

Systematic Transcriptome Wide Analysis of lncRNA-miRNA Interactions

Saakshi Jalali¹*, Deeksha Bhartiya¹*, Mukesh Kumar Lalwani², Sridhar Sivasubbu^{2*}, Vinod Scaria^{1*}

1 GN Ramachandran Knowledge Center for Genome Informatics, CSIR-Institute of Genomics and Integrative Biology (CSIR-IGIB), Delhi, India, **2** Genomics and Molecular Medicine, CSIR-Institute of Genomics and Integrative Biology (CSIR-IGIB), Delhi, India

Abstract

Background: Long noncoding RNAs (lncRNAs) are a recently discovered class of non-protein coding RNAs, which have now increasingly been shown to be involved in a wide variety of biological processes as regulatory molecules. The functional role of many of the members of this class has been an enigma, except a few of them like Malat and HOTAIR. Little is known regarding the regulatory interactions between noncoding RNA classes. Recent reports have suggested that lncRNAs could potentially interact with other classes of non-coding RNAs including microRNAs (miRNAs) and modulate their regulatory role through interactions. We hypothesized that lncRNAs could participate as a layer of regulatory interactions with miRNAs. The availability of genome-scale datasets for Argonaute targets across human transcriptome has prompted us to reconstruct a genome-scale network of interactions between miRNAs and lncRNAs.

Results: We used well characterized experimental Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP) datasets and the recent genome-wide annotations for lncRNAs in public domain to construct a comprehensive transcriptome-wide map of miRNA regulatory elements. Comparative analysis revealed that in addition to targeting protein-coding transcripts, miRNAs could also potentially target lncRNAs, thus participating in a novel layer of regulatory interactions between noncoding RNA classes. Furthermore, we have modeled one example of miRNA-lncRNA interaction using a zebrafish model. We have also found that the miRNA regulatory elements have a positional preference, clustering towards the mid regions and 3' ends of the long noncoding transcripts. We also further reconstruct a genome-wide map of miRNA interactions with lncRNAs as well as messenger RNAs.

Conclusions: This analysis suggests widespread regulatory interactions between noncoding RNAs classes and suggests a novel functional role for lncRNAs. We also present the first transcriptome scale study on miRNA-lncRNA interactions and the first report of a genome-scale reconstruction of a noncoding RNA regulatory interactome involving lncRNAs.

Citation: Jalali S, Bhartiya D, Lalwani MK, Sivasubbu S, Scaria V (2013) Systematic Transcriptome Wide Analysis of lncRNA-miRNA Interactions. PLoS ONE 8(2): e53823. doi:10.1371/journal.pone.0053823

Editor: Lin Zhang, University of Pennsylvania School of Medicine, United States of America

Received: February 27, 2012; **Accepted:** December 6, 2012; **Published:** February 6, 2013

Copyright: © 2013 Jalali 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 acknowledge funding from the Council of Scientific and Industrial Research (CSIR), India through the NWP0036 and FAC002 Grants. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: vinods@igib.res.in (VS); s.sivasubbu@igib.res.in (SS)

† These authors contributed equally to this work.

Introduction

Recent advances in sequencing technologies have enabled genome-scale mapping of transcriptional potential feasible at single-nucleotide resolutions [1,2]. A number of reports in the immediate past have hinted at the pervasive transcriptional potential of eukaryotic genomes [3], in addition to implicating a large number of novel genomic loci which were previously not known to encode any functional transcripts [4]. Though consistent evidence now suggests the existence of a large number of previously uncharacterized transcripts, a vast majority of the transcripts surprisingly do not code for a functional protein, and are grouped into a general class of noncoding RNAs (ncRNAs) [5–7]. The amenability of transcriptome mapping on a genome scale has enormously added to the number of ncRNAs, with members in the class presently outnumbering protein coding genes by a few folds. Recent efforts globally, including collaborative efforts such as the ENCODE [8] have been successful in characterizing the

transcript structures and their expression profiles in a wide number of cell types. This has also been supplemented by the availability of computational tools to map and identify transcript isoforms and quantify expression profiles in different cell types and conditions [9].

The ncRNA class now encompasses a wide diversity of subclasses organized based on their size, structure, function and conservation. MicroRNAs (miRNAs) have been one of the recently discovered and well characterized classes of ncRNAs and are small regulatory RNA molecules processed from larger precursors through a highly coordinated pathway [10,11]. Presently they are known to mediate post-transcriptional control of gene expression by binding to the 3'-untranslated regions of protein coding genes [12–14]. This is modulated through a ribonucleoprotein complex called the RNA induced silencing complex (RISC) [15]. Argonaute (Ago) proteins are the catalytic component of RISC complex. The small ncRNA bound to Ago

Table 1. The number of reads mapping in each of the individual datasets.

Protein Name	Read Length	Total No. of Reads	Reads mapped
Ago1	32	4155911	104075
Ago2	32	4588647	197799
Ago3	32	4122692	146122
Ago4	32	6434465	171212

doi:10.1371/journal.pone.0053823.t001

guides it to sequence complementary target sites [16]. The regulatory role of miRNAs encompass a wide variety of biological processes including development, organogenesis [17,18] and pathophysiology of a number of diseases including neuronal [19], cardiovascular [20] and infectious diseases [21–23], just to name a few. Another major class of recently discovered ncRNAs is the long noncoding RNAs (lncRNAs) [24,25]. Unlike the miRNAs, lncRNAs are longer, and by definition >200 nucleotides in length and usually show less sequence conservation. lncRNAs presently also encompass the previously known classes of processed pseudogenes, antisense transcripts and the recently discovered regulatory large intergenic noncoding RNAs (lincRNAs) [26]. Though a small number of lncRNAs have been functionally characterized like HOTAIR [27,28], Xist [29] and Malat [30], a large number of members in the class remain functionally uncharacterized. Presently well-characterized members of the class have been shown to be involved in regulatory roles in diverse processes like imprinting, X-inactivation and development. In addition, lncRNAs are also known to be associated with pathogenesis of a number of disease [31], recent reports from our lab also suggests potential of lncRNAs to be processed to smaller RNAs [32].

