



InCeDB: Database of Human Long Noncoding RNA Acting as Competing Endogenous RNA

Shaoli Das¹*, Suman Ghosal¹*, Rituparno Sen², Jayprokas Chakrabarti^{1,2*}

1 Computational Biology Group, Indian Association for the Cultivation of Science, Kolkata, West Bengal, India, **2** Gyanxet, Kolkata, West Bengal, India

Abstract

Long noncoding RNA (lncRNA) influences post-transcriptional regulation by interfering with the microRNA (miRNA) pathways, acting as competing endogenous RNA (ceRNA). These lncRNAs have miRNA responsive elements (MRE) in them, and control endogenous miRNAs available for binding with their target mRNAs, thus reducing the repression of these mRNAs. InCeDB provides a database of human lncRNAs (from GENCODE 19 version) that can potentially act as ceRNAs. The putative mRNA targets of human miRNAs and the targets mapped to AGO clipped regions are collected from TargetScan and StarBase respectively. The lncRNA targets of human miRNAs (up to GENCODE 11) are downloaded from miRCode database. miRNA targets on the rest of the GENCODE 19 lncRNAs are predicted by our algorithm for finding seed-matched target sites. These putative miRNA-lncRNA interactions are mapped to the Ago interacting regions within lncRNAs. To find out the likelihood of an lncRNA-mRNA pair for actually being ceRNA we take recourse to two methods. First, a ceRNA score is calculated from the ratio of the number of shared MREs between the pair with the total number of MREs of the individual candidate gene. Second, the P-value for each ceRNA pair is determined by hypergeometric test using the number of shared miRNAs between the ceRNA pair against the number of miRNAs interacting with the individual RNAs. Typically, in a pair of RNAs being targeted by common miRNA(s), there should be a correlation of expression so that the increase in level of one ceRNA results in the increased level of the other ceRNA. Near-equimolar concentration of the competing RNAs is associated with more profound ceRNA effect. In InCeDB one can not only browse for lncRNA-mRNA pairs having common targeting miRNAs, but also compare the expression of the pair in 22 human tissues to estimate the chances of the pair for actually being ceRNAs. **Availability:** Downloadable freely from <http://gyanxet-beta.com/lncedb/>.

Citation: Das S, Ghosal S, Sen R, Chakrabarti J (2014) InCeDB: Database of Human Long Noncoding RNA Acting as Competing Endogenous RNA. PLoS ONE 9(6): e98965. doi:10.1371/journal.pone.0098965

Editor: Sandro Banfi, Telethon Institute of Genetics and Medicine, Italy

Received: November 19, 2013; **Accepted:** May 9, 2014; **Published:** June 13, 2014

Copyright: © 2014 Das et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors received financial support from the intramural funds of their host institutes. SD, SG and JC received financial support from the Indian Association for the Cultivation of Science. RS was hosted by Gyanxet, an independent non-profit body. Funds for the website were contributed by Gyanxet. The fund giving organizations, namely, the Indian Association for the Cultivation of Science and Gyanxet, had no role in the design, data collection and analysis, the decision to publish, or in the preparation of the manuscript.

Competing Interests: Rituparno Sen is an employee of Gyanxet, which hosts the database InCeDB, together with several freely downloadable software. There are no further patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials. The database is freely accessible from <http://gyanxet-beta.com/lncedb/>.

* E-mail: j.chakrabarti@gyanxet.com

† These authors contributed equally to this work.

Introduction

A major part of the transcriptome of higher eukaryote does not code for any protein, but functions as regulatory RNAs. Long noncoding RNAs (lncRNAs) are long transcripts (generally more than 200 bases) that in many ways are similar to protein coding transcripts, except for the lack of a meaningful coding sequence (CDS) or open reading frame (ORF) [1–2]. In the last decade it has become possible to annotate a greater number of lncRNA transcripts. Consequently the study of them has gained much significance as many of them have been linked with epigenetic, transcriptional and post-transcriptional regulation of gene expression [3]. Despite having generally a lower level of concentration than protein coding transcripts, lncRNAs exhibit more tissue specific expressions [4–5]. A vast set of lncRNA transcripts are differentially expressed during development where many of them play critical roles [6–9]. lncRNAs are now known to have a major involvement in cancer [10–11]. Though, till now, a majority of the lncRNAs have been linked with epigenetic modulation of gene expressions [12–13], they can also regulate gene expression by

transcriptional or post transcriptional modes [14–15]. lncRNAs can influence post-transcriptional regulation by interfering with the miRNA pathways, by acting as competing endogenous RNAs (ceRNAs) [16]. In recent years it has been discovered that endogenous RNAs (mRNAs, pseudogenes, long noncoding RNAs or circular RNAs) compete with each other for the limited pool of cellular microRNAs (miRNAs) and thus affect the competing RNA's level. These RNAs have miRNA responsive elements (MRE), i.e., the miRNA binding sites in them, and act as miRNA sponges to control endogenous miRNAs available for binding with their target mRNAs, thus reducing the repression of these mRNAs [17–24]. This phenomenon adds a significant new dimension to the miRNA mediated regulation of gene expression in cells. ceRNAs are important regulators in cell cycle control and tumor suppression (e.g. PTEN-P1 blocking miR-19b and miR-20a from binding to PTEN tumor suppressor [17–19]), modulating self-regulation in hepatocellular carcinoma (HCC) (HULC lncRNA acts as ceRNA of the protein coding gene PRKACB that induces activation of CREB which in turn is involved in upregulation of

