

SOFTWARE

Open Access



# CeNet Omnibus: an R/Shiny application to the construction and analysis of competing endogenous RNA network

Xiao Wen, Lin Gao<sup>\*</sup> , Tuo Song and Chaoqun Jiang

\*Correspondence:  
lgao@mail.xidian.edu.cn  
School of Computer Science  
and Technology, Xidian  
University, Xi'an, Shaanxi,  
China

## Abstract

**Background:** The competing endogenous RNA (ceRNA) regulation is a newly discovered post-transcriptional regulation mechanism and plays significant roles in physiological and pathological progress. CeRNA networks provide global views to help understand the regulation of ceRNAs. CeRNA networks have been widely used to detect survival biomarkers, select candidate regulators of disease genes, and predict long noncoding RNA functions. However, there is no software platform to provide overall functions from the construction to analysis of ceRNA networks.

**Results:** To fill this gap, we introduce CeNet Omnibus, an R/Shiny application, which provides a unified framework for the construction and analysis of ceRNA network. CeNet Omnibus enables users to select multiple measurements, such as Pearson correlation coefficient (PCC), mutual information (MI), and liquid association (LA), to identify ceRNA pairs and construct ceRNA networks. Furthermore, CeNet Omnibus provides a one-stop solution to analyze the topological properties of ceRNA networks, detect modules, and perform gene enrichment analysis and survival analysis. CeNet Omnibus intends to cover comprehensiveness, high efficiency, high expandability, and user customizability, and it also offers a web-based user-friendly interface to users to obtain the output intuitively.

**Conclusion:** CeNet Omnibus is a comprehensive platform for the construction and analysis of ceRNA networks. It is highly customizable and outputs the results in intuitive and interactive. We expect that CeNet Omnibus will assist researchers to understand the property of ceRNA networks and associated biological phenomena. CeNet Omnibus is an R/Shiny application based on the Shiny framework developed by RStudio. The R package and detailed tutorial are available on our GitHub page with the URL <https://github.com/GaoLabXDU/CeNetOmnibus>.

**Keywords:** CeRNA, Shiny application, Network analysis

## Background

MicroRNAs are ~22nt small non-coding RNAs, which can bind to microRNA response elements (MREs) on target RNA sequences (e.g. mRNAs, lincRNAs, pseudogenes, and circle RNAs) through the RNA-induced silencing complex (RISC) and



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

lead to transcripts degraded or translation repressed [1–3]. Different transcripts with the same MRE can regulate each other via competitively binding shared microRNAs, which are called competing endogenous RNAs (ceRNAs). The ceRNA regulatory is a new kind of post-transcriptional regulation mechanism and is also considered the “Rosetta Stone of a hidden RNA language” [4, 5]. It has been demonstrated that the complex crosstalk among ceRNAs involves numerous physiological and pathological progress [6–9], such as cell proliferation, differentiation, invasion, and metastasis [10–12]. Because individual microRNAs can target multiple transcripts and the same transcript may contain MREs of different microRNAs, there are giant and complicated networks in cells between microRNAs and their target transcripts, as well as between ceRNAs themselves. An increasing number of studies about ceRNAs and ceRNA networks have been published in the past few years [13, 14], which are widely used to detect survival biomarkers, select candidate regulators of disease genes, and predict long noncoding RNA functions [15].

As the result of the huge scale of the ceRNA networks, computational methods have become efficient approaches to the construction of the ceRNA networks [13, 16]. Current methods are developed based on two basic principles: (1) ceRNA pairs should share a sufficient number of microRNAs; and (2) ceRNA pairs should be co-expressed. For the first measurement, the hyper-geometric test is used to evaluate the enrichment significance of microRNAs of a ceRNA pair. For the second one, statistical indexes can be categorized into two classes: (1) the pair-wised correlation, such as Pearson correlation coefficient (PCC) and mutual information (MI); and (2) partial associated correlation, such as sensitivity correlation (SI) [17], multiple sensitivity correlation [18] and conditional mutual information (CMI) [19]. A part from those methods, Zhang et al. [20] proposed LncmiRSRN to construct a lncRNA-mRNA ceRNA network via estimating the causal effects of lncRNAs on mRNAs with the IDA method [21], where IDA refers to the intervention calculus when the DAG is absent. Recently, we proposed that liquid association (LA) is another new measurement for the identification of ceRNA pairs [22], where LA can be used to estimate the correlation sensitivity of ceRNAs to microRNAs. Besides those computational methods, researchers have established a set of ceRNA databases, such as lncCeDB [23], LncCeR-Base [24], miRSponge [25], LncACTdb [26] and PceRBase [27].

Although there has been such great progress in ceRNA, there is no software platform to provide overall functions to construct and analyze the ceRNA networks. In this study, we developed CeNet Omnibus, an R/Shiny based application, which achieves a unified framework to construct and analyze ceRNA network. CeNet Omnibus intends to cover comprehensiveness, high efficiency, high expandability, and user customizability. CeNet Omnibus integrates complete functions for the construction and analysis of ceRNA networks. In order to accommodate to the different usage scenarios, CeNet Omnibus allows users to define some functions and set parameters in different situations. With the support of the R/Shiny framework, CeNet Omnibus offers a web-based user-friendly interface for users to obtain the output intuitively. We expect that CeNet Omnibus will become an efficient and convenient tool to investigate the ceRNA networks for researchers, especially those, who are not familiar with programming.