Though much of the focus in ncRNA research is directed towards understanding the regulation of protein-coding genes mediated by them, it has been suggested previously that ncRNAs could form a well-orchestrated regulatory interaction network [33]. There have been reports previously, which have suggested examples of such regulation, for example miRNA-miRNA interaction [34]. In addition, computational prediction of miRNA target sites suggest a widespread network of miRNA-lncRNA interaction [35]. Recently there have also been reports suggesting the possibility of a widespread interaction network involving competitive endogenous RNAs (ceRNAs) where ncRNAs could modulate regulatory RNA by binding and titrating them off their binding sites on protein coding messengers [36]. An example for this type of regulation is exemplified by HULC, an lncRNA expressed in hepatocellular carcinoma, which binds miR-372 and forms a regulatory interaction [37]. Similarly reports suggest that linc-MD1 could interact with miR-133 and miR-135 and promote muscle differentiation [38]. Similar examples have been described in plant systems and encompass target mimicry [39,40]. The entire spectrum of ncRNA regulatory layer, especially the possibility of regulatory miRNAs-lncRNAs interactions remains unexplored.

We employed datasets in public domain for lncRNA annotations recently made available from the ENCODE project along with genome-wide cross-linking with immunoprecipitation (CLIP-Seq) datasets for four Ago proteins [41]. We hypothesized that miRNAs and lncRNAs could form an interacting regulatory layer exemplifying the ceRNA hypothesis where multiple targets with multiple binding affinities compete for the regulatory RNAs and the members could modulate the regulation of each other by

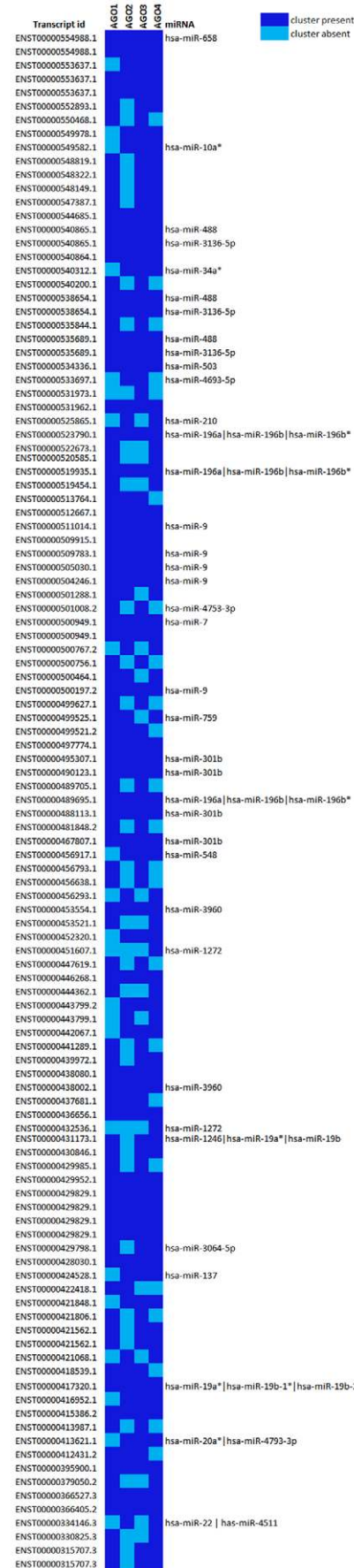


Figure 1. Depicting lncRNA transcripts targeted by miRNAs. The figure shows the lncRNA transcripts targeted by miRNAs along with the distribution of Ago types in the targeting complex. The presence and absence of specific Ago type in the RISC complex is represented by the difference in the colors in the heatmap. The targeting miRNA has also been mentioned.
doi:10.1371/journal.pone.0053823.g001

modulation of their relative concentrations (Figure S1). We show that a subset of the annotated lncRNAs could indeed harbor miRNA recognition elements and be a part of the miRNA interaction network, thus working as a critical regulatory layer of interaction. In addition, we have experimentally validated the interaction of *7sl* lncRNA with miR-125b in a zebrafish model, thus adding confidence to our hypothesis. Furthermore, we have also integrated miRNA-target information to suggest a network of interactions between lncRNAs, miRNAs and protein-coding transcripts. To our knowledge this is one of the first transcriptome scale study on miRNA-lncRNA interactions and the first report of a genome-scale reconstruction of a noncoding RNA regulatory interactome involving lncRNAs.

Results and Discussion

Identification of High-confidence microRNA Recognition Elements Mapping to Long Noncoding RNAs

The PAR-CLIP reads for HEK293 cells were downloaded from the NCBI Short Read Archive (SRA IDs: SRX020783, SRX020784, SRX020785, SRX020786), which amounted to a total of 19,301,715 CLIP-seq reads for Ago 1-4. The reads were further mapped to hg19 version of the human genome using a quality-aware reference mapping algorithm: Mapping and Assembly with qualities (MAQ). Mapping was performed with stringent criteria and allowed no mismatches. The Agos significantly overlap in their biological function. Ago1 has been shown to be involved in RNA-mediated post-transcriptional gene silencing, Ago2 shows slicer activity, Ago3 and Ago4 are required for RNA-mediated gene silencing (RNAi). Since there have been no significant differences reported between the mappings of reads for each of the Ago datasets, and since biologically their functions overlap, we merged the datasets as previously used for the analysis

[42]. A total of 619,208 hits were obtained for the Ago datasets. The number of reads mapping in each of the individual datasets is available in Table 1. To weed out any potential noisy loci, we applied stringent criteria that all loci should be supported by at least 5 overlapping reads. The mappings were parsed to identify high-quality miRNA recognition elements using custom scripts. We defined the recognition element as a stretch of contiguous positions in the genome, which matched the 5× read mapping criteria. A total of 15,983 such consensus high confidence miRNA recognition element clusters were identified. These mapped to 113 lncRNA exons and 4,367 mRNA exons from the GENCODE annotation datasets [43]. We further also analyzed the relative contributions of each of the Ago datasets for each of the high-confidence miRNA recognition elements.