HULC [20]) as well as in developmental stages (e.g. linc-MD1 blocking miR-133 from binding to transcription factors involved in myogenic differentiation [21] and H19 blocking the miRNA let-7 to affect muscle differentiation in vitro [22]). Circular RNAs have recently been shown to be involved in pathways of cancer and many other diseases [23,24].

Due to the availability of huge lncRNA datasets from recent GENCODE versions [25], 13870 lncRNA genes in GENCODE 19, it has become imperative to uncover the potential functions of these transcripts. In the light of new findings on ceRNAs and lncRNA-miRNA interactions, we developed a database, *lnCeDB*, of human lncRNAs that can potentially act as ceRNAs (based on the newly available GENCODE 19 annotated lncRNAs). Recently, databases describing lncRNA-miRNA interactions, like miRCode [26], Diana-lncBase [27], lncRNome [28] and StarBase v2.0 [29], have become available. But none of them documents miRNA interactions with lncRNAs annotated past GENCODE 17. We used lncRNA-mRNA interaction pairs from miRCode database of miRNA targets for lncRNAs in GENCODE 11 [26], and for the newly enlisted lncRNAs in GENCODE 19, we predicted seed-matched miRNA targets using our algorithm. We mapped these putative miRNA-lncRNA interactions into the Ago-interacting regions within lncRNAs, collected from a recent study [30]. In *lnCeDB*, the users can also browse for miRNA targets on recently available GENCODE 19 lncRNAs not available from other databases. Moreover, the objective of *lnCeDB* is not just describing lncRNA-miRNA interactions, but providing researchers with a database of human lncRNAs that can potentially act as ceRNAs to protein coding genes. The chances of an lncRNA-mRNA pair for actually being ceRNA depend not only on the fact that they are targeted by common miRNA(s), but also other factors like relative concentrations of individual component ceRNAs and the number of shared MREs. *lnCeDB* is built by taking into consideration these varied and complex features.

A previously published database of ceRNAs, ceRDB, provides data of mRNAs that can putatively act as ceRNAs, but it does not have information about lncRNAs [31]. It should also be noted that unlike the ceRDB database, that used putative miRNA-mRNA interactions predicted by TargetScan [32], we include mRNA-miRNA and miRNA-lncRNA interactions predicted from AGO CLIP-Seq data [30,33]. The user can limit the target search within regions of AGO interaction, significantly reducing false-positive target. Another recently published database, StarBase v2.0 include ceRNA pairs predicted from available AGO PAR-CLIP datasets [29]. However, the use of only the PAR-CLIP data for prediction of ceRNA pairs limits the result set to only a few cell lines where the AGO PAR-CLIP was performed. As mentioned earlier, our dataset includes, but is not limited to predictions from just the AGO CLIP-SEQ data. This gives the user a broader set of probable ceRNA pairs in many human tissues, and also the option to narrow down the search to only the AGO interacting regions as available from AGO CLIP-SEQ data.

As mentioned earlier, the chances of an lncRNA-mRNA pair actually being ceRNA depend not just on the fact that they are targeted by common miRNA(s), but also on the number of miRNA responsive elements (MRE), number of distinct miRNAs targeting both transcripts and the concentration and cellular levels of the competing RNAs. A study by Ala et al provides a comprehensive view on the ceRNA network and the possible outcome of perturbation in the components (miRNA and competing RNA levels) of the network [34]. The authors showed that the relative expressions of competing RNAs play a vital part in determining the ceRNA effect. While, for a pair of ceRNAs, it is seen that the competing RNA with higher expression has greater

ceRNA effect on the other competing RNA, it has also been observed that competing RNAs with near-equal expression exhibit more robust ceRNA effect than other ceRNA pairs having largely different expressions [34]. Thus, the information on the concentration levels of the two RNAs making the ceRNA pair is very crucial. Also, to determine the potential cross-regulation of a ceRNA pair, it is very important to check the co-expression of shared miRNAs along with the ceRNA pair. One major drawback of the existing ceRNA databases, other than *lnCeDB*, is that they do not offer the option to the users to check the co-expression of the ceRNA pair and the shared miRNAs. Following from the observation by Ala et al, to estimate the chances of an lncRNA-mRNA pair for actually being ceRNAs in particular tissues, *lnCeDB* offers users the possibility to browse for lncRNA-mRNA pairs targeted by common miRNAs (sorted by the number of shared miRNAs of the pair) and compare the expression of the pair in 22 human tissues (from RNA-Seq expression data from Cabili et al [35]). Moreover, *lnCeDB* also provides users with the information on the shared miRNAs co-expressed in each of the 22 different tissues. This feature is not offered by any other ceRNA database.