### Implementation

CeNet Omnibus is implemented with R programming language (Release 3.6) for the most parts of computation. For some specific functions, we use Java (Version 1.7) codes to improve efficiency. The web-based user interface is developed based on the Shiny framework (Version 1.4) from RStudio, as well as a set of associated packages, such as shinydashboard (Version 0.7), shinyWidgets (Version 0.5). Network construction and analysis are based on igraph package (Version 1.2.4), while the visualization of networks depends on javascript plugin Cytoscape.js (Version 3.13) and visNetwork package (Version 2.0). Other dependency packages are listed on the homepage of CeNet Omnibus.

CeNet Omnibus consists of five components, including data input, data processing, network construction, network visualization, and network analysis. Figure 1 illustrates the framework of CeNet Omnibus.

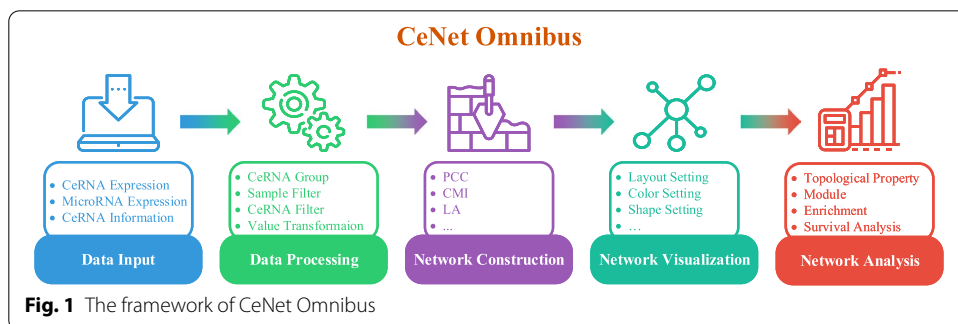
### Data input

Users are expected to upload the expression profiles of candidate ceRNAs, the expression profiles of microRNAs, microRNA- candidate ceRNA interaction data, and candidate ceRNA information as inputs before the further operation. All these data should be plain text delimited by tab, comma, space, semicolon, or any other practicable marks. The expression profiles of candidate ceRNA and microRNA should have sample names. The candidate ceRNA and microRNA symbols should meet the same standard in all these files. In addition to direct uploading, we also provide an interface for users to obtain candidate ceRNA information from Ensembl database [28] with the support of biomaRt [29] package, which enables collect large amounts of data in a uniform way.

### Data processing

At the beginning of data processing, the program will automatically obtain three symbol sets, including (1) candidate ceRNAs symbols that are available in all of candidate ceRNA expression profiles data, microRNA- candidate ceRNA interaction data and candidate ceRNA information data, (2) microRNA symbols available in both microRNA expression profiles data and microRNA- candidate ceRNA interaction data, and (3) sample names available in expression profiles of both candidate ceRNAs and microRNAs.

Considering that there are significant differences between the expression levels of different RNA types and that users may want to investigate a part of candidate ceRNA relationships in the datasets, users are allowed to divide candidate ceRNAs into different



groups to filter non-expressed RNAs or construct a ceRNA subnetwork. For example, the candidate ceRNAs can be divided into the “Noncoding” RNAs group and the “Coding” RNAs group according to their biotypes.

Users are allowed to remove low-quality samples. For this aim, users are required to provide the thresholds for the detected microRNAs and candidate ceRNAs respectively. Then, CeNet Omnibus plots the histograms of the detected RNAs ratios of samples. Users may determine how many samples should be remained based on the histograms. Furthermore, to remove the non-expressed RNAs, users should input the minimal expression levels of microRNAs and each candidate ceRNA group respectively. Then, CeNet Omnibus creates the histograms of the expressing sample ratios of RNAs. Users may also determine how many RNAs should be remained based on the histograms.

Finally, users can transform the expression values of microRNAs and candidate ceRNAs. We provide the log transformation and the normalization transformation. The two transformations can be executed at the same time, in the order of log transformation first and normalization the next. We have defined two specific normalization functions, which are min–max scaling and z-score scaling. Besides, users can define the normalization function themselves.

### Network construction

The construction of ceRNA network is associated with the identification of ceRNA pairs. CeNet Omnibus provides five different measurements for the detection of ceRNA pairs, including shared microRNA enrichment significance (MS), Pearson correlation coefficient (PCC), mutual information (MI), conditional mutual information (CMI), and liquid association (LA). MS is used to test if a candidate ceRNA pair shares a sufficient number of microRNAs. LA is used to estimate the correlation sensitivity of candidate ceRNAs to shared microRNAs. PCC, MI, and CMI are all used to evaluate the levels of co-expression, where PCC can evaluate the linear correlation, while MI and CMI can evaluate the non-linear correlation.

For a given candidate ceRNA pair (e.g.,  $R_1$  and  $R_2$ ), the MS is defined as:

$$MS(R_1, R_2) = 1 - \frac{\sum_{i=0}^{t-1} \binom{T_1}{i} \times \binom{Q - T_1}{T_2 - i}}{\binom{Q}{T_2}}$$

where  $Q$  is the total number of considered microRNAs,  $T_1$  is the number of microRNAs targeting  $R_1$ ,  $T_2$  is the number of microRNAs targeting  $R_2$ , and  $t$  is the number of microRNAs targeting both  $R_1$  and  $R_2$ .