Identification of the Cognate microRNAs for the microRNA Recognition Elements

Ago proteins are a component of the RNA induced silencing complex and are intricate components modulating the function of the complex. CLIP-seq allows one to identify the miRNA recognition elements in the transcriptome. To understand the biological interactions between miRNAs and their cognate target transcripts, it is imperative to identify the miRNAs corresponding to each of the miRNA recognition elements. We used miRanda [44] a popular computational algorithm to identify miRNAs binding to cognate miRNA recognition elements. Analysis revealed 51 miRNA-lncRNA interactions with 29 miRNAs targeting 25 lncRNA transcripts (Figure 1) (Table 2). We also predicted miRNA targets for all GENCODE annotated transcripts predicted to be bound with Agos. Out of the total 4,367 mRNA clusters, 416 protein coding transcripts were found to be targeted by 1,253 miRNA (Figure S2). The paucity in the number of known human miRNAs from the present analysis could potentially be an effect of the paucity in the biological understanding of the repertoire of smallRNA species in humans and their biological mechanisms. Apart from miRNAs, other smallRNA species also has potential to be associated with Ago proteins [45]. We hope future in-depth understanding of smallRNA species, including their biogenesis and mechanism of action could potentially throw

Table 2. Contribution of each Ago class in targeting lncRNA.

Argonaute	Number of Ago binding sites present in lncRNA	Number of lncRNA transcripts	miRNAs
Ago1	91	80	hsa-miR-1246, hsa-miR-196a, hsa-miR-196b, hsa-miR-196b*, hsa-miR-19a*, hsa-miR-19b-1*, hsa-miR-19b-2*, hsa-miR-301b, hsa-miR-3064-5p, hsa-miR-3136-5p, hsa-miR-3960, hsa-miR-4753-3p, hsa-miR-488, hsa-miR-503, hsa-miR-658, hsa-miR-7, hsa-miR-759, hsa-miR-9
Ago2	75	65	has-miR-4511, hsa-miR-10a*, hsa-miR-137, hsa-miR-196a, hsa-miR-196b, hsa-miR-196b*, hsa-miR-19a*, hsa-miR-19b-1*, hsa-miR-19b-2*, hsa-miR-20a*, hsa-miR-210, hsa-miR-22, hsa-miR-301b, hsa-miR-3136-5p, hsa-miR-34a*, hsa-miR-3960, hsa-miR-4693-5p, hsa-miR-4753-3p, hsa-miR-488, hsa-miR-503, hsa-miR-548, hsa-miR-658, hsa-miR-7, hsa-miR-759, hsa-miR-9
Ago3	94	82	hsa-miR-10a*, hsa-miR-1246, hsa-miR-137, hsa-miR-196a, hsa-miR-196b, hsa-miR-196b*, hsa-miR-19a*, hsa-miR-19b-1*, hsa-miR-19b-2*, hsa-miR-20a*, hsa-miR-301b, hsa-miR-3064-5p, hsa-miR-3136-5p, hsa-miR-34a*, hsa-miR-3960, hsa-miR-4693-5p, hsa-miR-4753-3p, hsa-miR-4793-3p, hsa-miR-488, hsa-miR-503, hsa-miR-548, hsa-miR-658, hsa-miR-7, hsa-miR-9
Ago4	90	78	has-miR-4511, hsa-miR-10a*, hsa-miR-1246, hsa-miR-1272, hsa-miR-137, hsa-miR-196a, hsa-miR-196b, hsa-miR-196b*, hsa-miR-19a*, hsa-miR-19b-1*, hsa-miR-19b-2*, hsa-miR-20a*, hsa-miR-210, hsa-miR-22, hsa-miR-301b, hsa-miR-3064-5p, hsa-miR-3136-5p, hsa-miR-34a*, hsa-miR-3960, hsa-miR-4693-5p, hsa-miR-4753-3p, hsa-miR-4793-3p, hsa-miR-488, hsa-miR-503, hsa-miR-548, hsa-miR-658, hsa-miR-7, hsa-miR-759, hsa-miR-9

doi:10.1371/journal.pone.0053823.t002

```

7sl lncRNA_Dre  CCGGGCGCAGTGGCGTGGCGCTGTAATCCAAGCTACT-GGGAGGCTGAGGCTGGGGGACC 59
7sl lncRNA_Dre  GCTTGAGCTCAGGGGTTCTGGGCTGCAGTGGACTATGCCGATCGGGTGTCCGCACTAAGT 119
7sl lncRNA_Dre  TCGGTATCGATATGGTG-C-TCCTGGGGGAGCCITGGACTACCAGGTCGTCTAAGGAGGG 177
7sl lncRNA_Dre  GTGAACGGGCCAGGTCGGAGACGGAGCAGGTCAAAGCCOCCGTGCCGATCAGTAGTGGG 237
7sl lncRNA_Dre  ATCGCGCTGTGAATAGACACTGCAGTGCAGCCTG 272