To assess the likelihood of an lncRNA-mRNA pair to act as ceRNAs, we provide two different methods. In the first approach, we calculate the P-value for each ceRNA pair by hypergeometric test, similar to the study by Sumazin et al [19] and StarBase v2.0 [29], considering the number of shared miRNAs between the pair against the total number of miRNAs targeting each component in the pair, i.e. the lncRNA and the mRNA. But in the second approach, unlike other ceRNA databases that predict the likelihood for being ceRNAs by the number of shared miRNAs [19,29], we calculate a ceRNA score for each probable ceRNA pair by taking into consideration the number of shared MREs against the total number of MREs for the candidate lncRNA. A major drawback of the other ceRNA databases [29] is that they calculate the likelihood of a pair of genes to act as ceRNA by considering only the number of shared miRNAs between the pair. But there is a certain importance to the number of shared MREs compared to the number of shared miRNAs between the ceRNA pair. Thus, a candidate lncRNA having 100 MREs for 1 shared miRNA would be considered lesser than a candidate lncRNA with 2 MREs for 2 shared miRNAs in some of the existing databases. We believe that the number of shared MREs would be more appropriate instead of the number of shared miRNAs between the ceRNA pair. And that makes *lnCeDB* different from the existing databases on ceRNA. In *lnCeDB*, the ceRNA score, along with the provision for checking relative expressions of the ceRNA pair over different tissues, offer users a better assessment of the potential ceRNAs. We believe, therefore, that *lnCeDB* addresses the finer aspects of post transcriptional gene expression regulation in human.

Results

lnCeDB contains dataset for human genome wide lncRNAs that potentially can act as ceRNAs. Presently we have miRNA targets on 22286 lncRNAs out of 25978 lncRNA transcripts in GENCODE 19. miRNA targets for lncRNAs in GENCODE 11 were collected from miRCode database and targets for newly enlisted lncRNAs in GENCODE 19 were predicted by our algorithm (see section 2). Among the 22286 lncRNAs with putative miRNA target sites, 15842 lncRNAs were predicted to be working as potential ceRNAs. Figure 1 shows a statistics of the database contents in *lnCeDB*.