Let  $e_{R_1}$  and  $e_{R_2}$  represent the expression of  $R_1$  and  $R_2$  respectively, and the PCC of  $R_1$  and  $R_2$  is defined as:

$$PCC(R_1, R_2) = \frac{E[(e_{R_1} - E(e_{R_1})) \times (e_{R_2} - E(e_{R_2}))]}{\sqrt{Var(e_{R_1}) \times Var(e_{R_2})}},$$

where  $E(\cdot)$  and  $Var(\cdot)$  represent the expectation and variance of a random variable respectively.

Let  $disc(e_{R_1})$  and  $disc(e_{R_2})$  represent the discretized expression of  $R_1$  and  $R_2$ , the MI of  $R_1$  and  $R_2$  is defined as:

$$MI(R_1, R_2) = \sum_{r_1 \in disc(e_{R_1})} \sum_{r_2 \in disc(e_{R_2})} p_{R_1, R_2}(r_1, r_2) \log \frac{p_{R_1, R_2}(r_1, r_2)}{p_{R_1}(r_1)p_{R_2}(r_2)},$$

where  $p_{R_1, R_2}(r_1, r_2)$  is the joint probability distribution of  $R_1$  and  $R_2$ ,  $p_{R_1}(r_1)$  and  $p_{R_2}(r_2)$  are the marginal distributions of  $R_1$  and  $R_2$  respectively.

Similarly, let  $R_{mi}$  indicate the shared microRNAs between  $R_1$  and  $R_2$ ,  $e_{R_{mi}}$  represent the sum of the expression profile of  $R_{mi}$ , and  $disc(e_{R_{mi}})$  represent the discretized value of  $e_{R_{mi}}$ . The CMI of  $R_1$  and  $R_2$  is defined as:

$$CMI(R_1; R_2 | R_{mi}) = \sum_{r_{mi} \in disc(e_{R_{mi}})} \sum_{r_1 \in disc(e_{R_1})} \sum_{r_2 \in disc(e_{R_2})} p_{R_1, R_2, R_{mi}}(r_1, r_2, r_{mi}) \log \frac{p_{R_{mi}}(r_{mi})p_{R_1, R_2, R_{mi}}(r_1, r_2, r_{mi})}{p_{R_1, R_{mi}}(r_1, r_{mi})p_{R_2, R_{mi}}(r_2, r_{mi})}$$

where  $p_{R_1, R_2, R_{mi}}(r_1, r_2, r_{mi})$  is the joint probability distribution of  $R_1$ ,  $R_2$ , and  $R_{mi}$ .

The LA of  $R_1$  and  $R_2$  is defined as:

$$LA(R_1, R_2 | R_{mi}) = \frac{\sum_{i=1}^N z_i(e_{R_1}) \times z_i(e_{R_2}) \times norm_i(e_{R_{mi}})}{N},$$

where  $N$  is the total number of samples,  $z_i(e_{R_1})$  and  $z_i(e_{R_2})$  are the expression of  $R_1$  and  $R_2$  normalized by z-score transformation in  $i$ th sample respectively, and  $norm_i(e_{R_{mi}})$  is the  $i$ th value of the expression of  $R_{mi}$  normalized via Van der Waerden's method [30].

CeNet Omnibus also allows users to define new measurements for the detection of ceRNA pairs. Users can select one or a set of measurements to detect ceRNA pairs according to their situation.

To improve the efficiency of the calculation, CeNet Omnibus can calculate these measurements parallelly with the parallel package (Version 3.6). In addition, users can only calculate a part of relations in the datasets depending on the results of gene grouping in Data Processing Section.

After the calculation of selected measurements, CeNet Omnibus will create the density plots for each measurement. Users can select a specific threshold for each computed measurement according to their situation. CeNet Omnibus will collect interactions, which are satisfied with all the thresholds, to construct networks. After the construction of the ceRNA network, CeNet Omnibus will summarize the basic information of the network, including the number of connected nodes, isolated nodes, edges, and connected components.

### Network visualization

Users can use the network visualization component to view the overview structure of the constructed ceRNA network. Users are able to modify the layout of the network, the size, color, and shape of specific nodes, as well as the displayed symbols of nodes. Users can also search for nodes they are interested in for the next analysis.

### Network analysis

For further analyzing the ceRNA network, CeNet Omnibus supplies a one-stop solution, including network topological property analysis, network module detection, gene enrichment, and survival analysis.

In the part of network topological property analysis, CeNet Omnibus provides four topological properties of nodes, which include degree, betweenness, clustering coefficient, and closeness, as well as an topological property of edges—edge betweenness. The analysis of network topological properties can help users to find the important nodes and edges in the ceRNA networks for further analysis.

In the part of network module detection, CeNet Omnibus embeds the popular module detection algorithms, such as Louvain method [31], MCL [32], MCODE [33], and EPCA [34]. Because some algorithms require parameters, CeNet Omnibus allows users to test the performance of different parameters, and it will give a report, including the count of modules, the modularity of modules, the average density of the modules, the modules size distribution, the edge counts of modules distribution, and the modules density distribution, to help users determine the final parameters. The microRNAs associated with edges in modules are considered as the module associated microRNAs, and these microRNAs are also listed in the results of network modules. CeRNAs may form dense subnetworks to regulate each other, and current studies have suggested that the ceRNA modules may be potential biomarkers for cancer therapy. These results may be used for further analysis.