```

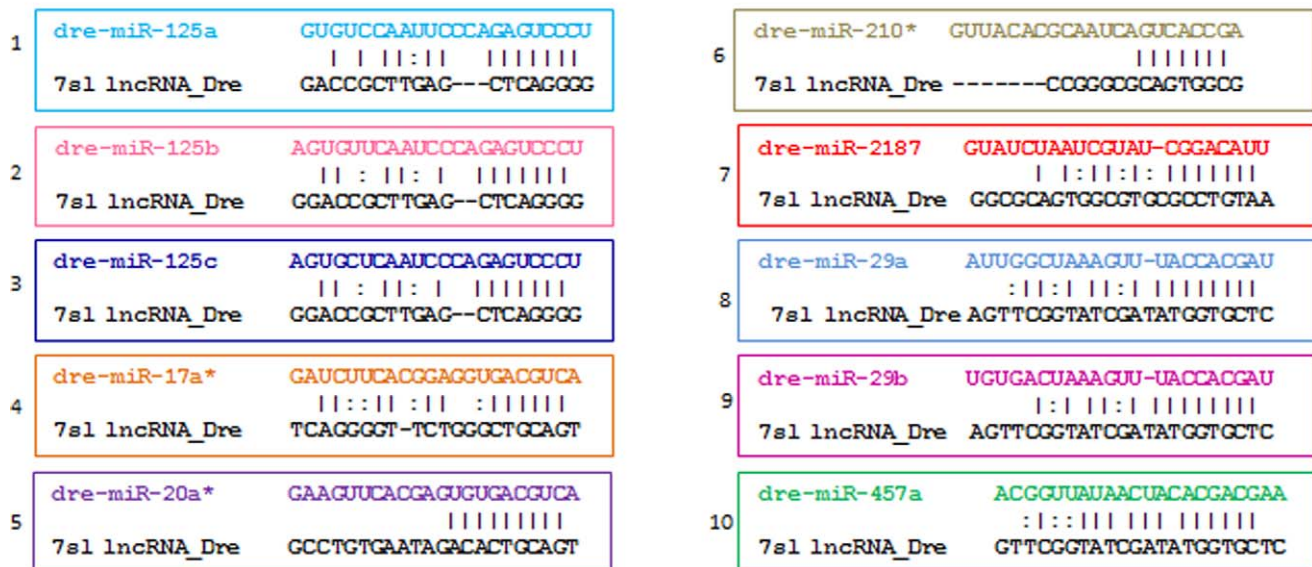


Figure 2. Binding alignments of microRNAs targeting predicted zebrafish lncRNA. The figure shows predicted binding alignment of miR-125a, miR-125b, miR-125c, miR-17a*, miR-20a*, miR-210*, miR-2187, miR-29a, miR-29b and miR-457a with predicted zebrafish lncRNA. doi:10.1371/journal.pone.0053823.g002

light into the biological relevance of the remaining regulatory motifs.

7sl RNA lncRNA is Down Regulated by mir-125b in Zebrafish

We have attempted to model one example of miRNA-lncRNA interaction in a zebrafish model. Initially, we performed BLAT analysis of human lncRNA against zebrafish genome for identification of putative lncRNA in zebrafish as described in method section. The BLAT analysis generated a single hit on zebrafish chromosome no. 3 (chr3:30399858-30400130), which had 91.64% similarity with human 7sl lncRNA. The miRanda software identified potential binding sites of miR-125a, miR-125b, miR-125c, miR-17a*, miR-20a*, miR-210*, miR-2187, miR-29a, miR-29b and miR-457a in the predicted zebrafish 7sl lncRNA (Figure 2). MiR-125b is known to express ubiquitously during zebrafish embryogenesis [46]. In addition, miR-125b is known to targets *p53* mRNA to regulate embryonic development and stress response in zebrafish [46]. Therefore, we were intrigued to study the interaction of miR-125b with the predicted 7sl lncRNA in developing zebrafish embryos.

Prior to examining the interaction of miR-125b with the predicted 7sl lncRNA in developing zebrafish embryos, we

investigated the expression of the predicted zebrafish 7sl lncRNA during zebrafish embryonic development using reverse transcriptase PCR technique. The predicted zebrafish 7sl lncRNA was expressed in 24 and 48 hpf (hour post fertilization) zebrafish embryos (data not shown). The interaction of miR-125b and 7sl lncRNA was investigated by ectopic overexpression of miR-125b mimic in zebrafish embryos [47]. Ectopic overexpression of miR-125b (10 μ M) resulted in anterior-posterior axis curvature defects in approximately 80% of the injected zebrafish embryos at 2 dpf ($n = 225/283$) (Figure 3C and 3D). A sequence unrelated miRNA (miR-144) was used as control. Ectopic overexpression of control miR-144 (10 μ M) did not result in any observable phenotype in zebrafish embryos at 2 dpf. Further expression of the predicted 7sl lncRNA was assayed in miR-125b injected embryos using quantitative real-time PCR. MiR-125b injected embryos displayed reduction in expression of the predicted 7sl lncRNA by approximately 1.72 fold (0.42 ± 0.26 SD) compared to non-injected embryos (NIC) (Figure 3E). However, the expression levels of the predicted 7sl lncRNA were not significantly affected in case of control injected miRNA (Figure 3E). In summary, we show that over expression of miR-125b resulted a corresponding decrease in the expression levels of the predicted 7sl lncRNA in zebrafish embryos.

