OXFORD

*Briefings in Bioinformatics*, 2014, 1–7

**doi: 10.1093/bib/bbu048**
Paper

# LncTar: a tool for predicting the RNA targets of long noncoding RNAs

Jianwei Li, Wei Ma, Pan Zeng, Junyi Wang, Bin Geng, Jichun Yang and Qinghua Cui

Corresponding authors. Qinghua Cui, Department of Biomedical Informatics, Peking University Health Science Center, Beijing, China. Tel: +8610-82801585; Fax: +8610-82801001; E-mail: cuiqinghua@hsc.pku.edu.cn; Jianwei Li, School of Computer Science, Hebei University of Technology, Tianjin, China. Tel: +8610-13821072360; Fax: +8622-60435867; E-mail: lijianwei@hebut.edu.cn



## Abstract

Long noncoding RNAs (lncRNAs) represent a big category of noncoding RNA molecules, and increasing studies have shown that they play important roles in various critical biological processes. They show a diversity of functions through diverse mechanisms, among which regulating RNA molecules is one of the most popular ones. Given the big number of lncRNAs, it becomes urgent and important to predict the RNA targets of lncRNAs in a large scale for the comprehensive understanding of lncRNA functions and action mechanisms. Although several methods have been developed to predict RNA–RNA interactions, none of them can be used to predict the RNA targets of lncRNAs in a large scale. Here we presented a tool, LncTar, which shows the ability to efficiently predict the RNA targets of lncRNAs in a large scale. To test the accuracy of LncTar, we applied it to 10 experimentally supported lncRNA–mRNA interactions. As a result, LncTar successfully predicted 8 (80%) of the 10 lncRNA–mRNA pairs, suggesting that LncTar has a reliable accuracy. Finally, we believe that LncTar could be an efficient tool for the fast identification of the RNA targets of lncRNAs. LncTar is freely available at http://www.cuilab.cn/lnctar.

**Key words:** long noncoding RNAs; RNA–RNA interaction

## Introduction

The recent development of high-throughput technologies showed that most of the human genome are transcribed and discovered a big number of noncoding transcripts [1]. It was reported that ∼75% of the human genome is transcribed to RNAs; however, only ∼2% of the RNAs can be translated to proteins [2, 3]. Thus, the majority of the human genome transcripts are noncoding RNAs, especially long noncoding RNAs (lncRNAs)

**Jianwei Li** is an associate professor at the School of Computer Science, Hebei University of Technology, Tianjin, China. He has been working in the field of translational biomedical informatics since 2011.
**Wei Ma** is a PhD student at the Department of Biomedical Informatics, Peking University Health Science Center. He has been working in the field of Bioinformatics since 2012.
**Pan Zeng** is a PhD student at the Department of Biomedical Informatics, Peking University Health Science Center. He has been working in the field of Bioinformatics since 2013.
**Junyi Wang** is a visiting student at the Department of Biomedical Informatics, Peking University Health Science Center. He has been working in the field of Bioinformatics since 2013.
**Bin Geng** is an associate professor at the Department of Physiology and Pathophysiology, Peking University Health Science Center, Beijing, China. His research interest is noncoding RNAs and cardiovascular disease.
**Jichun Yang** is a professor at the Department of Physiology and Pathophysiology, Peking University Health Science Center, Beijing, China. His research interest is noncoding RNAs and diabetes.
**Qinghua Cui** is a professor at the Department of Biomedical Informatics, and a PI at the Institute of Systems Biomedicine and the MOE Key Lab of Molecular Cardiovascular Sciences, Peking University Health Science Center, Beijing, China. His research interests include systems biology of noncoding RNA, systems pharmacology, biological network simulation and modeling.
**Submitted:** 19 September 2014; **Received (in revised form):** 13 November 2014