For further understanding the biological and medical principles in the ceRNA networks, CeNet Omnibus provides users the functions of enrichment analysis and survival analysis. The current version has integrated the functional enrichment with g:Profiler [35, 36]. Besides, users can input user-defined datasets for other enrichment analysis. Furthermore, the current version has integrated two survival analysis models to investigate the relationships between ceRNAs and disease therapy. Kaplan–Meier survival estimator is used to study the effect of a single factor on survival probability, while the Cox proportional hazards regression model can estimate multiple factors simultaneously.

## Results

### Comparison with related package

CeRNA networks are effective tools to study the ceRNA regulation. However, there are not enough available computational tools to help researchers construct and analyze the ceRNA networks. We surveyed the published computational tools, and only found the CeRNASeek [37], which is an R package to identify and analyze ceRNA-ceRNA interactions. Comparing with CeRNASeek, CeNet Omnibus is more comprehensive, more expandable, and more user-friendly.

CeNet Omnibus and CeRNASeek are both R-based packages. Different from CeRNASeek, CeNet Omnibus provides a user-friendly web-based interface to eliminate the gap between users and programming. CeNet Omnibus provides comprehensive toolkits, from data processing to network analysis. However, CeRNASeek depends on users to input processed data and lacks the analysis of the network properties. Finally, CeNet Omnibus uses the parallel computing technique to speed up the calculation, and

it provides interfaces to expand application scenarios. The comparison details are shown in Table 1.

**Case study: K562 single-cell microRNA-mRNA co-sequencing data**

To examine the performance of CeNet Omnibus, we applied CeNet Omnibus on single-cell microRNA-mRNA co-sequencing data from [38]. The data set contains 23,284 mRNAs/noncoding RNAs FPKM values and 2822 mature microRNAs log2-transformed miRNA expression of 19 K562 single cells. The interactions between ceRNAs and microRNAs are downloaded from mirwalk2.0 database [39]. To obtain the gene biotypes, we downloaded the gene information from Ensembl database [28]. The sample data is shown in Additional file 1: Table S1.

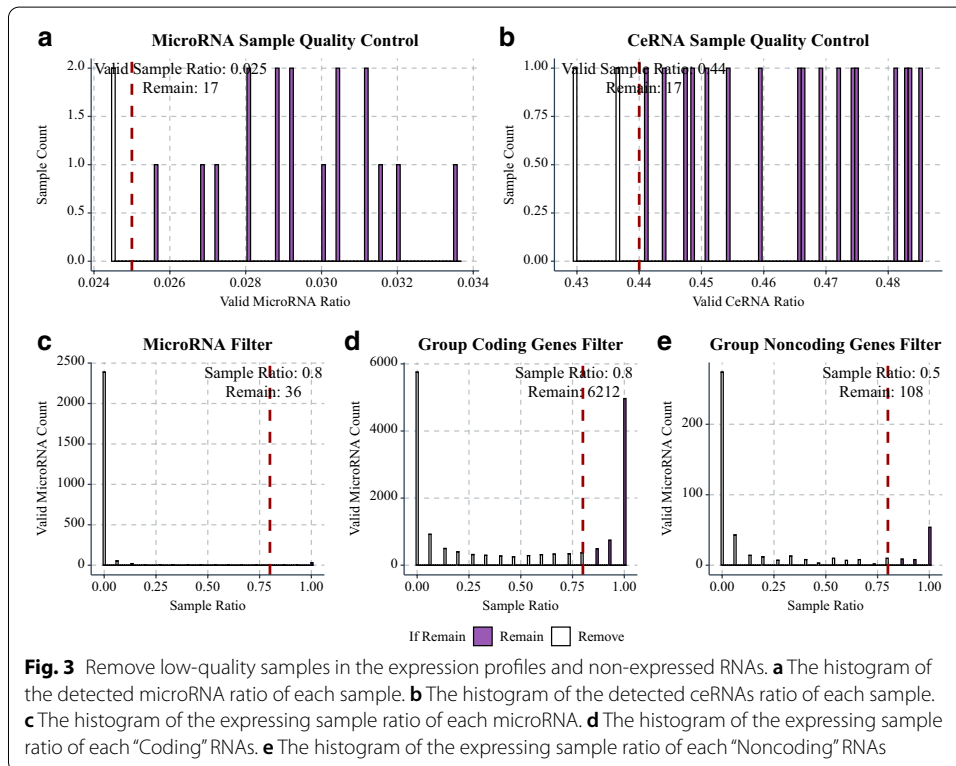
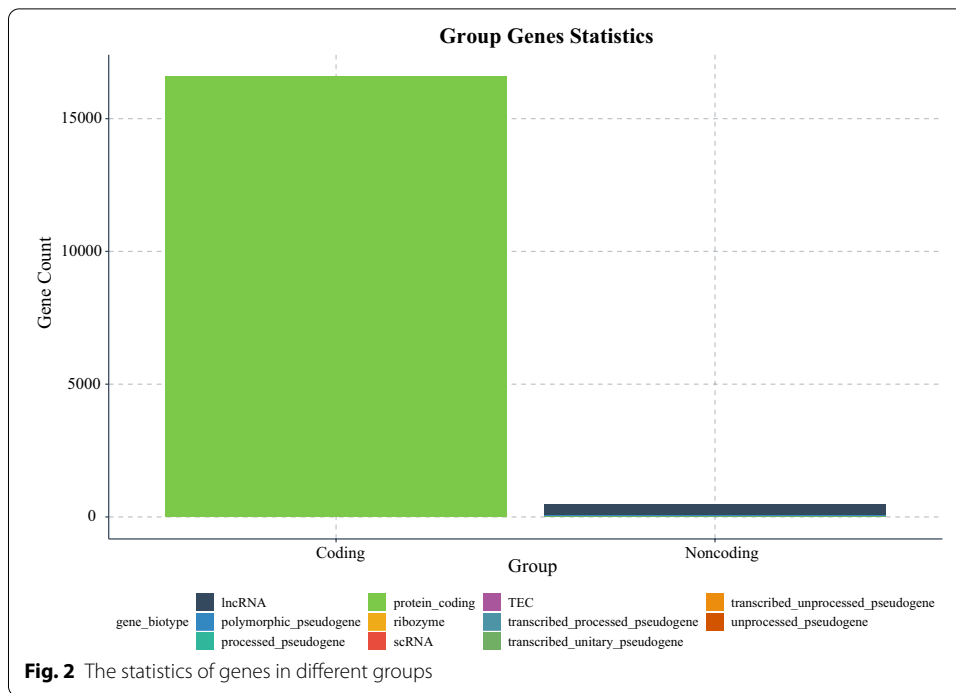
After uploading the four data files into CeNet Omnibus, the program obtains 17,058 valid mRNAs/noncoding RNAs, and 2532 valid microRNAs of 19 valid samples. We group the mRNAs/noncoding RNAs according to gene biotype attribution in gene information files. We put RNAs, whose biotypes are “protein\_coding”, into the “Coding” group, while other RNAs are put into the “Noncoding” group. The gene count statistics are shown in Fig. 2.