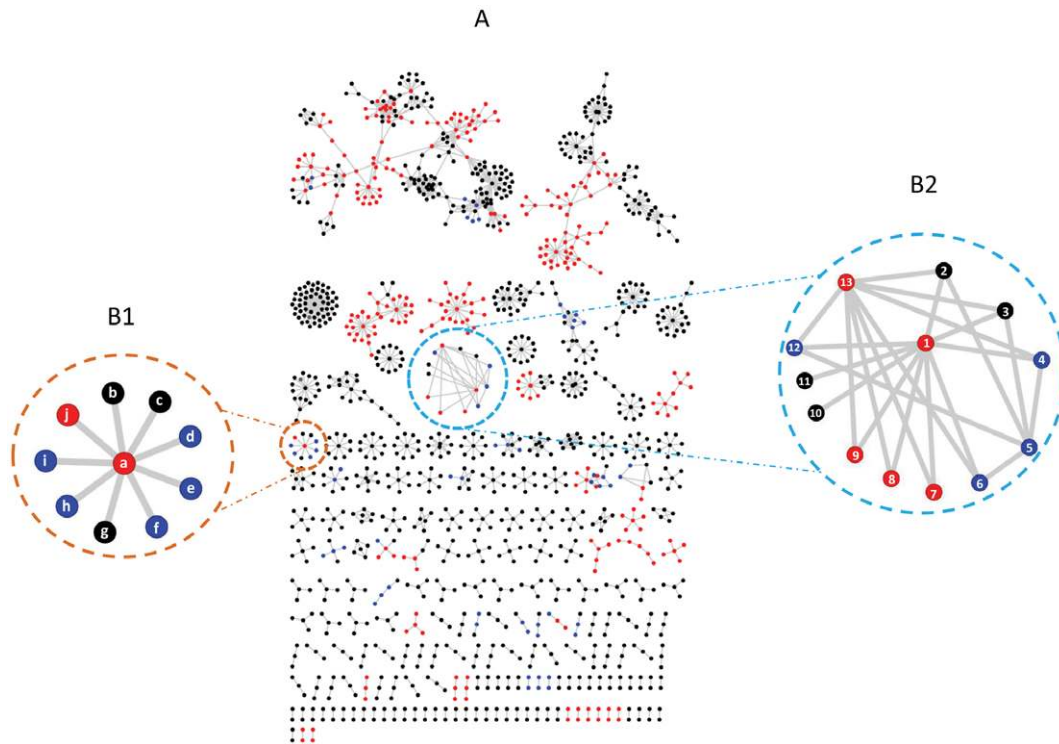


Figure 4. A: Complete interaction network of lncRNA, miRNA and mRNA. The interaction network wherein the experimental miRNA-mRNA interactions are represented in red nodes, the predicted miRNA-lncRNA interactions are represented in blue nodes and total miRNA-mRNA interactions represented as black nodes. **B1:** An interesting example from the network highlighted in orange showing interactions between of network highlighting **miRNA:** (a) hsa-mir-9; **lncRNA:** (d) ENST00000500197.2, (e) ENST00000509783.1, (f) ENST00000511014.1, (h) ENST00000505030.1, (i) ENST00000504246.1; **mRNA:** (b) ENST00000384838.1, (c) ENST00000262095.2, (g) ENST00000491143.1, (j) ENST00000226574. **B2:** Another interesting example from the network highlighted in blue showing interactions between **miRNA:** (1) hsa-miR-196a, (5) hsa-miR-196b*, (13) hsa-miR-196b; **lncRNA:** (4) ENST00000523790.1; (6) ENST00000489695.1, (12) ENST00000519935.1; **mRNA:** (2) ENST00000354032.4, (3) ENST00000384852.1, (7) ENST00000313173, (8) ENST0000024215, (9) ENST0000040584, (10) ENST00000304786.7, (11) ENST00000366839.4. doi:10.1371/journal.pone.0053823.g004

Materials and Methods

Ethics Statement

Zebrafish experiments were performed in strict accordance with the recommendations and guidelines laid down by the CSIR-Institute of Genomics and Integrative Biology, India. The protocol was approved by the Institutional Animal Ethics Committee (IAEC) of the CSIR Institute of Genomics and Integrative Biology, India (Proposal No 45a). All necessary efforts were made to minimize animal suffering.

Zebrafish Husbandry

Zebrafish used in this study were housed at the CSIR-Institute of Genomics and Integrative Biology following standard husbandry practices [52]. Zebrafish embryos were obtained by pair-wise mating of adult fish.

PAR-CLIP Datasets

We used transcriptome wide interaction data for miRNA-Ago available from PAR-CLIP studies [53]. The dataset provides the RNA-binding proteins (miRNA-Ago) target sites identified from the human embryonic kidney (HEK) 293 cell lines stable expressing Ago1-4. The read sequences (SRA IDs: SRX020783, SRX020784, SRX020785, SRX020786) of this study were fetched from Sequence Read Archive (SRA) from NCBI and were assembled using MAQ [42] reference assembly software to the hg19 version of the human genome from UCSC Genome Browser

[54]. To identify only high-confidence miRNA recognition elements, we used cutoff of $5\times$ and all loci with equal to or more than $5X$ coverage were filtered for the further analysis.

Long Noncoding RNA Datasets

The manually annotated lncRNAs from GENCODE version 9 (May 2011 freeze, GRCh37) (<http://www.encodegenes.org/>) [43] were used for the analysis. The dataset contains both Ensembl and Havanna annotations and the transcripts fall into six biotypes, viz, processed transcripts, antisense, lincRNA, ncRNA host, noncoding and retained intron. The dataset had a total of 18878 long noncoding transcripts encompassing by 11004 genes. The former dataset of high-confidence miRNA recognition elements were mapped on the lncRNA exon positions using custom scripts.

MicroRNA Sequence Datasets

The dataset of mature miRNA sequences was derived from miRBase17 [55,56] for all annotated Human miRNAs. This dataset comprised of 1731 sequences. The dataset of validated miRNA targets were derived from miRecords [57].

MicroRNA Target Predictions

To identify the subset of miRNAs for the corresponding high-confidence Ago sites, we used a computational approach using a miRNA target prediction algorithm miRanda [44]. MiRanda uses dynamic programming and incorporates the free energy of the RNA duplex and miRNA binding rules to predict potential