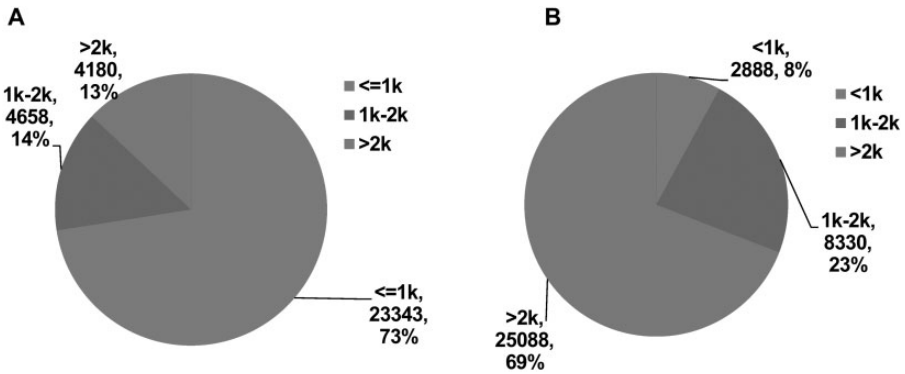


**Figure 1.** Distributions of the RNA size of the human lncRNAs (A) and the human mRNAs (B).

[4]. LncRNAs are generally defined as noncoding RNA molecules >200 nucleotides (nt) in length [5]. Because lncRNAs usually show low cross-species conservation, low expression levels and high tissue specificity, historically people often argue against their functionality [6]. In recent years, however, accumulating studies showed that lncRNAs have complex and diverse functions. And thus their dysfunctions could implicate in a wide spectrum of diseases [7, 8]. To play various functions, lncRNAs could select a diversity of mechanisms by interacting with different biological molecules, such as proteins [9], RNAs [10] and DNA [11]. Among these lncRNA-interacting molecules, RNAs represent one of the most popular ones. It has been shown that lncRNAs could regulate a number of key biological processes by interacting with their target RNAs. For instance, the binding of lncRNA BACE1-AS with its target BACE1 mRNA increases BACE1 mRNA stability and thus regulates BACE1 mRNA and subsequently BACE1 protein expression *in vitro* and *in vivo* [12, 13]. Therefore, identifying the RNA targets of lncRNAs will be of great help for the understanding of lncRNA functions and mechanisms. It is thus becoming emergent and important to develop efficient bioinformatics tools to predict RNA targets for lncRNAs. Currently, our knowledge about the nature of lncRNA–RNA interaction is very limited. lncRNAs are long and thus could form complex tertiary structures for interacting with other biological molecules. Because the tertiary structure of an lncRNA plays key roles in its interacting with other molecules, the prediction of interactions between lncRNAs and other molecules will become to be extremely difficult because no computational methods can solve the problems of lncRNA tertiary structures. After careful checking the reported interactions of lncRNAs with other molecules, we found that the tertiary structures play key roles in lncRNA–protein interactions. However, we found that base pairing could play key roles in the lncRNA–mRNA interactions [8, 14]. Moreover, it seems that basically the base pairing is the most simple and convenient way for RNA–RNA interaction. Although the tertiary structures of lncRNAs play some roles in lncRNA–RNA interaction, based on the above analysis, we hypothesized that base pairing could play key roles in lncRNA–RNA interaction.

Although some bioinformatics tools for predicting RNA–RNA interactions have been developed, such as IntaRNA [15], GUUGle [16], RactIP [17] and RNAup [18], none of them can be used for the large-scale prediction of the RNA targets of lncRNAs for the following reasons. Firstly, all of these methods have limits to RNA size. For example, IntaRNA, a tool to predict interacting regions between two RNA molecules by incorporating the accessibility of both interaction sites and the presence of a seed interaction, requires the RNA size ≤2 kb. RactIP can integrate approximate information on an ensemble of equilibrium joint structures into the objective function of integer programming using posterior internal and external base-paring probabilities. It limits the max RNA size to 1 kb. RNAup can be combined with other faster methods for assessing RNA–RNA interactions. The RNA size for RNAup is not bigger than 5 kb. However, many lncRNAs and other RNA molecules are bigger than the RNA size limits for the above tools. For example, 4180 (13%) of the lncRNAs and 25 088 (69%) of the human mRNAs are bigger than 2 kb (Figure 1). As a result, these tools can not be applied to a big number of RNA molecules. The second reason is that these tools do not have a quantitative standard to determine whether two RNA molecules interact with each other automatically. They only output the matching status of two RNA molecules. Users need to manually judge whether two RNA molecules interact with each other based on their expert knowledge. As a result, these methods are not feasible for large-scale prediction of RNA–RNA interactions. The third reason is that the input parameters of these tools are complex and the calculation results are heavily dependent on these parameters. However, for most users, it is very difficult to determine suitable values for these parameters. For example, RNAup requests users to set the parameter of maximal length of the region of interaction for two RNAs. However, it is difficult to know the length of the interaction region between two RNA molecules. Taken together, novel bioinformatics tools are needed for efficiently predicting lncRNA–RNA interactions.