We use  $\log_2(1e-4)$  and 1 as the thresholds of detected microRNAs and mRNAs/noncoding RNAs samples respectively, and remove two samples with the lowest detected RNA ratios for each data set (Fig. 3a, b). Then, we obtain 15 high-quality samples for the next step. To exclude lowly expressed microRNAs, we only retain microRNAs with expression value  $> \log_2(2e-4)$  in  $> 80\%$  of cells (Fig. 3c). For RNAs in group “Coding”, we retain RNAs with expression value  $> 1$  in  $> 80\%$  of cells, while for RNAs in group “Noncoding”, we retain RNAs with expression value  $> 0.5$  in  $> 50\%$  of cells (Fig. 3d, e). Finally, we obtain 36 microRNAs, 6212 “Coding” RNAs, and 108

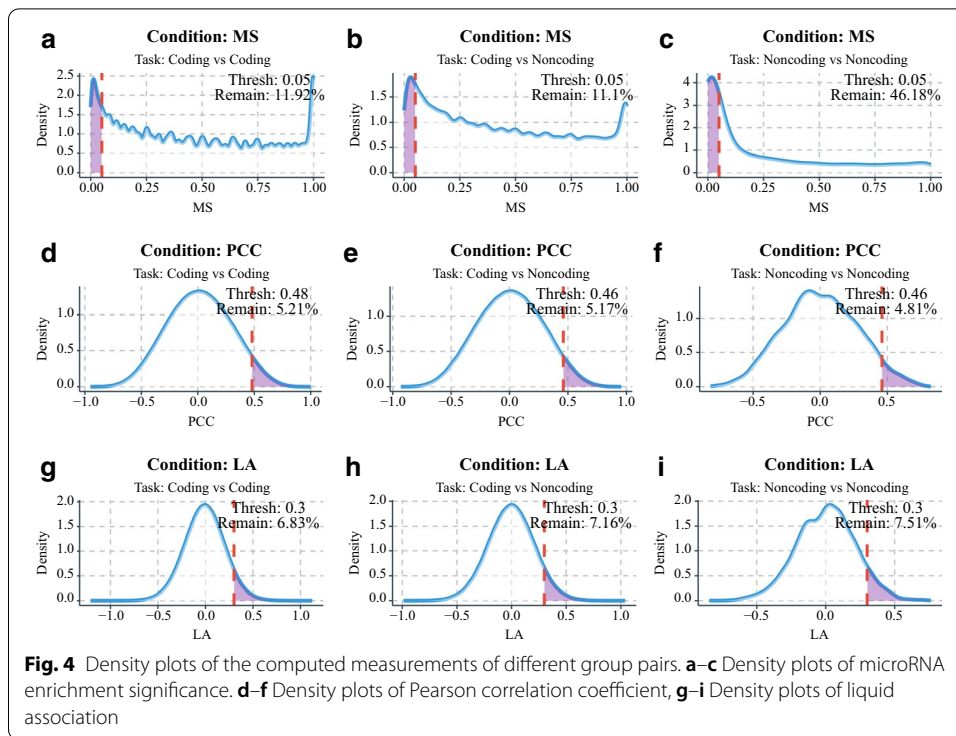
**Table 1 The comparison between CeNet Omnibus and CeRNASeek**

Items		CeNet Omnibus	CeRNASeek
Environment	Language	R	R
	Operation interface	web-base	command line
Functionality	Data processing	Yes	No
	Network Construction	PCC, MS, MI, CMI, LA, and self define	Ratio, HyperT, HyperC, SC, CMI, and Cernia
	Network Visualization	Yes	Yes
	Network topological Property	Yes	No
	Network module	Yes	No
	Enrichment analysis	Yes	Yes
	Survival analysis	Yes	Yes
Feature	Usability	GUI with document tutorial	R function with document tutorial
	Efficiency	Yes	No
	Expandability	Yes	No

“Noncoding” RNAs to construct ceRNA networks. The FPKM values of “Coding” and “Noncoding” RNAs are log2-transformed.





