaluate the specificity of a primer across the whole genomic/cDNA sequences. Third, TaqMan probes were designed. TaqMan probe is extensively applied in small ncRNA quantification assays. The amplification products can be detected by either specific or universal TaqMan probe. For specific TaqMan-based probe, the binding site was designed in the 3′-end of small ncRNA and RT primer stem region (Supplementary Figure S5C). sRNAPrimerDB shifted the universal TaqMan probe toward the conserved sequence of the RT primer and eliminated any overlap with the target sequence (Supplementary Figure S5D). The default length of TaqMan probe was set to 14-mer in sRNAPrimerDB. The designed result for each SYBR Green or TaqMan-based small ncRNA assay contained (1) RT primer; (2) specific forward PCR primers; (3) reverse PCR primer; (4) specific and/or universal TaqMan MGB probe (optional).

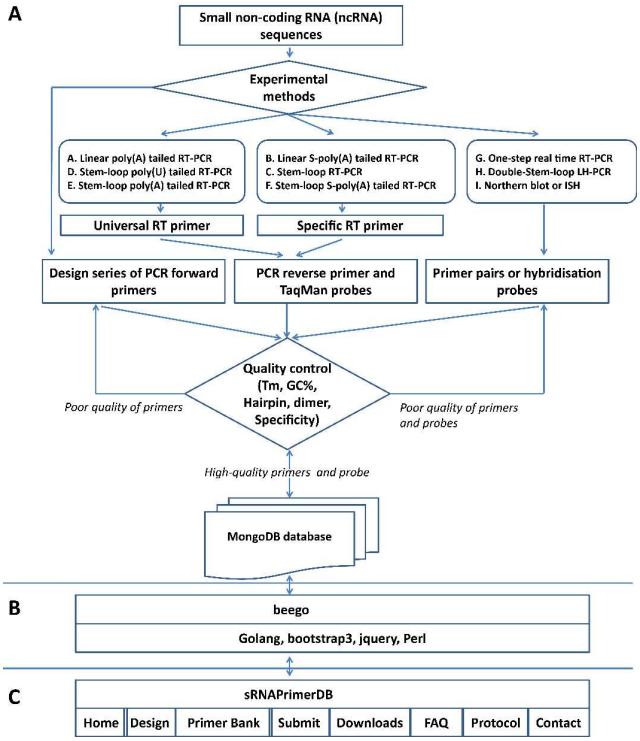

Fig. 1. Workflow for the development of sRNAPrimerDB. A. Workflow for the generation of small ncRNA-specific RT-qPCR primer pairs and probes. The best primer sequences are loaded into the MongoDB database. B. Implementation of sRNAPrimerDB through the integration of different programs. C. Organization of sRNAPrimerDB. ncRNA is the non-coding RNA, *T*m is the melting temperature, GC% is the guanine-cytosine content percentage, RT-PCR stands for reverse transcription polymerase chain reaction, and ISH stands for in situ hybridization.

(2) Algorithms of primer and probe design for hybridization-based methods

The features of primer and probe design for hybridization-based methods in sRNAPrimerDB are as follows. (1) One-step real time RT-PCR assay is performed, and sRNAPrimerDB designs four primers for the whole detection procedure, namely, RP1, RP2, P1, and P2 (Supplementary Figure S1G). P1 and P2 are primer pairs for PCR. RP1 contains the P1 and an 11-base sequence that is complimentary to the 3′-end of target small ncRNA. RP2 contains P2 and an 11-base sequence that is the same with the 5′-end of target small ncRNA. (2) Double-Stem-loop LH-PCR assay is performed. sRNAPrimerDB designs five primers and probes: double-hairpin probe (specific), forward primer (universal), reverse primer (universal), 5′ end for PCR bridging (universal), 3′ end for PCR bridging (specific) (Supplementary Figure S1H). (3) Northern blot or ISH assay is performed, and the module must only design anti-sense oligo probes targeting the small ncRNA of interest (Supplementary Figure S1I).