In this article, we present a novel bioinformatics tool, LncTar, to explore lncRNA–RNA interactions by finding the minimum free energy joint structure of two RNA molecules based on base pairing. LncTar overwhelmed the existing RNA–RNA prediction tools on the following aspects. LncTar does not have a limit to RNA size and can process all length of current RNA molecules. More importantly, LncTar provides a quantitative standard to automatically determine whether two RNA molecules interact with each other. Moreover, users only need to input several exact parameters for LncTar, such as input file name and output file name. In addition, the algorithm of LncTar has a running time of $o(n^2)$, which makes it possible for fast and global prediction of the RNA targets for given lncRNAs. When an lncRNA binds to an RNA target, it probably has an unpaired conformation [19]. However, most of the current RNA–RNA interaction prediction methods do not take account of multiple binding sites when evaluating the unpaired conformation [19–21]. LncTar takes account of multiple binding sites using a matching algorithm, which finds the region of the minimum free energy joint structure between the input RNA sequences. Finally, to evaluate the accuracy of LncTar, we applied it to 10

experimentally supported lncRNA–mRNA interactions. As a result, LncTar successfully predicted 8 (80%) of the 10 lncRNA–mRNA pairs, suggesting that LncTar has a reliable accuracy.

## Methods and materials

### Experimentally supported lncRNA–mRNA interaction data

We downloaded the interaction data between lncRNAs and other molecules from the LncRNADisease database [22] and the NPinter database [23]. We next curated the lncRNA–mRNA interactions if original references clearly described that the lncRNA and the mRNA interact with each other directly. As a result, we obtained 10 lncRNA–mRNA interactions (Table 1), which were taken as the experimentally supported lncRNA–mRNA interaction data set to test the prediction accuracy of LncTar.

### Algorithms and implementation

Basically, we hypothesized that base pairing plays the critical roles in RNA–RNA interactions. Given that, one important step in real-time polymerase chain reaction (PCR) design is detecting primer–dimer, which is also a process of base pairing in nature. Therefore, the primer–dimer prediction algorithms such as Autodimer [24] and MPprimer [25] shed light on the prediction of RNA–RNA interaction. For predicting lncRNA–RNA interactions, here LncTar modified the primer–dimer prediction algorithm of PerlPrimer [26], which is an open source code software for designing primers in standard, bisulphate and real-time PCR. LncTar integrated the precise melting-temperature and primer–dimer prediction algorithms for lncRNA–RNA interaction prediction [27]. LncTar also integrated the primer–dimer checking program for calculating bimolecular secondary structures of input RNA molecules PCR.

LncTar first takes the input lncRNA and the other RNA molecule as the forward and reverse primers, respectively. And then LncTar predicts RNA–RNA interactions by creating a two-dimension binding matrix $M_i$ between paired RNA molecules. The complementarity of each pairing combination between the two RNAs is recorded in $M_i$. For searching binding regions, AutoDimer and Primer3 use the matching score [28], which is not accuracy and reliable. Therefore, LncTar predicts binding regions between paired RNAs using a Nearest-Neighbor method base on thermodynamic parameters [29–31]. The major advantage of this method is that it is more reliable and efficient than the matching score method [24, 28]. LncTar evaluates the free energy by walking through every binding region between the two input RNA molecules. It reads the complementarity between the two RNAs from the matrix $M_i$ and maximizes the number of base pairs among interacting sequences. LncTar calculates the approximate binding free energy, deltaG (dG), of each pairing with the recent data that indicate the stability of complementarity [29].

It has been reported that thermodynamics for double helix formation between two-paired RNAs can be calculated with Nearest-Neighbor parameters [32]. In the Nearest-Neighbor model, enthalpy change ($\Delta H^\circ$), entropy change ($\Delta S^\circ$), free energy change ($\Delta G^\circ$) and melting temperature $T_m$ were calculated [33]. Sequence dependent stability of two paired RNAs is determined by Nearest-Neighbor doublets. For example, the 10 unique internal doublets adopted in duplex RNA are AT/AT, TA/TA, AA/TT, AC/GT, CA/TG, TC/GA, CT/AG, CG/CG, GC/GC and GG/CC [32].