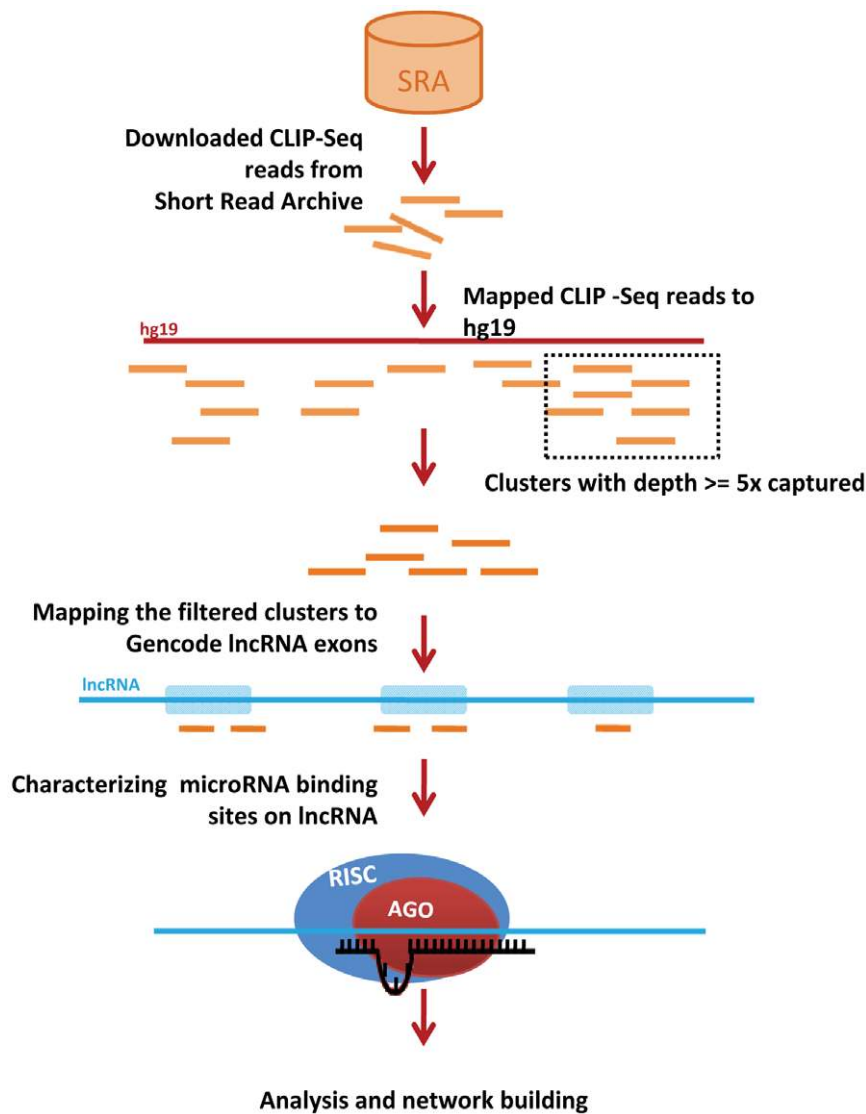


Figure 5. Workflow for the analysis of potential microRNA binding sites on lncRNAs and re-construction of the regulatory interactome.

doi:10.1371/journal.pone.0053823.g005

miRNA target sites. We used an empirical alignment score of 160 and minimum free energy of -20 kcal/mol. Similarly miRNA targets were predicted for all those messenger RNAs which harbored high confidence miRNA recognition elements.

Positional Preference of microRNA Regulatory Elements in Long Noncoding RNAs

The positional preference of miRNAs across the lncRNA length was computed using in-house scripts. We divided the entire length of the lncRNA into non-overlapping bins of 10 percent each. This was done to accommodate the diversity in lengths of the lncRNA. The frequencies of mapping miRNA regulatory elements were plotted for each of the bins.

MicroRNA-interaction Network

Both the interactions, viz, miRNA-lncRNA and miRNA-mRNA were compared and an unweight network displaying miRNA-lncRNA and miRNA-mRNA interactions were created using Cytoscape 2.8.2 [58]. The targets were then searched on

miRecords to check whether any of the high-confidence miRNA recognition elements overlap with validated target sites. The miRNA-mRNA interactions were also taken from literature. The entire workflow of the computational analysis is summarized in Figure 5.

Prediction of lncRNA in Zebrafish

Human long noncoding RNA (lncRNAs) sequences were downloaded from lncRNAdb database [59]. These sequences were further aligned to zebrafish genome (Zv9) using UCSC BLAT tool to retrieve sequences with $\geq 90\%$ identity score. The alignment percentage was calculated by dividing the aligned length by query (lncRNA) length. The zebrafish genomic regions with identical human lncRNA sequences were further analyzed with ensembl genome browser for expressed sequence tag (EST). The resultant sequences were considered as putative lncRNA in zebrafish. These sequences were further analyzed for miRNA binding site as mentioned in previous section (microRNA target predictions).

Microinjections into Zebrafish Embryos

MiRNA injections were performed following a published protocols [47,60]. Glass capillary (World Precision) micropipettes were pulled using Sutter Instrument (USA) and clipped appropriately for microinjection to deliver 2–3 nl solution into one cell stage zebrafish embryos. For overexpression of miR-125b mimic (Dharmacon), two to three nanoliter (nl) containing 10 μ M miRNA mimic was injected at one cell stage zebrafish embryos. The concentration and dose of miRNA that was microinjected in zebrafish embryos were based on the previous published studies [47,61,62].

Imaging

The injected and non-injected control embryos were observed and imaged in white light with an upright Zeiss AxioScope 40 microscope (Carl Zeiss, Germany) using 2.5X magnification and 0.075 numerical aperture. Images were processed with Zeiss AxioVision 4.6 and Adobe Photoshop CS software. Identical modifications and adjustment were applied to all the images in the same experiment.

Quantitative Real-Time PCR (QRT-PCR)

RNA derived from miRNA injected and control embryos were obtained using Qiagen miRNeasy Mini Kit (cat no. 217004) according to manufacturer described protocol. One microgram of RNA was reverse transcribed into cDNA using Superscript II (Invitrogen, USA), and diluted by 1:5 for RT-PCR assay. The assay was performed using Sybr Green mix (Roche, Germany). The oligonucleotide primers used for RT-PCR assay of predicted zebrafish *7sl* lncRNA was p1942F (5'-GTAATCCAAGC-TACTGGGAGGC-3') and p1943R (5'-AGGCGCGATCC-CACTACTGATC-3'). The fold change were calculated by $\Delta\Delta$ CT method [63]. U6 used as an internal control using primers U6F (5'-ACTAAAATTGGAACGATACAGAGA-3') and U6R (5'-AAAGATGGAACGCTTCACG-3').