3.2 Database organization

sRNAPrimerDB comprises eight sections, including "Home", "Design", "Primer Bank", "Submit", "Downloads", "FAQ", "Protocol", and "Contact" functions (Figure 1). From the "Home" page, users can view the brief introduction, primer statistics, schematic strategy, and workflow for the development of sRNAPrimerDB. In the "Design" section, users can design primers and probes for nine small ncRNA detection methods (A to I). In the "Primer Bank" section, sRNAPrimerDB contains 531,306, 2,941,669 pre-designed primers and probes for 35,828 miRNAs (Supplementary Table S3) and 100,639 piRNAs (Supplementary Table S4), respectively. This database can be queried using either the small ncRNA name or symbol, or sequences. In the "Protocol" section, the detailed experimental procedures and related resources are provided for each detection method. On the "FAQ" page, frequently asked questions and answers are provided. In the "Downloads" section, sRNAPrimerDB offers a user-friendly interface to batch download all selected primer or probe sequences for the target organism and different experimental approaches. Users can submit their validated primers or probes for nine small ncRNA detection methods through "Submit" pages. Feedback information can also be submitted online. We appreciate users that submit their validated primer and probe sequence to sRNAPrimerDB to update this module regularly based on user feedback.

3.3 Web interface and usage

Design tool. Nine detection methods can be selected from the "design" drop-down menu to design primers and probes for a given small ncRNA (Figure 2A). The default parameters for each small ncRNA detection method are provided to facilitate the design of primers. Thus, users only need to select an interesting method and input the target sequences in FASTA format. The primer result can be obtained in a few click "run" steps (Figure 2A). For advanced users, the "Advance Parameters" was set for nine primer designed methods, such as "RT primer, Specific Length (bp), Gap Length (bp), Optimal *T*m (°C), *T*m cutoff (°C), Max difference *T*m (°C), Primer size (bp), and Primer GC (%)" (Figure 2B). The detailed explanations of such parameters are provided in the "FAQ" menu of sRNAPrimerDB. "Background DNA" parameter should be selected to evaluate the specificity of PCR primers. For a given small ncRNA, the result is given in one RT primer, one reverse PCR primer, five candidates forward PCR primers, and one TaqMan probe (Figure 2C). In addition, the nucleotide composition and the size of the amplicon sequences are displayed. The results can be downloaded using the "Download primers for small ncRNA" button (Figure 2C).

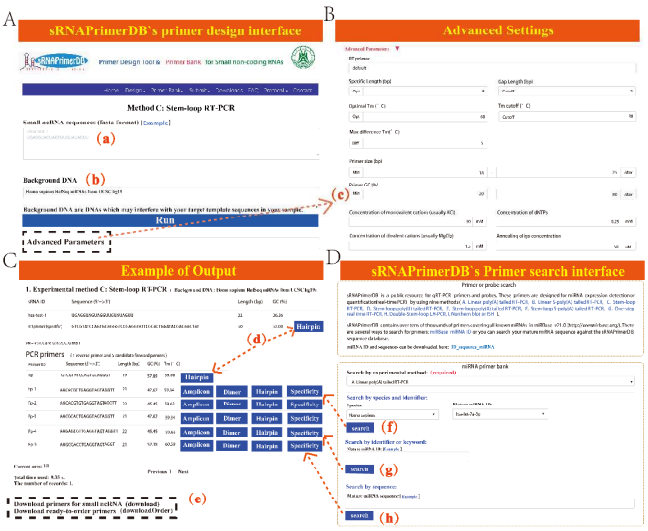

Fig. 2. Screenshots of the navigation bar, design and search module in sRNAPrimerDB. A. sRNAPrimerDB' s primer design interface. B. Advanced settings. C. Example of output. D. sRNAPrimerDB' s primer search interface. (a) Input box; (b) Reference genome; (c) Advanced parameters; (d) Primer or probe quality assessment; (e) Download hyperlinks for primer design results; (f) Search by species and identity number of small non-coding RNA; (g) Search by identity number of small non-coding RNA; (H) Search by sequence of small non-coding RNA.