The total difference of free energy $\Delta G^\circ$ between two paired RNAs can be calculated from $\Delta H^\circ$ and $\Delta S^\circ$ using following formula [29].

$$\Delta G^\circ_T = \Delta H^\circ - T\Delta S^\circ$$

For a specific temperature, LncTar can compute the total difference of free energy $\Delta G^\circ$ using Nearest-Neighbor doublets [30].

In the primer–dimer prediction, two kinds of primer–dimer stability are calculated. One is the extensible dimer, which will reduce the amplification of products. The other is the non-extensible dimer, which can reduce the free primer population in a reaction. In this study, our aim is to predict the interaction of lncRNAs and RNAs, only the non-extensible dimers are retained in LncTar, which can largely reduce running time of algorithm. The flowchart of LncTar is shown as Figure 2.

### User-defined parameters

LncTar introduced a novel parameter, normalized free energy (ndG), which reflects the relative stability of internal base pairs in the paired RNAs [34–36]. Given that normally longer RNA molecules often have lower dG, it seems not suitable to take dG as a standard to determine whether two RNA molecules interact

**Table 1**: The LncTar prediction results of the 10 experimentally supported pairs of lncRNAs and their target mRNAs

| lncRNA | mRNA target | dG | ndG |
|---|---|---|---|
| HIF1A-AS2 | HIF1A_NM_181054 | −1182.5 | −0.5765 |
| HIF1A-AS1 | HIF1A_NM_001530 | −168.35 | −0.2582 |
| BACE1-AS | BACE1_NM_012104 | −137.79 | −0.1727 |
| BC200 | MAP1B | −34.4 | −0.1720 |
| BC200 | ARC | −27.46 | −0.1373 |
| BDNF-AS_NR_033312.1 | BDNF_NM_170734 | −277.26 | −0.1362 |
| BC200 | CAMK2A_NM_171825 | −26.75 | −0.1338 |
| PEG3-AS1 | PEG3_NM_006210 | −1395.39 | −1.0619 |
| NPPA-AS1 | NPPA | −75.92 | −0.0885 |
| EMX2OS | EMX2_NM_004098 | −92.61 | −0.0318 |

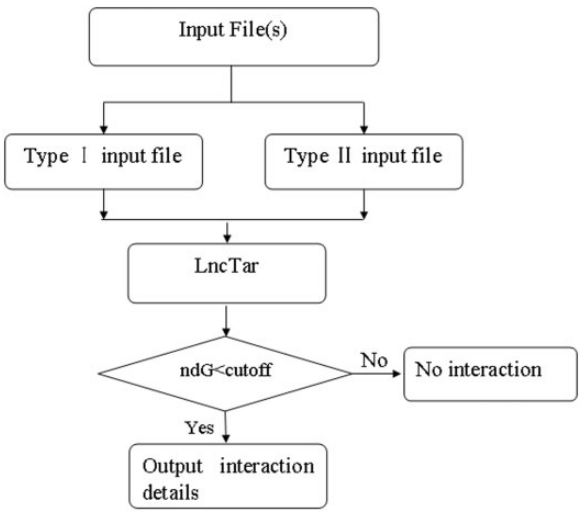


**Figure 2.** Flowchart of LncTar on predicting lncRNA-mRNA interactions.

with each other. The ndG normalized by the size of RNA molecules will be better than the original dG. Therefore, in LncTar, we introduced ndG, which was calculated by the following equation:

$$ndG = dG/\min\left(length_{\text{ln cRNA}}, length_{mRNA}\right)$$

where $length_{\text{ln cRNA}}$ and $length_{mRNA}$ were the lengths of the candidate lncRNA and the other RNA sequences, respectively. The *ndG* was set as a float number and could be a cutoff to determine whether two RNA molecules interact with each other. LncTar will determine two input RNA molecules interact with each other if the calculated ndG is equal to or less than the ndG cutoff (e.g. −0.1). One key point for the successful determination of two interacting RNA molecules is the selecting of an appropriate cutoff. In the simulation for 5000 random lncRNA–mRNA pairs, we found the −0.1 is a cutoff point for the top 5% least ndG. Moreover, the original primer–dimer detection algorithm suggests a dG of −2.0 as the cutoff. Given that normally primers are 20∼30 nt in length, the ndG cutoff for primier–dimer detection is −0.1∼−0.15. Here, therefore, we suggest a normal ndG cutoff, −0.1. The users could select a higher or lower cutoff, for example −0.08, −0.13, −0.15 and even −0.20. It should be noted that a lower cutoff will decrease predicted positives and thus will miss more true positives. A higher cutoff, however, will increase predicted positives and thus will get more false positives.