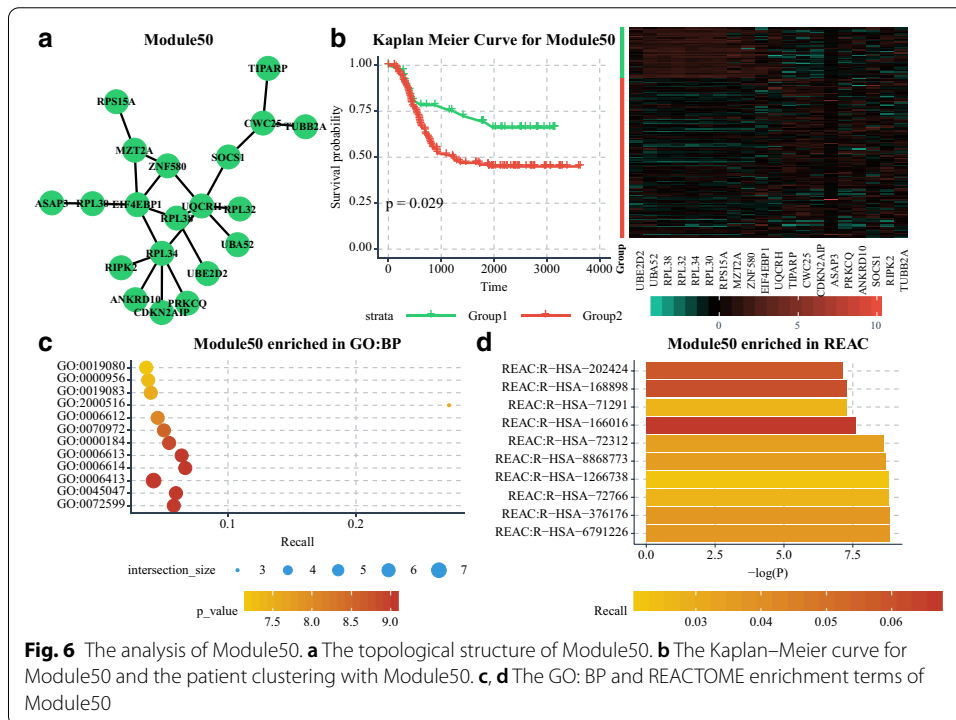
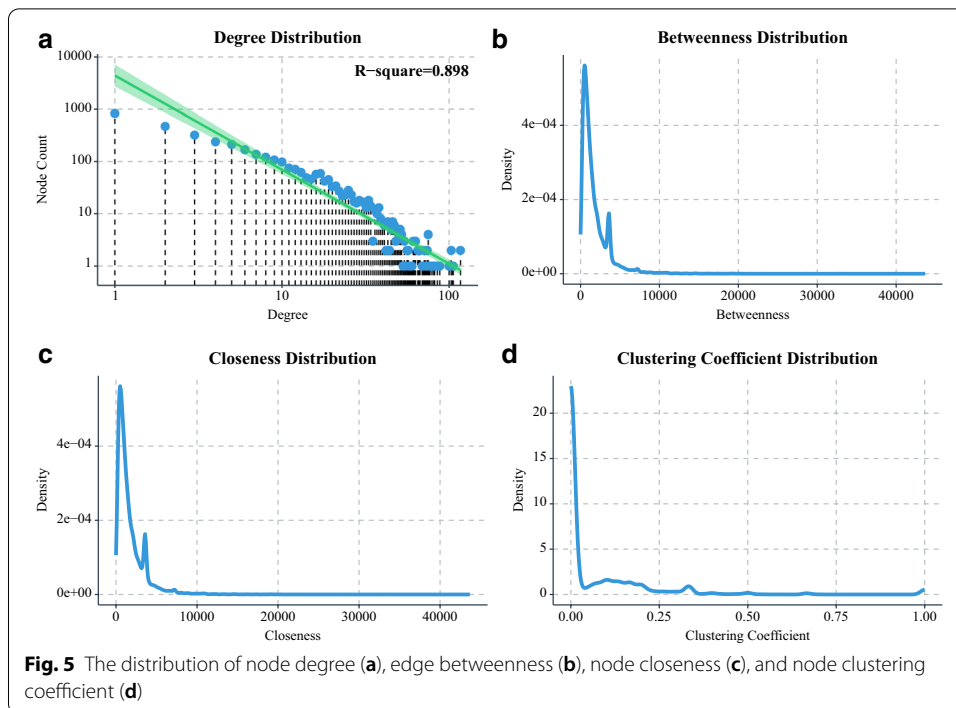


**Table 2** The thresholds used of MS, PCC and LA for each group pair

Measurement	Group pair	Direction	Thresh
MS	Coding versus Coding	<	0.05
	Coding versus Noncoding		
	Noncoding versus Noncoding		
PCC	Coding versus Coding	>	0.48
	Coding versus Noncoding	>	0.46
	Noncoding versus Noncoding	>	0.46
LA	Coding versus Coding	>	0.3
	Coding versus Noncoding		
	Noncoding versus Noncoding		

As CeNet Omnibus provides five different measurements for constructing ceRNA networks, here we use MS, PCC, and LA for the network construction. We calculate these three measurements for each RNA group pair, and the density plots are shown in Fig. 4. We choose the thresholds for each measurement as shown in Table 2. Finally, we obtain a ceRNA network with 3695 nodes, 16,879 edges, and 42 connected components.