Conclusions

The present analysis illustrates the potential integration of genome-wide experimental datasets to provide insights into novel biological regulatory interactions in the cell. Interactions for miRNAs with lncRNAs have been suggested to play an important role in modulating regulation mediated through miRNAs. This analysis also provides a genome-scale report and reconstruction of miRNA interaction with mRNAs as well as lncRNAs. We also show how such a network approach could potentially suggest the lncRNAs for functional validation. In addition we have also experimentally validated a miRNA-lncRNA interaction using a zebrafish model, which suggest that miRNA-lncRNA interaction networks are biologically active (Figure 3). We have earlier suggested distinct patterns of epigenetic marks across lncRNAs with potential significance in the transcriptional regulation of

References

- Li Y, Zheng H, Luo R, Wu H, Zhu H, et al. (2011) Structural variation in two human genomes mapped at single-nucleotide resolution by whole genome de novo assembly. *Nat Biotechnol* 29: 723–730.
- Hollox EJ (2012) The challenges of studying complex and dynamic regions of the human genome. *Methods Mol Biol* 838: 187–207.
- Dinger ME, Amaral PP, Mercer TR, Mattick JS (2009) Pervasive transcription of the eukaryotic genome: functional indices and conceptual implications. *Brief Funct Genomic Proteomic* 8: 407–423.
- Lander ES (2011) Initial impact of the sequencing of the human genome. *Nature* 470: 187–197.
- Roberts A, Pimentel H, Trapnell C, Pachter L (2011) Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics* 27: 2325–2329.
- Severin AJ, Peiffer GA, Xu WW, Hyten DL, Bucciarelli B, et al. (2010) An integrative approach to genomic introgression mapping. *Plant Physiol* 154: 3–12.
- 1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. *Nature* 467: 1061–1073.
- ENCODE Project Consortium (2004) The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* 306: 636–640.
- Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10: 57–63.
- Janga SC, Vallabhaneni S (2011) MicroRNAs as post-transcriptional machines and their interplay with cellular networks. *Adv Exp Med Biol* 722: 59–74.

lncRNAs [64]. This present analysis adds to the understanding of yet another regulatory layer at the post transcriptional level modulating lncRNA expression and thereby the function.

The study is not without caveats. The first being limitation in the diversity of CLIP-seq datasets available in public domain. The datasets used in the analysis are limited to HEK293 cells. The second major caveat is that majority of miRNA recognition elements could not be tracked back to their cognate miRNAs. With the advent of faster and cheaper sequencing technologies [65,66] we hope newer genome-wide CLIP-seq datasets would emerge. This would potentially provide a better picture of the depth of interactions and a perspective on the spatio-temporal organization of these networks. We hope the advancements in technology together with further insights into the biological function of small RNA classes including discovery of other members of this class would provide a better insight into the regulatory ramifications. This could also potentially pave the way to identify cognate RNAs for the regulatory elements discovered through genome-wide screens. We also hope the availability of technology to identify targets of RNA binding proteins would open up newer avenues for understanding the functions and targets of many of the RNA binding proteins. This would also provide for the much needed base to understand biological processes and regulatory interactions of ncRNAs with coding transcripts as well as between them in addition to deciphering how these proteins modulate these through binding to their respective partners.

Supporting Information

Figure S1 Schematic of the proposed hypothesis.

(TIF)

Figure S2 List of all transcripts with the Ago binding sites.

(DOC)

Figure S3 Distribution of Ago binding sites across the length of lncRNAs divided in bins of 10 percent. A) The panel shows the number of Ago sites in each bin. B) The panel shows the individual Ago binding sites in lncRNAs transcripts.

(TIF)

Acknowledgments

Authors thank Dr. Chetana Sachidanandan for scientific inputs, which helped in formulating the study and preparation of the manuscript. DB and MKL acknowledge Senior Research Fellowship from CSIR, India.

Author Contributions

Conceived and designed the experiments: VS SS SJ DB. Performed the experiments: MKL. Analyzed the data: SJ DB. Wrote the paper: VS SJ DB.