Search tool. Three search boxes are provided in the "Primer Bank" menu. To search primers or probes for a given small ncRNA. These boxes contain three independent query strategies: (1) Search by experimental method; (2) Search by identifier or keyword; (3) Search by sequence (Figure 2D). Regardless of the kind of query strategy used, the detection method should be selected by first clicking the "Search by experimental method" button. For "Search by identifier or keyword" box as example, when using the search box, users should enter mature miRNA ID and then clicking the "search" button to search primer pairs. Considering the different identifiers and symbols for small ncRNAs across different species with the same sequence, we provided a "Regular expression" search server in the "Search by sequence" box. Therefore, users can perform "Regular expression" searches against multiple organisms. If the output format is set to table format, then the users can download the detailed information of primers in the "Search" page through a hyperlink under Subject heading.

Submit function. We provided the user with submit and feedback forms. sRNAPrimerDB requires users to upload verified primers according to correspondence detection methods because of the different primers used in the selected methods. Related experimental methods can be identified from the "Submit" drop-down menu in the navigation bar. To increase the running efficiency, we also welcome links to annotation files or published articles. Reliable primer pairs or probes validated by RT-qPCR or Northern blotting are marked in the database and are strongly recommended, whereas non-specific primer pairs or probes are removed in subsequent versions of sRNAPrimerDB. We also provided a "Comments or Questions" form in the "Submit" menu, wherein users can submit comments or suggestions about the content and structure of the database.

Batch download tool. In sRNAPrimerDB, all RT and qPCR primers, probes designed for each organism were deposited in the database in different experimental detection methods. Users can download these datasets for detection of miRNAs and piRNAs by click hyperlinks in the "Downloads" menu. More than 17 columns (for instance, RT primer, forward and reverse primer IDs, and sequence) are included in the downloaded file. Users can select multiple RT and qPCR primers for a given small ncRNA. These primers are directly used for oligo synthesis.

Protocol and FAQ. Guidelines for experimental procedures, sample handling, and data analysis are provided for each method in the sRNAPrimerDB. A user's manual (version 1.6) was also provided to describe how to use sRNAPrimerDB. A step-by-step guide for performing different assays can be found in the "Protocol" menu. The FAQs and corresponding answers can be viewed in the "FAQ" menu, including the introduction and usage of sRNAPrimerDB. This information will help users efficiently find information in the database.

3.4 Validation of primers designed by sRNAPrimerDB

Two miRNAs, mmu-miR-10a-5p and mmu-miR-10b-5p, which differ by only one base were selected to perform stem-loop RT-qPCR assay (Chen *et al*., 2015) to test the reliability of sRNAPrimerDB (Figure 3A). Total RNA from HEK293T cells which transfected with miRNA mimic were extracted and Northern blotting assays were performed first, the result showed that mature mmu-miR-10a-5p and mmu-miR-10b-5p can be detected after transient transfection of miRNA mimic (Supplementary Figure S6). Subsequently, stem-loop RT-qPCR for detecting mmu-miR-10a-5p was conducted using the same RNA sample as the Northern blotting assays. As shown in Figure 3B, the result showed that mmu-miR-10a-5p and mmu-miR-10b-5p can be accurately distinguished using primer pairs which designed by sRNAPrimerDB. The result also showed that stem-loop RT-qPCR method can be used to measurement of miRNA after transient transfection with miRNA mimic. To evaluate the quality of the probe through experiments, the tissue expression profile of mmu-miR-200a in mouse was detected by Northern blotting. As shown in Supplementary Figure S7, the precursor and mature of miR-200a can only be simultaneously detected in the lung, kidney, stomach, small intestine and epididymis. This result showed that miRNA detection probe designed by sRNAPrimerDB is ideal for sensitive and specific detection of miRNAs by Northern blotting assay.