In summary, although LncTar used the mathematical algorithm of the primer–dimer detection program from PerlPrimer, it has the following novel and significant contributions. (i) Based on the observation of the nature of lncRNA–RNA interaction as described above, we believe that the primer–dimer detection algorithm in PerlPrimer can be used in lncRNA–RNA interaction although the original program is for short sequences (normally 20–30 nt). It is the first time that the PerlPrimer algorithm was adopted for lncRNA–RNA interaction prediction, which is a totally novel scientific goal. (ii) LncTar introduced a novel metric, normalized deltaG (ndG), which is not presented in PerlPrimer. The ndG metric is a critical parameter in LncTar because this parameter is used to automatically and quantitatively determine whether an lncRNA–RNA pair interacts with each other.

## Results

### Input file format

To increase the generality of LncTar, two types of input file formats are used. The first file format includes two input text files. Each text file contains one list of RNA sequences. Each file has two columns separated by the key 'TAB'. Each line represents one RNA molecule. The first column is the RNA name and the second column is the RNA sequence. Figure 3 shows this type of input file format. For this type of input file format, running LncTar needs two text files of this format. In this case, each sequence in one text file will be paired with every sequence in the other text file.

The second file format only needs one input text file. Each line in the text file includes four columns separated by the key of 'TAB'. The first and the third columns are the names of two RNA molecules. The second and the fourth columns are the sequences of the two RNAs. In this case, LncTar will pair the two RNA molecules given in each line. This type of input file format is shown as Figure 4.

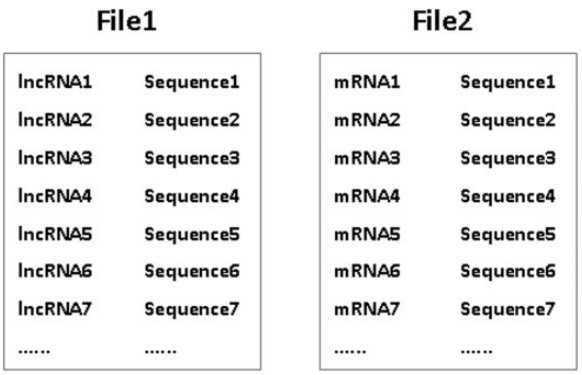


**Figure 3.** The first type of file format. Each file contains two columns separated by TAB key. Each line represents one RNA. For each line, the first column is the RNA name and the second column is the RNA sequence.

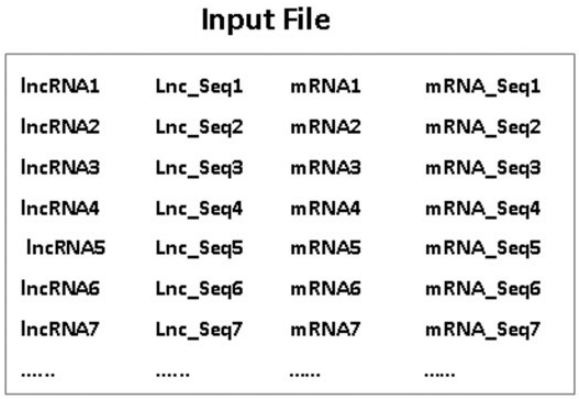


**Figure 4.** The second type of file format. Each file contains four columns separated by TAB key. Each line represents one pair of RNAs. For each line, the first and the third columns are the RNA names. The second and the fourth columns are the RNA sequences.

### Output file format

In LncTar, the predicted results are saved in an output text file and both input file formats generate the same format of output files. The outputs of the prediction results that passed the ndG cutoff will be saved into one output file. Each line in the output file represents one prediction for the paired RNAs, including the names and length of them. In addition, the dG and ndG are provided as well. Moreover, if the users want to investigate the positions of paired bases between the two RNAs, LncTar can also write the pairing graphical display to the output file, as shown in Figure 5.