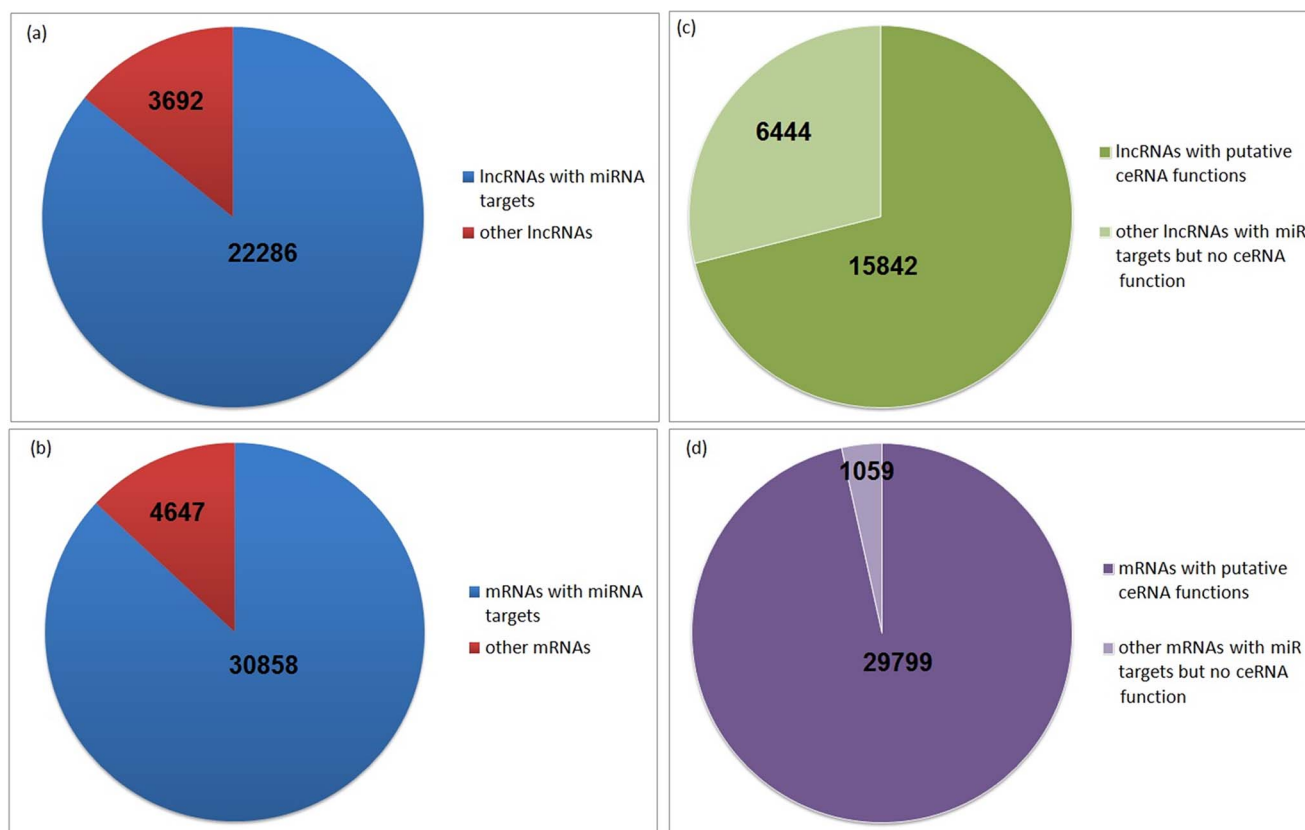


Figure 1. A statistics of the database contents in InCeDB. (a) Fraction of GENCODE 19 lncRNA transcripts with putative miRNA targets. (b) Fraction of mRNAs with predicted miRNA targets. (c) Fraction of lncRNAs with predicted ceRNA function compared to all lncRNAs with putative miRNA targets. (d) Fraction of mRNAs with predicted ceRNA function compared to all mRNAs with putative miRNA targets. doi:10.1371/journal.pone.0098965.g001

lncRNA-miRNA interactions mapping into Ago interaction sites within lncRNA loci

Among the putative lncRNA-miRNA interactions, 72 interaction pairs were mapped to the Ago-interaction sites within lncRNA transcript loci which contained 41 miRNAs targeting 27 lncRNA transcripts. ceRNAs specifically corresponding to these PAR-CLIP supported miRNA target interactions are browsable in InCeDB.

Putative mRNA-miRNA interactions from TargetScan

Conserved human miRNA targets on mRNAs were collected from TargetScan [32]. The current version stores miRNA targets on 30858 mRNA transcripts for human.

mRNA-miRNA interactions derived from Ago interacting regions within mRNA

miRNA targets on 6090 mRNAs were collected from StarBase; these targets were predicted from CLIP-Seq data by five most widely used miRNA target prediction tools (see section 2).

Expression data for a ceRNA pair over 22 tissues

Users can compare the cellular levels of the competing RNAs to estimate the chances of the pair for actually being ceRNAs by viewing the tissue specific expression of a ceRNA pair over 22 human tissues. Presently we have tissue specific expressions for 7017 lncRNA genes and 17701 protein coding genes.

Experimentally verified lncRNAs acting as ceRNAs

A small number of lncRNAs are reported and experimentally verified as ceRNAs. Experimentally validated lncRNA ceRNAs are identified in InCeDB. The reported lncRNA-ceRNAs include lncRNA HULC acting as ceRNA of the gene PRKACB [20], lincRNA MD1 working as ceRNA of the gene MAML1 [21] and linc-RoR acting as ceRNA of the pluripotency associated transcription factors SOX2 and NANOG [36]. Table 1 lists the experimentally verified lncRNA ceRNAs in InCeDB.

Utility

Users can browse for ceRNA candidates for a protein coding gene (by gene symbol, gene id or refseq accession) and/or lncRNA gene (by gene name, ensemble gene id or ensemble transcript id) and/or miRNA name. lncRNA-mRNA pairs targeted by common miRNAs are displayed as potential ceRNAs. The list of ceRNAs is sorted by the number of shared miRNAs. The miRNA targets for a particular lncRNA or mRNA can also be browsed in our database by choosing a particular lncRNA/mRNA from the ceRNA list. By choosing a particular record for a ceRNA pair, users can view the RNA-Seq expressions of the lncRNA-mRNA pair over 22 human tissues in form of a heat map (if the expression data for the corresponding lncRNA and mRNA is available) and a bar chart with actual FPKM values appearing in tool tip text. The users can also view the co-expressed miRNAs shared between the ceRNA pair in each of the 22 tissues. Figure 2 shows a flow diagram for the usability of InCeDB.