Furthermore, a well-known high-expression miRNA ssc-let-7a from pig was selected to perform linear poly(A) tailed RT-PCR, linear S-poly(A) tailed RT-PCR, stem-loop RT-PCR, stem-loop poly(U) tailed RT-PCR, stem-loop poly(A) tailed RT-PCR and stem-loop S-poly(A) tailed RT-PCR assays (Supplementary Figures S1 A to F) to test the reliability of sRNAPrimerDB. Pooled total RNA from pig was reverse transcribed to cDNA, and PCR reaction was then performed. After approximately 25 cycles of DNA synthesis, the PCR product was detected by 3.5% gel electrophoresis. In the meantime, annealing temperature ranging from 55 °C to 63 °C in the PCR reaction was tested. As shown in Supplementary Figure S8, we found that all six RT-PCR-based assays can produce a specific ethidium bromide-stained PCR product. By contrast, the output of production varied under different annealing temperatures. Subsequently, the product of stem-loop RT-PCR for ssc-let-7a was validated by TA cloing and sequencing (Supplementary Figure S9). The result showed that the primers designed by sRNAPrimerDB are reliable.

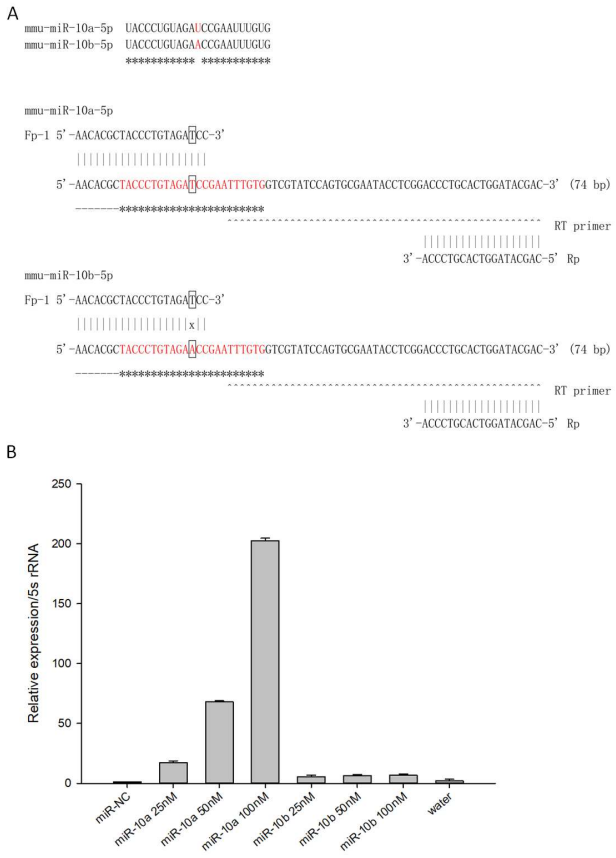

Fig. 3. Evaluation of the specificity of primer pairs for detecting highly similar miRNAs. A. Sequence alignment and schematic diagram of primer pair for miRNAs. B. Stem-loop RT-qPCR for mmu-miR-10a-5p. FP-1 stands for No.1 forward primer, Rp stands for reverse primer, RT stands for reverse transcription, miR-NC stands for miRNA negative control, water using as No-template-control (NTC) in real-time PCR, 5s *rRNA* using as reference gene. Mature miRNA sequences are labeled in red. Analysis of relative mmu-miR-10a-5p expression level using real-time qPCR and the $2^{-\Delta\Delta CT}$ Method. The asterisks (*) represent the highly conserved sequence and mature miRNA. The base in the box (□) indicates a base difference. bp stands for base pairs.

4 Discussion

4.1 Comparison with existing tools

Primer design is a crucial step in RT-qPCR experiment to target and amplify a known small ncRNA. To our knowledge, only three web and/or standalone primer designing tools, including "miRNA Primer Design Tool" (Czimmerer *et al.*, 2013), miRprimer (Busk, 2014), and miPrimer (Kang *et al.*, 2017), were developed for miRNA. Each primer designing tool can only be used for one type of small ncRNA. For instance, the miRprimer, which is developed in Ruby, is a stand-alone tool that can only be used to design primers for "miR-specific RT-qPCR". The specificity of miRprimer was increased by designing reverse primers with 3-8 nucleotides complementary to the miRNA. Candidate primers were adjusted at a *T*m value of 59 °C by trimming or addition of nucleotides. Optimum primers were selected to minimize the formation of secondary structures and primer dimers. "miRNA Primer Design Tool" is an online tool that requires login requirements and can be used to design Universal Probe Library (UPL) probes and specific RT primers for stem-loop RT-qPCR method. miPrimer, which is an empirical-based qPCR primer design method for "miR-specific RT-qPCR", was developed. Unfortunately, this tool is not currently available for the public.