### Validation of LncTar accuracy

To test the accuracy of LncTar, we applied it to known lncRNA–RNA interactions. For this purpose, we first curated the LncRNADisease database and the NPInter database for experimentally supported lncRNA–mRNA interactions. As a result, we obtained 10 lncRNA–mRNA interactions. We next applied LncTar to the 10 lncRNA–mRNA interactions. To set a quantitative ndG cutoff, we calculated the ndG of 5000 random lncRNA–mRNA pairs using LncTar. The 5000 random lncRNA–mRNA pairs were selected as the following process. We first randomly selected 5000 lncRNAs from the whole human lncRNAs and 5000 mRNAs from the whole human mRNAs.


**Figure 5.** The format of output file. For each interaction, LncTar outputs the name and length of query RNA and target RNA. The dG, ndG, start positions and end positions of binding region in query RNA and target RNA. If the user selects to output the interaction details, a character graph will be shown below the above parameters. A dot in the graph indicates there is a paired of bases in the two sequences. For this case, because the two input sequences are too big, only part of the binding region is shown in this figure.

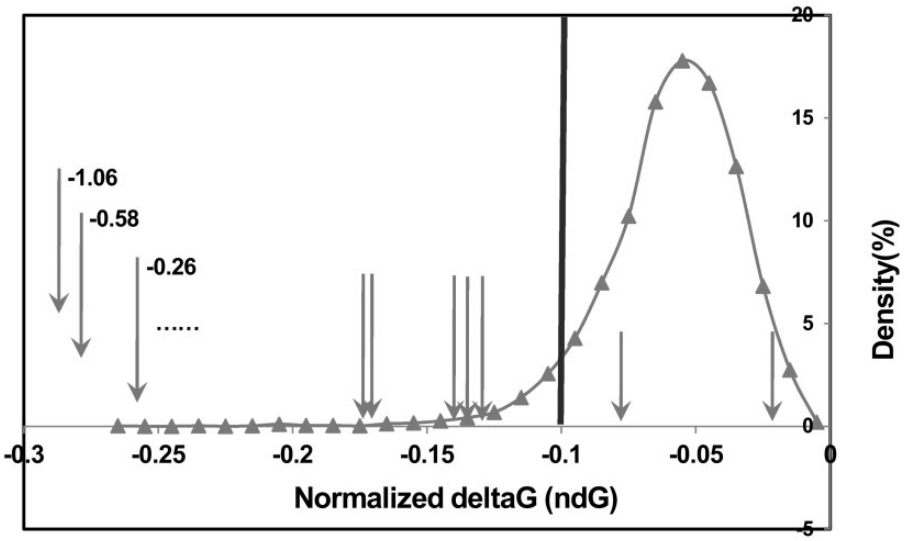


**Figure 6.** The LncTar prediction results of the 10 pairs of experimentally supported lncRNA–mRNA interactions. Density means the percentage of lncRNA–mRNA pairs with specific ndG in the total random lncRNA–mRNA pairs.

We then paired the 5000 lncRNAs and the 5000 mRNAs one by one. By this process, finally we obtained 5000 random lncRNA–mRNA pairs. As a result, the cutoff for the lowest 5% ndG of the 5000 random pairs is −0.1 (the green curve and the black line in Figure 6). Therefore, here we set −0.1 as the cutoff to determine whether two RNA molecules interact with each other. That is, the lncRNA–mRNA pairs with a ndG $\leq -0.1$ will be predicted to be interacting RNA pairs. Otherwise, they do not interact with each other. As a result, 8 (80%) of the 10 experimentally supported lncRNA–mRNA interactions (red arrows in Figure 6) are successfully predicted to be interacting RNA pairs (Table 1), suggesting that these mRNAs could be the RNA targets of the given lncRNAs. These results further indicate that LncTar has a reliable accuracy.

It is very good if comparisons can be made between LncTar and other tools. However, as we described above, none of the current tools can run automatically. The users of these tools need to manually set and understand a number of parameters to determine whether the input RNAs interact with each other. We take RNAup as an example. RNAup needs the users to manually determine the maximal length of the region of interaction as an input parameter. However, it is difficult for users to determine a suitable 'maximal length of the region of interaction'. Moreover, different lncRNA–mRNA pairs should have different 'maximal length of the region of interaction'. In addition, these tools do not have a quantitative standard to tell the users whether the input RNAs interact. The users need to manually determine whether the input RNAs interact with each other based on their expert knowledge. For example, for the pair of HIF1A-AS1::NM_001530, the binding region detected by RNAup is only 25 nt using the default input parameters (the default value for 'maximal length of the region of interaction' is 25). If we set a bigger value to the parameter 'maximal length of the region of interaction', for example 50, RNAup will predict a 50 nt long binding region. LncTar detected a much longer binding region (423 nt), which perfectly covered the predicted binding region by RNAup. However, it is extremely difficult for the RNAup users to determine exactly 423 as the value of the input parameter, 'maximal length of the region of interaction'. Moreover, when we set 'maximal length of the region of interaction' as 423, the RNAup server does not work for unknown reasons.