Table 1. Experimentally verified lncRNA ceRNAs in InCeDB.

lncRNA acting as ceRNA	Competing protein coding gene	Shared miRNA	ceRNA score	Reference
HULC (Highly Upregulated in Liver Cancer)	PRKACB	miR-372	0.026 (p-value = 0.001)	[20]
lincRNA MD1	MAML1	miR-133	0.022 (p-value = 0.02)	[21]
H19	Targets of hsa-let-7	Let-7	-	[22]
Linc-RoR (Regulator of Reprogramming)	SOX2 and NANOG	miR-145	0.038 (p-value = 0.008)	[36]
PTCSC3	Targets of miR-574-5p in thyroid cancer cell line	miR-574-5p	-	[37]

doi:10.1371/journal.pone.0098965.t001

Discussion

As competing endogenous RNAs are crucial new determinants of gene expression regulation [38], new data sources are needed. Following the availability of huge dataset of annotated lncRNA transcripts from GENCODE project, the possible functions of the transcripts have to be addressed. In InCeDB, we have tried to explore the potential sponge activity of recently annotated lncRNAs in the miRNA mediated gene regulation networks at human genome wide scale. Unlike other ceRNA databases, InCeDB includes but is not limited to miRNA targets on protein coding and lncRNA transcripts predicted from Ago interaction sites within them. The advantage is that it reduces the false positive target detection in our miRNA-target interaction dataset and enhances the reliability of the prediction. At the same time, the dataset is not limited to predictions from only a few cell lines where AGO PAR-CLIP were performed. Also, in InCeDB we considered, for the first time, that ceRNA activity largely depends on the relative concentration of the components of a ceRNA network, i.e. the pair of competing RNAs and also the miRNAs they compete for. The provision for checking the tissue specific expression for a potential ceRNA pair (whenever available) along with the co-expression of shared miRNAs, gives the user a higher chance of identifying the most likely ceRNA candidates in a tissue of interest.

One interesting example of a putative ceRNA pair identified by InCeDB is the lncRNA maternally expressed 3 (MEG3) and a transcript (the longer isoform) of the protein coding gene Myeloid Cell Leukemia Sequence 1 (MCL1). MEG3 is a maternally expressed imprinted gene encoding a number of alternatively spliced lncRNA transcripts. It interacts with the tumour suppressor P53, and is supposedly a tumour suppressor itself. MEG3 is expressed in many normal tissues including breast, colon, liver, ovary but its expression is lost in many tumour cells. Interestingly, it has been reported that MEG3 is targeted by miRNAs [39]. MCL1 is a member of the BCL2 family and it has three isoforms. The longer isoform is anti-apoptotic whereas the shorter isoforms are pro-apoptotic. The ceRNA pair MEG3-MCL1 putatively shares 16 common miRNAs including miR-28, miR-181d, miR-520a, miR-520b and miR-876-3p and show comparable high co-expressions in breast and colon, especially in colon (MEG3 (66.2): MCL1 (37.9)). The co-expression pattern of MEG3 and MCL1, along with the co-expressed shared miRNAs indicates that MEG3 may act as a ceRNA to MCL1 in colon. Interestingly, the MEG3 gene locus has been reported to be hypermethylated in colorectal cancer cells [40] indicating the possible perturbation of MEG3 lncRNA expression in colorectal carcinoma. Furthermore, in a colorectal cancer cell, the anti-apoptotic MCL1 has been reported to be regulated (in vitro) by a number of miRNAs, including miR-876-3p [41] which we predicted to be shared by MEG3 and MCL1. Together, these observations suggest that there may be a

disruption of the potential MEG3-MCL1 ceRNA network in colorectal cancer cells as opposed to the normal colon cells. This observation, however, needs to be validated by further investigations. This example shows the importance of InCeDB over other ceRNA databases as no other ceRNA database allows the users to check the co-expression patterns of the competing RNAs and the shared miRNAs in different tissue types. Some other interesting observations from InCeDB are MEG3 and CMPK1 as potential ceRNA pair with near-equal expression signature in colon and ovary and MALAT1-PRKACB potential ceRNA pair in liver.

We believe this database will help researchers in deciphering the larger and more complex scenario of miRNA mediated gene regulatory networks in human in the real world of ceRNAs.

Materials and Methods

Data Source

We collected predicted miRNA target sites on human mRNA transcripts from TargetScan. The miRNA-mRNA interactions mapped to the AGO interacting regions were collected from StarBase database [33]. StarBase houses datasets of mRNA-miRNA interactions predicted from AGO CLIP-Seq data (AGO interacting regions) by widely used miRNA target prediction tools like TargetScan [32], PITA, PicTar, RNA22 and miRanda. The lncRNA-miRNA interaction dataset (for lncRNAs in GENCODE 11 version) was downloaded from miRCode database [26]. For newly enlisted lncRNAs in GENCODE 19 version, we used our miRNA-target prediction algorithm for finding seed matched miRNA target sites on lncRNAs. The dataset for Ago interacting regions within lncRNA transcript loci was collected from Jalali et al [30]. Tissue specific expressions (RNA-Seq) for lncRNAs and mRNAs in 22 human tissues were used from the dataset of Cabili et al [35]. Tissue specific miRNA expressions (QRT-PCR) were collected from miRNA body-map [42].