11. Zhao Y, He S, Liu C, Ru S, Zhao H, et al. (2008) MicroRNA regulation of messenger-like noncoding RNAs: a network of mutual microRNA control. *Trends Genet* 24: 323–327.
12. Zhang R, Su B (2009) Small but influential: the role of microRNAs on gene regulatory network and 3'UTR evolution. *J Genet Genomics* 36: 1–6.
13. de Planell-Saguer M, Rodicio MC (2011) Analytical aspects of microRNA in diagnostics: a review. *Anal Chim Acta* 699: 134–152.
14. Esteller M (2011) Non-coding RNAs in human disease. *Nat Rev Genet* 12: 861–874.
15. Gregory RI, Chendrimada TP, Cooch N, Shiekhattar R (2005) Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* 123: 631–640.
16. Iwasaki S, Tomari Y (2009) Argonaute-mediated translational repression (and activation). *Fly (Austin)* 3: 204–206.
17. Sayed D, Abdellatif M (2011) MicroRNAs in development and disease. *Physiol Rev* 91: 827–887.
18. Wienholds E, Plasterk RH (2005) MicroRNA function in animal development. *FEBS Lett* 579: 5911–5922.
19. Fiore R, Khudayberdiev S, Saba R, Schrott G (2011) MicroRNA function in the nervous system. *Prog Mol Biol Transl Sci* 102: 47–100.
20. Fichtlscherer S, Zeiher AM, Dimmeler S (2011) Circulating microRNAs: biomarkers or mediators of cardiovascular diseases? *Arterioscler Thromb Vasc Biol* 31: 2383–2390.
21. Rederstorff M, Huttenhofer A (2010) Small non-coding RNAs in disease development and host-pathogen interactions. *Curr Opin Mol Ther* 12: 684–694.
22. Scaria V, Hariharan M, Maiti S, Pillai B, Brahmachari SK (2006) Host-virus interaction: a new role for microRNAs. *Retrovirology* 3: 68.
23. Scaria V, Jadhav V (2007) microRNAs in viral oncogenesis. *Retrovirology* 4: 82.
24. Hung T, Chang HY (2010) Long noncoding RNA in genome regulation: prospects and mechanisms. *RNA Biol* 7: 582–585.
25. Mercer TR, Dinger ME, Mattick JS (2009) Long non-coding RNAs: insights into functions. *Nat Rev Genet* 10: 155–159.
26. Liao Q, Liu C, Yuan X, Kang S, Miao R, et al. (2011) Large-scale prediction of long non-coding RNA functions in a coding-non-coding gene co-expression network. *Nucleic Acids Res* 39: 3864–3878.
27. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, et al. (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464: 1071–1076.
28. Wan Y, Chang HY (2010) HOTAIR: Flight of noncoding RNAs in cancer metastasis. *Cell Cycle* 9: 3391–3392.
29. Pontier DB, Gribnau J (2011) Xist regulation and function explored. *Hum Genet* 130: 223–236.
30. Xu C, Yang M, Tian J, Wang X, Li Z (2011) MALAT-1: a long non-coding RNA and its important 3' end functional motif in colorectal cancer metastasis. *Int J Oncol* 39: 169–175.
31. Bhartiya D, Kapoor S, Jalali S, Sati S, Kaushik K, et al. (2012) Conceptual approaches for lncRNA drug discovery and future strategies. *Expert Opin Drug Discov* 7: 503–513.
32. Jalali S, Jayaraj GG, Scaria V (2012) Integrative transcriptome analysis suggest processing of a subset of long non-coding RNAs to small RNAs. *Biol Direct* 7: 25.
33. Collins IJ (2011) The RNA infrastructure: an introduction to ncRNA networks. *Adv Exp Med Biol* 722: 1–19.
34. Xu J, Li CX, Li YS, Lv JY, Ma Y, et al. (2011) MiRNA-miRNA synergistic network: construction via co-regulating functional modules and disease miRNA topological features. *Nucleic Acids Res* 39: 825–836.
35. 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.
36. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP (2011) A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 146: 353–358.
37. 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.
38. Cesana M, Cacchiarelli D, Legnini I, Santini T, Shandier O, et al. (2011) A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* 147: 358–369.
39. Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, et al. (2007) Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat Genet* 39: 1033–1037.
40. Todesco M, Rubio-Somoza I, Paz-Ares J, Weigel D (2010) A collection of target mimics for comprehensive analysis of microRNA function in *Arabidopsis thaliana*. *PLoS Genet* 6: e1001031.
41. Jungkamp AC, Stoeckius M, Mecnas D, Grun D, Mastrobuoni G, et al. (2011) In vivo and transcriptome-wide identification of RNA binding protein target sites. *Mol Cell* 44: 828–840.
42. Li H, Ruan J, Durbin R (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* 18: 1851–1858.
43. Harrow J, Denoeud F, Frankish A, Reymond A, Chen CK, et al. (2006) GENCODE: producing a reference annotation for ENCODE. *Genome Biol* 7 Suppl 1: S4.1–9. Epub; 2006 Aug 7: S4–S9.
44. John B, Enright AJ, Aravin A, Tuschl T, Sander C, et al. (2004) Human MicroRNA targets. *PLoS Biol* 2: e363.
45. Czech B, Hannon GJ (2011) Small RNA sorting: matchmaking for Argonautes. *Nat Rev Genet* 12: 19–31.
46. Le MT, Teh C, Shyh-Chang N, Xie H, Zhou B, et al. (2009) MicroRNA-125b is a novel negative regulator of p53. *Genes Dev* 23: 862–876.
47. Lalwani MK, Sharma M, Singh AR, Chauhan RK, Patowary A, et al. (2012) Reverse genetics screen in zebrafish identifies a role of miR-142a-3p in vascular development and integrity. *PLoS one*. DOI 10.1371/journal.pone.0052588.
48. Majoros WH, Ohler U (2007) Spatial preferences of microRNA targets in 3' untranslated regions. *BMC Genomics* 8: 152.
49. Yekta S, Shih IH, Bartel DP (2004) MicroRNA-directed cleavage of HOXB8 mRNA. *Science* 304: 594–596.
50. He X, Yan YL, Eberhart JK, Herpin A, Wagner TU, et al. (2011) miR-196 regulates axial patterning and pectoral appendage initiation. *Dev Biol* 357: 463–477.
51. Schimanski CC, Frerichs K, Rahman F, Berger M, Lang H, et al. (2009) High miR-196a levels promote the oncogenic phenotype of colorectal cancer cells. *World J Gastroenterol* 15: 2089–2096.
52. Westerfield M (2000) The zebrafish book. A guide for the laboratory use of zebrafish (*Danio rerio*). Eugene: Univ. of Oregon Press.
53. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, et al. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141: 129–141.
54. Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, et al. (2011) The UCSC Genome Browser database: update 2011. *Nucleic Acids Res* 39: D876–D882.
55. Griffiths-Jones S, Saini HK, van DS, Enright AJ (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Res* 36: D154–D158.
56. Kozomara A, Griffiths-Jones S (2011) miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* 39: D152–D157.
57. Xiao F, Zuo Z, Cai G, Kang S, Gao X, et al. (2009) miRecords: an integrated resource for microRNA-target interactions. *Nucleic Acids Res* 37: D105–D110.
58. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27: 431–432.
59. Amaral PP, Clark MB, Gascoigne DK, Dinger ME, Mattick JS (2011) lncRNAdb: a reference database for long noncoding RNAs. *Nucleic Acids Res* 39: D146–D151.
60. Hyatt TM, Ekker SC (1999) Vectors and techniques for ectopic gene expression in zebrafish. *Methods Cell Biol* 59: 117–26.
61. Giraldez AJ, Cinalli RM, Glasner ME, Enright AJ, Thomson JM, et al. (2005) MicroRNAs regulate brain morphogenesis in zebrafish. *Science* 308: 833–838.
62. Pase L, Lieschke GJ (2009) Validating microRNA Target Transcripts Using Zebrafish Assays. *Methods Mol Biol* 546: 227–40.
63. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29: e45.
64. Sati S, Ghosh S, Jain V, Scaria V, Sengupta S (2012) Genome-wide analysis reveals distinct patterns of epigenetic features in long non-coding RNA loci. *Nucleic Acids Res*.
65. Zhou X, Ren L, Meng Q, Li Y, Yu Y, et al. (2010) The next-generation sequencing technology and application. *Protein Cell* 1: 520–536.
66. Nagele P (2011) Perioperative genomics. *Best Pract Res Clin Anaesthesiol* 25: 549–555.