## Discussion

As one big class of important noncoding RNA molecules, lncRNAs play critical roles in a number of biological processes. They show a diversity of functions and mechanisms by binding

with diverse molecules, in which RNAs are one of the most important and popular ones. Uncovering the RNA targets of lncRNAs will be an important way to understanding the functions of lncRNAs and their roles in disease. Therefore, there is a great demand for developing efficient bioinformatics methods to globally predict the lncRNA–RNA interactions. Unfortunately, none of the current methods can satisfy the above purpose. In this study, we presented an efficient tool, LncTar, to predict the specific interaction between an lncRNA and its target RNAs. The main advantage of LncTar is its ability to efficiently predict the lncRNA–RNA interactions, which is important and necessary for the understanding of lncRNA functions and molecular mechanisms. The results confirmed that LncTar has a reliable accuracy. However, LncTar has some limitations. For example, it overlooks the stacked pair energy and loop energy in the process of searching the stable joint structures formed by interacting RNA pairs. Predicting joint secondary structures with all kinds of loop interactions is a NP-complete problem [37]. In addition, LncTar does not consider RNA tertiary structure, which could play roles in RNA–RNA interactions. Although it seems that base pairing plays a key role in lncRNA–RNA interaction according to current publications [8, 14], both the tertiary structure [38] and the secondary structure such as loop [38], stacking base pairs [9] and helices [39] have roles in RNA–RNA interactions. This may be the reason that two pairs of lncRNA–mRNA (EMX2OS::NM_004098 and NPPA-AS1::NPPA) were not successfully predicted by LncTar. Therefore, it is important to integrate the structure information to improve LncTar in the future. Although limitations exist, we believe that LncTar provides lncRNA researchers a valuable and efficient tool to predict the RNA targets of candidate lncRNAs.

**Key Points**

- We presented a method and developed a tool, LncTar, to predict the RNA targets of long noncoding RNAs.
- LncTar runs fast and therefore can be used for large-scale identification of the RNA targets for long noncoding RNAs.
- LncTar does not have limit to RNA size, indicating LncTar can be used to all RNAs.
- LncTar has a high prediction accuracy.

## Funding

This study was supported by National Nature Science Foundation of China (Nos. 91339106, 81422006) and the 863 project (2014AA021102).