Prediction of miRNA target sites on lncRNAs

We used a 25 base window on the target to run a modified version of the Smith-Waterman alignment (as used in the miRanda algorithm [43]) to find complementary alignment of the 25 base target region with the miRNA. For reducing runtime, at the first step we searched for 6-mer seed region (position 2-7 on miRNA) complementarities on the target. If a match was found, then a 25 base window around the seed-matched site was used for further alignment. For prediction of different types of miRNA target sites (6-mer, 7-mer, 7-merA1 and 8-mer), we changed the seed region definition in each case [43]. We considered transcripts with perfect seed complementarity as well as one base mismatch tolerance (in position 2- 8 or 2-7) in the seed region with 3' compensatory complementarity (at position 13-18).

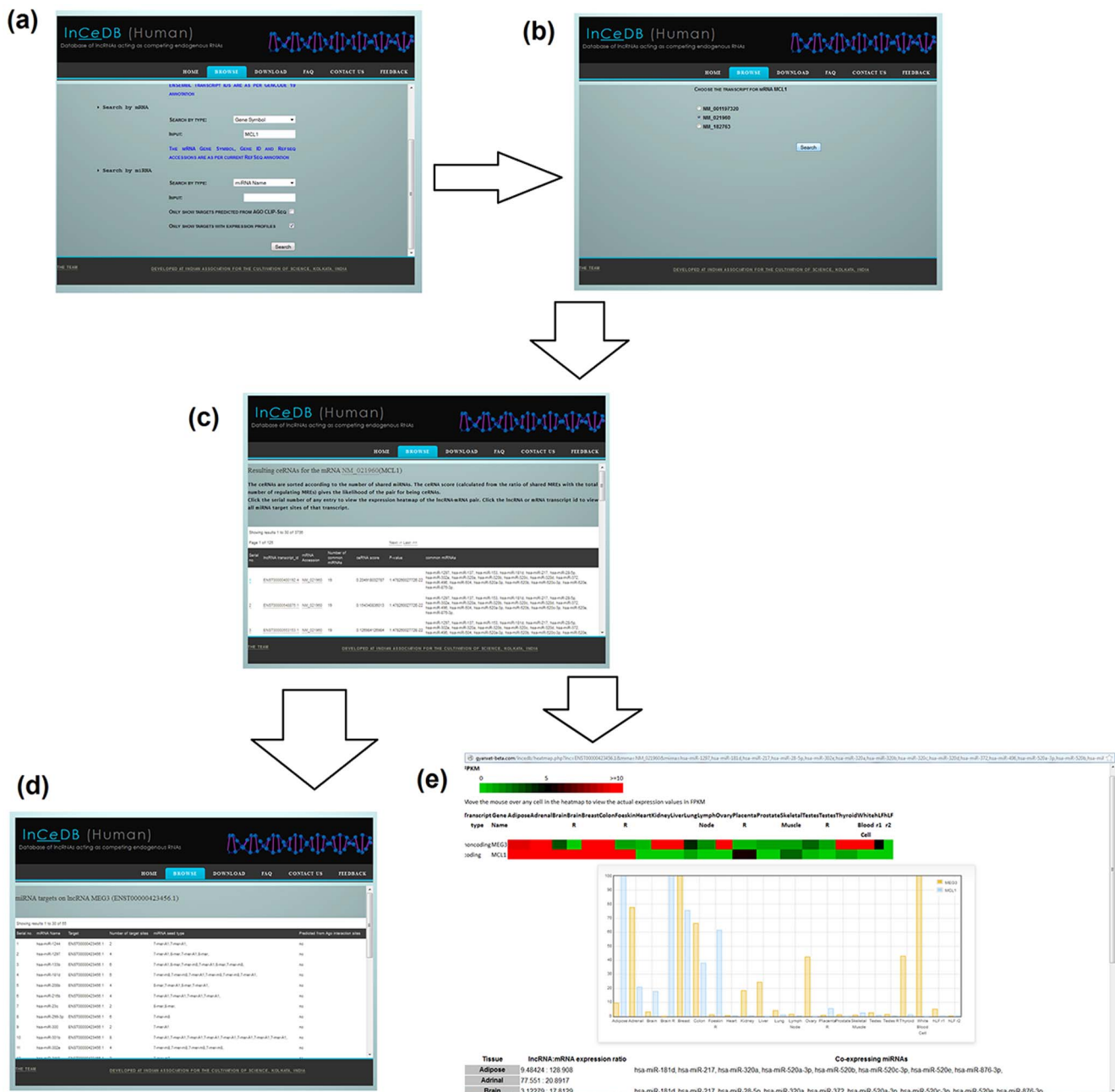


Figure 2. The flow diagram for browsing ceRNAs in InCeDB. (a) InCeDB is searched by the gene symbol MCL1. (b) An intermediate page shows the different transcripts of the gene MCL1. (c) Upon choosing a transcript, the results page shows the potential lncRNAs working as ceRNA for the chosen transcript of MCL1. (d) By clicking on a lncRNA or mRNA id in the ceRNA table, the user views all miRNA targets on the chosen transcript. (e) By clicking on a serial number in the ceRNA table, the user views the expression heatmap and bar chart for the ceRNA pair along with co-expressed shared miRNAs in 22 human tissues. doi:10.1371/journal.pone.0098965.g002