The detailed comparison of the performance and features of sRNAPrimerDB, "miRNA Primer Design Tool", miRprimer and miPrimer are described and summarized in Supplementary Table S5. sRNAPrimerDB is a comprehensive primer and probe design and search web service for small ncRNA. The three features of sRNAPrimerDB (Supplementary Figure S10) are as follows. First, sRNAPrimerDB is a fast and easy to use server with five types of primers including: (a) generic or specific reverse transcription primers, (b) PCR forward and reverse primer pairs, (c) TaqMan probe, (d) double hairpin probe, and (e) hybridization oligos for various small ncRNA detection assays. In addition, the quality of specificity, dimers, and hairpins for each PCR primer are evaluated in sRNAPrimerDB. The specific sequence and length of the PCR amplicon are also provided for further PCR-based experiments. Second, sRNAPrimerDB contains a primer database providing the query, upload, and download for primers. A total of 531,306 and 2,941,669 pre-designed primers and probes were available for miRNAs and piRNAs, respectively. Users can easily share their experimental validated primers by uploading their validated primers to the sRNAPrimerDB. The best-hit primers of sRNAPrimerDB can reduce the failure rate of experiments and avoid misleading results in the amplification of products. Third, the validated experimental protocols for each method are provided in sRNAPrimerDB. Furthermore, RT and PCR primers, which are designed by sRNAPrimerDB, were successfully confirmed by experiment. In summary, the sRNAPrimerDB database provides a valuable time-saving resource and platform for the design, searching, and sharing of small ncRNA primers.

4.2 Intended user groups

Small ncRNAs regulate a wide range of developmental and physiological pathways in plants and animals (Ma *et al.*, 2013). Accurately detecting their expression is important to reveal the functions of small ncRNAs. To our knowledge, sRNAPrimerDB is the first comprehensive primer designing and search web service for small ncRNAs. sRNAPrimerDB contains a pre-designed RT, PCR primers, and probes for small ncRNAs across 223 organisms. sRNAPrimerDB is a valuable resource and can improve the efficient and specific performance of RT-qPCR or hybridization-based assays for small ncRNAs. Numerous studies have shown that miRNAs show potential as biomarkers for a wide range of medical and veterinary applications, including diagnostics, prognostics, nucleic-acid-based therapeutics, and theranostics (Perez-Sanchez *et al.*, 2018, Cardoso *et al.*, 2018). sRNAPrimerBD can assist in designing primers and probes to detect various miRNAs. Another feature of miRNAs is that they form families with similar sequences, making it difficult to design PCR assays with specificity to discriminate family members. Furthermore, individual miRNAs exist as a series of isoforms or "isomiRs", which vary in length and/or sequence and exhibit diagnostic potential (Nejad *et al.*, 2018). Here, sRNAPrimerDB was able to design primer pairs for many different small ncRNA detection methods, making it easy for users to choose the best way to assess isomiRs expression. We plan to update the database periodically and incorporate information for additional organisms and novel small ncRNAs in the future. Most importantly, small ncRNA-specific primers that have been experimentally validated will be highlighted in the database.

Funding

This work was supported by the National Transgenic Project of China [grant numbers 2016ZX08006003-004], the National Natural Science Foundation of China [grant numbers 31301226, 31501076, 31361140365], the National Key Research and Development Program of China, Stem Cell and Translational Research [grant numbers 2016YFA0100200] and the Fundamental Research Funds for the Central Universities [grant numbers 2662018JC002].

Conflict of Interest: none declared.

References

Backes, C., *et al.* (2018) miRCarta: a central repository for collecting miRNA candidates, *Nucleic Acids Res*, **46**, 160-167.

Bartel, D.P. (2018) Metazoan MicroRNAs, *Cell*, **173**, 20-51.

Busk, P.K. (2014) A tool for design of primers for microRNA-specific quantitative RT-qPCR, *BMC bioinformatics*, **15**, 29.