## References

1. Kapranov P, Cawley SE, Drenkow J, *et al*. Large-scale transcriptional activity in chromosomes 21 and 22. *Science* 2002;**296**:916–19.
2. International Human Genome Sequencing C. Finishing the euchromatic sequence of the human genome. *Nature* 2004;**431**:931–45.
3. Serviss JT, Johnsson P, Grander D. An emerging role for long non-coding RNAs in cancer metastasis. *Front Genet* 2014;**5**:234.
4. Nagano T, Fraser P. No-nonsense functions for long noncoding RNAs. *Cell* 2011;**145**:178–81.
5. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. *Cell* 2009;**136**:629–41.
6. Mercer TR, Dinger ME, Sunkin SM, *et al*. Specific expression of long noncoding RNAs in the mouse brain. *Proc Natl Acad Sci USA* 2008;**105**:716–21.
7. Schonrock N, Harvey RP, Mattick JS. Long noncoding RNAs in cardiac development and pathophysiology. *Circ Res* 2012;**111**:1349–62.
8. Wapinski O, Chang HY. Long noncoding RNAs and human disease. *Trends Cell Biol* 2011;**21**:354–61.
9. Yajima S, Inoue S, Ogawa T, *et al*. Structural basis for sequence-dependent recognition of colicin E5 tRNase by mimicking the mRNA-tRNA interaction. *Nucleic Acids Res* 2006;**34**:6074–82.
10. Tian D, Sun S, Lee JT. The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell* 2010;**143**:390–403.
11. Han P, Li W, Lin CH, *et al*. A long noncoding RNA protects the heart from pathological hypertrophy. *Nature* 2014;**514**:102–6.
12. Faghihi MA, Modarresi F, Khalil AM, *et al*. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat Med* 2008;**14**:723–30.
13. Faghihi MA, Zhang M, Huang J, *et al*. Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol* 2010;**11**:R56.
14. Johnson R, Guigo R. The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. *RNA* 2014;**20**:959–76.
15. Wright PR, Georg J, Mann M, *et al*. CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. *Nucleic Acids Res* 2014;**42**:W119–23.
16. Gerlach W, Giegerich R. GUUGle: a utility for fast exact matching under RNA complementary rules including G-U base pairing. *Bioinformatics* 2006;**22**:762–64.
17. Kato Y, Sato K, Hamada M, *et al*. RactIP: fast and accurate prediction of RNA-RNA interaction using integer programming. *Bioinformatics* 2010;**26**:i460–6.
18. Muckstein U, Tafer H, Hackermuller J, *et al*. Thermodynamics of RNA-RNA binding. *Bioinformatics* 2006;**22**:1177–82.
19. Salari R, Backofen R, Sahinalp SC. Fast prediction of RNA-RNA interaction. *Algorithms Mol Biol* 2010;**5**:5.
20. Andronescu M, Zhang ZC, Condon A. Secondary structure prediction of interacting RNA molecules. *J Mol Biol* 2005;**345**:987–1001.
21. Bernhart SH, Tafer H, Muckstein U, *et al*. Partition function and base pairing probabilities of RNA heterodimers. *Algorithms Mol Biol* 2006;**1**:3.
22. Chen G, Wang Z, Wang D, *et al*. LncRNADisease: a database for long-non-coding RNA-associated diseases. *Nucleic Acids Res* 2013;**41**:D983–6.
23. Yuan J, Wu W, Xie C, *et al*. NPInter v2.0: an updated database of ncRNA interactions. *Nucleic Acids Res* 2014;**42**:D104–8.
24. Vallone PM, Butler JM. AutoDimer: a screening tool for primer-dimer and hairpin structures. *Biotechniques* 2004;**37**:226–31.
25. Shen Z, Qu W, Wang W, *et al*. MPprimer: a program for reliable multiplex PCR primer design. *BMC Bioinformatics* 2010;**11**:143.
26. Marshall OJ. PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics* 2004;**20**:2471–2.
27. Marshall O. Graphical design of primers with PerlPrimer. *Methods Mol Biol* 2007;**402**:403–14.
28. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 2000;**132**:365–86.

29. SantaLucia J, Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci USA* 1998;**95**:1460–5.
30. Allawi HT, SantaLucia J, Jr. Thermodynamics and NMR of internal G.T mismatches in DNA. *Biochemistry* 1997;**36**: 10581–94.
31. Allawi HT, SantaLucia J, Jr. Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: sequence dependence and pH effects. *Biochemistry* 1998;**37**:9435–44.
32. Sugimoto N, Nakano S, Yoneyama M, *et al*. Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Res* 1996;**24**: 4501–5.
33. SantaLucia J, Jr., Allawi HT, Seneviratne PA. Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry* 1996;**35**:3555–62.
34. Rehmsmeier M, Steffen P, Hochsmann M, *et al*. Fast and effective prediction of microRNA/target duplexes. *RNA* 2004;**10**: 1507–17.
35. Dimitrov RA, Zuker M. Prediction of hybridization and melting for double-stranded nucleic acids. *Biophys J* 2004;**87**: 215–26.
36. Markham NR, Zuker M. UNAFold: software for nucleic acid folding and hybridization. *Methods Mol Biol* 2008;**453**:3–31.
37. Alkan C, Karakoc E, Nadeau JH, *et al*. RNA-RNA interaction prediction and antisense RNA target search. *J Comput Biol* 2006;**13**:267–82.
38. Keel AY, Jha BK, Kieft JS. Structural architecture of an RNA that competitively inhibits RNase L. *RNA* 2012;**18**:88–99.
39. Yin J, Huang Q, Pakhomova ON, *et al*. The conserved adenosine in helix 6 of Archaeoglobus fulgidus signal recognition particle RNA initiates SRP assembly. *Archaea* 2004;**1**:269–75.