Consideration for target site conservation

We treated conserved and non-conserved target sites separately. We searched for target region conserved among human, chimp, mouse, rat and dog. Genome wide conservation data generated using multiz 46-way alignment (for 46 vertebrate species) was downloaded from UCSC genome browser [44]. Genomic regions (within human genome) conserved within human, chimp, mouse, rat and dog were then extracted and mapped within the coordinates of human mRNAs (downloaded from UCSC genome browser) to get the location of the conserved regions within human

mRNAs. Conserved regions of length 8 bases or more were only considered.

The ceRNA score

We generated the ceRNA score of an lncRNA-mRNA pair targeted by common miRNAs to measure the likelihood of a lncRNA to act as ceRNA to a protein coding gene. The ceRNA score was calculated keeping in mind the observation by Ala et al [34] that genes that shared a large number of distinct miRNAs, as opposed to genes targeted by a small number of shared miRNAs,

compared to the total number of miRNAs targeting the individual genes, exhibit more profound ceRNA

$$ceRNAscore = \frac{\text{the number of MREs for the distinct shared miRNAs between the pair}}{\text{the total number MREs for all distinct miRNAs targeting the lncRNA}} \quad (1)$$

Calculating the probability of ceRNA pair to cross-regulate each other

We implemented another measure to assess the likelihood of a ceRNA pair to regulate each other via shared miRNAs. This approach was similar to what has been used in the study of Sumazin et al [19] and in StarBase v2.0 [29]. We calculated the p-value for each potential ceRNA pair by hypergeometric test considering the number of shared miRNAs between a ceRNA pair against the number of miRNAs targeting individual components of the ceRNA pair. The p-value was measured as:

$$p = \sum_{i=m_c}^{\min(m_p, m_n)} \frac{\binom{m_n}{i} \binom{M_T - m_n}{m_p - i}}{\binom{M_T}{m_p}} \quad (2)$$