We use the network analysis component to analyze the constructed network. Figure 5 shows a part of topological properties distribution of the network, and the specific topological properties of the partial nodes are shown in Additional file 2: Table S2. We use the Louvain method [31] to detect network modules, and the whole module information is listed in the Additional file 3: Table S3. We use Module50



(Fig. 6a) as an example for further analysis. The GO: BP and REACTOME enrichment terms of Module50 are shown in Fig. 6c, d, and Additional file 4: Table S4. We find Module50 is highly associated with RNA transcript and translation. Furthermore, we download the expression profiles and clinical information of 155 acute myeloid

leukemia (AML) patients from TCGA to evaluate if genes of Module50 can be a module biomarker to predict patient prognostic. As shown in Fig. 6b, the patient group with lower expressions of ZNF580, RPL34, RPL30, RPS15A, RPL32, RPL38, and UBA52 has relatively worse prognostic than the group with higher expression.

## Conclusion

CeNet Omnibus is a comprehensive platform for the construction and analysis of ceRNA networks. It provides a friendly-used framework to process uploaded data, apply different strategies to construct ceRNA networks, and investigate the biological and medical principles in the networks. It is highly customizable and outputs the results in intuitive and interactive. In testing on our data, our platform successfully constructed a single-cell ceRNA network and detected a prognostic module biomarker with abundant biological functions. Given its design as an open-source Shiny application with a web-based interface, CeNet Omnibus is an important tool for researchers to systematically study the properties of different ceRNA networks and their effects in different biological progresses.

## Availability and requirements

Project name: CeNet Omnibus

Project home page: <https://github.com/GaoLabXDU/CeNetOmnibus>

Operating system(s): Platform independent

Programming language: R

Other requirements: Java 1.7 or higher, web browsers, internet connectivity

License: MIT

Any restrictions to use by non-academics: None

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12859-021-04012-y>.

**Additional file 1: Table S1.** The data sample uploaded to the platform.

**Additional file 2: Table S2.** The information of the topological properties of a part of nodes.

**Additional file 3: Table S3.** The basic information of modules.

**Additional file 4: Table S4.** The enrichment terms of Module50.

## Abbreviations

AML: Acute myeloid leukemia; ceRNA: Competing endogenous RNA; CMI: Conditional mutual information; FPKM: Fragments per kilobase of transcript per million; LA: Liquid association; MI: Mutual information; MRE: MicroRNA response element; MS: Shared microRNA enrichment significance; PCC: Pearson correlation coefficient; RISC: RNA-induced silencing complex; TCGA: The Cancer Genome Atlas.

## Acknowledgements

The authors would like to thank Ran Duan and Peizhuo Wang for discussion about the design of the web page and test of the programs. The authors would like to thank Ran Duan for the modification of the manuscript.

## Authors' contributions

LG conceived the project. XW designed the structure of the project. XW, TS, and CQJ developed and tested the program. XW analyzed the interpreted the data and generated the figures in the manuscript. XW wrote the manuscript. LG supervised the research and edited the manuscript. All authors have read and approved the manuscript.

**Funding**

This work was supported by the National Natural Science Foundation of China (Grant Nos. 61532014, 61672407), the National Key R&D Program of China (Grant No. 2018YFC0910400), Shanghai Municipal Science and Technology Major Project (No.2018SHZDZX01), and LCNBI and ZJLab and the Fundamental Research Funds for the Central Universities (ZD2009). These funding bodies played no role in any of the following: study design, data collection, analysis, and interpretation; manuscript writing.

**Availability of data and materials**

All datasets used in this study have been previously published. The single cell RNA and microRNA expression profiles are available in Gene Expression Omnibus repository [<https://www.ncbi.nlm.nih.gov>] under GEO series GSE114071. The data of AML patients is available in GDC [<https://portal.gdc.cancer.gov/>].