Bellaousov, S., *et al.* (2013) RNAstructure: Web servers for RNA secondary structure prediction and analysis, *Nucleic Acids Res*, **41**, 471-474.

Cardoso, A.L. and Guedes, J.R. (2018) Quantifying miRNA Deregulation in Alzheimer's Disease, *Methods in molecular biology*, **1750**, 307-319.

Chen, C., *et al.* (2005) Real-time quantification of microRNAs by stem-loop RT-PCR, *Nucleic Acids Res*, **33**, e179.

Czimmerer, Z., *et al.* (2013) A versatile method to design stem-loop primer-based quantitative PCR assays for detecting small regulatory RNA molecules, *PLoS One*, **8**, e55168.

Dellett, M. and Simpson, D.A. (2016) Considerations for optimization of microRNA PCR assays for molecular diagnosis, *Expert review of molecular diagnostics*, **16**, 407-414.

Kalvari, I., *et al.* (2018) Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families, *Nucleic Acids Res*, **46**, 335-342.

Kang, K., *et al.* (2012) A novel real-time PCR assay of microRNAs using S-Poly(T), a specific oligo(dT) reverse transcription primer with excellent sensitivity and specificity, *PLoS One*, **7**, e48536.

Kang, S., *et al*. (2018) miPrimer: An empirical-based qPCR primer design method for small non-coding microRNA. *Rna*, **24**, 304-312.

Kozomara, A. and Griffiths-Jones, S. (2014) miRBase: annotating high confidence microRNAs using deep sequencing data, *Nucleic Acids Res*, **42**, 68-73.

Lagos-Quintana, M., *et al.* (2001) Identification of novel genes coding for small expressed RNAs, *Science*, **294**, 853-858.

Li, X., *et al.* (2014) A convenient system for highly specific and sensitive detection of miRNA expression, *Rna*, **20**, 252-259.

Ma, W., *et al.* (2013) An androgen receptor-microrna-29a regulatory circuitry in mouse epididymis, *The Journal of biological chemistry*, **288**, 29369-29381.

Mei, Q., *et al.* (2012) A facile and specific assay for quantifying microRNA by an optimized RT-qPCR approach, *PLoS One*, **7**, e46890.

Nejad, C., *et al.* (2018) miR-222 isoforms are differentially regulated by type-I interferon, *Rna*, **24**, 332-341.

Perez-Sanchez, C., *et al.* (2018) Circulating microRNAs as biomarkers of disease and typification of the atherothrombotic status in Antiphospholipid Syndrome, *Haematologica*, **103**, 908-918.

Qu, W., *et al.* (2012) MFEprimer-2.0: a fast thermodynamics-based program for checking PCR primer specificity, *Nucleic Acids Res*, **40**, 205-208.

Rodriguez, A., *et al.* (2015) Design of primers and probes for quantitative real-time PCR methods, *Methods in molecular biology*, **1275**, 31-56.

Sai Lakshmi, S. and Agrawal, S. (2008) piRNABank: a web resource on classified and clustered Piwi-interacting RNAs, *Nucleic Acids Res*, **36**, 173-177.

Shen, E.Z., *et al.* (2018) Identification of piRNA Binding Sites Reveals the Argonaute Regulatory Landscape of the C. elegans Germline, *Cell*, **172**, 937-951.

Shi, R. and Chiang, V.L. (2005) Facile means for quantifying microRNA expression by real-time PCR, *BioTechniques*, **39**, 519-525.

Untergasser, A., *et al.* (2007) Primer3Plus, an enhanced web interface to Primer3, *Nucleic Acids Res*, **35**, 71-74.

Valoczi, A., *et al*. (2004) Sensitive and specific detection of microRNAs by northern blot analysis using LNA-modified oligonucleotide probes. *Nucleic Acids Res*, **32**, e175.

Yan, J., *et al.* (2013) One-step real time RT-PCR for detection of microRNAs, *Talanta*, **110**, 190-195.

Yin, V.P. (2018) In Situ Detection of MicroRNA Expression with RNAscope Probes, *Methods in molecular biology*, **1649**, 197-208.

Zhang, D., *et al.* (2018) The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes, *Science*, **359**, 587-592.