Where,

References

- Ponting CP, Oliver PL, Reik W (2009) Evolution and functions of long noncoding RNAs, Cell 136: 629–641.
- Ponjavic J, Ponting CP, Lunter G (2007) Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. Genome Res. 17: 556–565.
- Mercer TR, Dinger ME, Mattick JS (2009) Long non-coding RNAs: insights into functions. Nat Rev Genet. 10: 155–159.
- Pang KC, Frith MC, Mattick JS (2006) Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. Trends Genet. 22: 1–5.
- Marques AC, Ponting CP (2009) Catalogues of mammalian long noncoding RNAs: modest conservation and incompleteness. Genome Biol. 10: R124.
- Amaral PP, Mattick JS (2008) Noncoding RNA in development. Mamm Genome. 19: 454–492.
- Meola N, Pizzo M, Alfano G, Surace EM, Banfi S (2012) The long noncoding RNA Vax2os1 controls the cell cycle progression of photoreceptor progenitors in the mouse retina. RNA 18: 111–123.
- Paralkar VR, Weiss MJ (2011) A new 'Linc' between noncoding RNAs and blood development. Genes Dev. 25: 2555–2558.
- Ghosal S, Das S, Chakrabarti J (2013) Long noncoding RNAs: new players in the molecular mechanism for maintenance and differentiation of pluripotent stem cells. Stem Cells Dev. 22:2240–2253.
- Aguilo F, Zhou MM, Walsh MJ (2011) Long noncoding RNA, polycomb, and the ghosts haunting INK4b-ARF-INK4a expression. Cancer Res. 71: 5365–5369.
- Lv XB, Lian GY, Wang HR, Song E, Yao H, et al. (2013) Long Noncoding RNA HOTAIR Is a Prognostic Marker for Esophageal Squamous Cell Carcinoma Progression and Survival. PLoS One. 8: e63516.
- Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 322: 750–756.
- Ng SY, Johnson R, Stanton LW (2011) Human long noncoding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. EMBO J. 31: 522–533.
- Hawkins PG, Morris KV (2010) Transcriptional regulation of Oct4 by a long noncoding RNA antisense to Oct4-pseudogene 5. Transcription 1: 165–175.
- Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, et al. (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell 39: 925–938.
- Salmena L, Polisenio L, Tay Y, Kats L, Pandolfi PP (2011) A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell 146: 353–358.
- Tay Y, Kats L, Salmena L, Weiss D, Tan SM, et al. (2011) Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. Cell 147: 344–357.
- Karreth FA, Tay Y, Perna D, Ala U, Tan SM, et al. (2011) In vivo identification of tumor-suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. Cell 147: 382–395.
- Sumazin P, Yang X, Chiu HS, Chung WJ, Iyer A, et al. (2011) An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma. Cell 147: 370–381.
- Wang J, Liu X, Wu H, Ni P, Gu Z, et al. (2010) CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. Nucleic Acids Res. 38: 5366–5383.
- Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, et al. (2011) A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell 147: 358–369.
- Kallen AN, Zhou XB, Xu J, Qiao C, Ma J, et al. (2013) The imprinted H19 lncRNA antagonizes let-7 microRNAs. Mol Cell. 52:101–112.
- Hansen TB, Kjems J, Damgaard CK (2013) Circular RNA and miR-7 in liver cancer. Cancer Res. 73, 5609–5612.
- Ghosal S, Das S, Sen R, Basak P, Chakrabarti J (2013) Circ2Traits: A comprehensive database for circular RNA potentially associated with disease and traits. Front Genet. Doi: 10.3389/fgene.2013.00283.
- Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, et al. (2012) The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. Genome Res 22: 1775–1789.
- Jeggari A, Marks DS, Larsson E (2012) miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. Bioinformatics. 28: 2062–2063.
- Paraskevopoulou MD, Georgakilas G, Kostoulas N, Reczko M, Maragkakis M, et al. (2013) DIANA-LncBase: experimentally verified and computationally predicted microRNA targets on long non-coding RNAs. Nucleic Acids Res. 41(Database issue): D239–45.
- Bhartiya D, Pal K, Ghosh S, Kapoor S, Jalali S, et al. (2013) lncRNome: a comprehensive knowledgebase of human long noncoding RNAs. Database (Oxford). 2013:bat034.
- Li JH, Liu S, Zhou H, Qu LH, Yang JH (2014) starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res. 42:D92–D97.
- Jalali S, Bhartiya D, Lalwani MK, Sivasubbu S, Scaria V (2013) Systematic transcriptome wide analysis of lncRNA-miRNA interactions. PLoS One. 8: e53823.
- Sarver AL, Subramanian S. (2012) Competing endogenous RNA database. Bioinformatics. 8: 731–733.
- Grimson A, Farh KK, Johnston WK, Garrett-Engle P, Lim LP, et al. (2007) MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. Molecular Cell. 27: 91–105.

33. Yang JH, Li JH, Shao P, Zhou H, Chen YQ, et al. (2011) starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. *Nucleic Acids Res.* 39(Database issue):D202–D209.
34. Ala U, Karreth FA, Bosia C, Pagnani A, Taulli R, et al. (2013) Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments. *Proc Natl Acad Sci U S A.* 110: 7154–7159.
35. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, et al. (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 25: 1915–1527.
36. Wang Y, Xu Z, Jiang J, Xu C, Kang J, et al. (2013) Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. *Dev Cell.* 25: 69–80.
37. Fan M, Li X, Jiang W, Huang Y, Li J, et al. (2013) A long non-coding RNA, PTCSC3, as a tumor suppressor and a target of miRNAs in thyroid cancer cells. *Exp Ther Med.* 5: 1143–1146.
38. Tay Y, Rinn J, Pandolfi PP (2014) The multilayered complexity of ceRNA crosstalk and competition. *Nature.* 505:344–352.
39. Braconi C, Kogure T, Valeri N, Huang N, Nuovo G, et al. (2011) microRNA-29 can regulate expression of the long non-coding RNA gene MEG3 in hepatocellular cancer. *Oncogene.* 30:4750–4756.
40. Menigatti M, Staiano T, Manser CN, Bauerfeind P, Komljenovic A, et al. (2013) Epigenetic silencing of monoallelically methylated miRNA loci in precancerous colorectal lesions. *Oncogenesis.* 2:e56.
41. Lam LT, Lu X, Zhang H, Lesniewski R, Rosenberg S, et al. (2010). A microRNA screen to identify modulators of sensitivity to BCL2 inhibitor ABT-263 (navitoclax). *Mol Cancer Ther.* 9:2943–2950.
42. Mestdagh P, Lefever S, Pattyn F, Ridzon D, Fredlund E, et al. (2011) The microRNA body map: dissecting microRNA function through integrative genomics. *Nucleic Acids Res.* 39:e136.
43. Enright AJ, John B, Gaul U, Tuschl T, Sander C, et al. (2003) MicroRNA targets in *Drosophila*. *Genome Biol.* 5: R1
44. Karolchik D, Hinrichs AS, Kent WJ (2012) The UCSC Genome Browser. *Curr Protoc Bioinformatics.* 40: 1.4.1–1.4.33.