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

Received: 3 August 2020 Accepted: 8 February 2021

Published online: 18 February 2021

**References**

- Ambros V. The functions of animal microRNAs. *Nature*. 2004;431(7006):350–5.
- Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*. 2004;116(2):281–97.
- Bartel DP. Metazoan MicroRNAs. *Cell*. 2018;173(1):20–51.
- Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell*. 2011;146(3):353–8.
- Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. *Nat Rev Genet*. 2016;17(5):272–83.
- Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature*. 2010;465(7301):1033–8.
- Conte F, Fiscon G, Chiara M, Colombo T, Farina L, Paci P. Role of the long non-coding RNA PVT1 in the dysregulation of the ceRNA-ceRNA network in human breast cancer. *PLoS ONE*. 2017;12(2):e0171661.
- Russo F, Fiscon G, Conte F, Rizzo M, Paci P, Pellegrini M. Interplay between long noncoding RNAs and microRNAs in cancer. *Methods Mol Biol*. 2018;1819:75–92.
- Qiannan E, Wang C, Gu X, Gan X, Zhang X, Wang S, Ma J, Zhang L, Zhang R, Su L. Competitive endogenous RNA (ceRNA) regulation network of lncRNA-miRNA-mRNA during the process of the nickel-induced steroidogenesis disturbance in rat Leydig cells. *Toxicol In Vitro*. 2020;63:104721.
- Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. *Nature*. 2014;505(7483):344–52.
- Sanchez-Mejias A, Tay Y. Competing endogenous RNA networks: tying the essential knots for cancer biology and therapeutics. *J Hematol Oncol*. 2015;8:30.
- Xia W, Mao Q, Chen B, Wang L, Ma W, Liang Y, Zhang T, Dong G, Xu L, Jiang F. The TWIST1-centered competing endogenous RNA network promotes proliferation, invasion, and migration of lung adenocarcinoma. *Oncogenesis*. 2019;8(11):62.
- Le TD, Zhang J, Liu L, Li J. Computational methods for identifying miRNA sponge interactions. *Brief Bioinform*. 2017;18(4):577–90.
- Sherin K, Nair A. Review of computational prediction of competing endogenous RNA. *J Proteomics Bioinform*. 2019;12:2.
- Liu K, Yan Z, Li Y, Sun Z. Linc2GO: a human LincRNA function annotation resource based on ceRNA hypothesis. *Bioinformatics*. 2013;29(17):2221–2.
- Qi X, Lin Y, Chen J, Shen B. Decoding competing endogenous RNA networks for cancer biomarker discovery. *Brief Bioinform*. 2020;21(2):441–57.
- Paci P, Colombo T, Farina L. Computational analysis identifies a sponge interaction network between long non-coding RNAs and messenger RNAs in human breast cancer. *BMC Syst Biol*. 2014;8:83.
- List M, Dehghani Amirabad A, Kostka D, Schulz MH. Large-scale inference of competing endogenous RNA networks with sparse partial correlation. *Bioinformatics*. 2019;35(14):596–604.
- Sumazin P, Yang X, Chiu HS, Chung WJ, Iyer A, Llobet-Navas D, Rajbhandari P, Bansal M, Guarnieri P, Silva J, et al. An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma. *Cell*. 2011;147(2):370–81.
- Zhang J, Liu L, Li J, Le TD. LncmiRSRN: identification and analysis of long non-coding RNA related miRNA sponge regulatory network in human cancer. *Bioinformatics*. 2018;34(24):4232–40.
- Maathuis MH, Kalisch M, Buhlmann P. Estimating high-dimensional intervention effects from observational data. *Ann Stat*. 2009;37:3133–64.
- Wen X, Gao L, Hu Y. LAceModule: identification of competing endogenous RNA modules by integrating dynamic correlation. *Front Genet*. 2020;11:235.

23. Das S, Ghosal S, Sen R, Chakrabarti J. InCeDB: database of human long noncoding RNA acting as competing endogenous RNA. *PLoS ONE*. 2014;9(6):e98965.
24. Pian C, Zhang G, Tu T, Ma X, Li F. LncCeRBase: a database of experimentally validated human competing endogenous long non-coding RNAs. *Database (Oxford)*. 2018;2018:bay061.
25. Wang P, Zhi H, Zhang Y, Liu Y, Zhang J, Gao Y, Guo M, Ning S, Li X. miRSponge: a manually curated database for experimentally supported miRNA sponges and ceRNAs. *Database (Oxford)*. 2015;2015:bav098.
26. Wang P, Li X, Gao Y, Guo Q, Wang Y, Fang Y, Ma X, Zhi H, Zhou D, Shen W, et al. LncACTdb 2.0: an updated database of experimentally supported ceRNA interactions curated from low- and high-throughput experiments. *Nucleic Acids Res*. 2019;47(D1):D121–7.
27. Yuan C, Meng X, Li X, Illing N, Ingle RA, Wang J, Chen M. PceRBase: a database of plant competing endogenous RNA. *Nucleic Acids Res*. 2017;45(D1):D1009–14.
28. Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, et al. Ensembl 2020. *Nucleic Acids Res*. 2020;48(D1):D682–8.
29. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc*. 2009;4(8):1184–91.
30. Li KC. Genome-wide coexpression dynamics: theory and application. *Proc Natl Acad Sci USA*. 2002;99(26):16875–80.
31. Blondel VD, Guillaume J, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech*. 2008;2008(10):10008.
32. Dongen S: Graph clustering by flow simulation. University of Utrecht; 2000.
33. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*. 2003;4:2.
34. Jia S, Gao L, Gao Y, Nastos J, Wang Y, Zhang X, Wang H. Defining and identifying cograph communities in complex networks. *New J Phys*. 2015;17(1):013044.
35. Reimand J, Kull M, Peterson H, Hansen J, Vilo J. g:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res*. 2007;35(Web Server issue):W193–200.
36. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res*. 2019;47(W1):W191–8.
37. Zhang M, Jin X, Li J, Tian Y, Wang Q, Li X, Xu J, Li Y, Li X. CeRNASeek: an R package for identification and analysis of ceRNA regulation. *Brief Bioinform*. 2020: bbaa048.
38. Wang N, Zheng J, Chen Z, Liu Y, Dura B, Kwak M, Xavier-Ferrucio J, Lu YC, Zhang M, Roden C, et al. Single-cell microRNA-mRNA co-sequencing reveals non-genetic heterogeneity and mechanisms of microRNA regulation. *Nat Commun*. 2019;10(1):95.
39. Dweep H, Gretz N. miRWalk2.0: a comprehensive atlas of microRNA-target interactions. *Nat Methods*. 2015;12(8):697.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

