IncRNASNP: a database of SNPs in IncRNAs and their potential functions in human and mouse

Jing Gong1,2,†, Wei Liu2,†, Jiayou Zhang3, Xiaoping Miao1,† and An-Yuan Guo2,*

1Department of Epidemiology and Biostatistics, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei 430030, PR China, 2Department of Biomedical Engineering, Key Laboratory of Molecular Biophysics of the Ministry of Education, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, Hubei 430074, PR China and 3Department of influenza vaccine research, Wuhan Institute of Biological Products, Wuhan, Hubei 430207, PR China

Received August 08, 2014; Revised September 29, 2014; Accepted October 07, 2014

ABSTRACT

Long non-coding RNAs (IncRNAs) play key roles in various cellular contexts and diseases by diverse mechanisms. With the rapid growth of identified IncRNAs and disease-associated single nucleotide polymorphisms (SNPs), there is a great demand to study SNPs in IncRNAs. Aiming to provide a useful resource about IncRNA SNPs, we systematically identified SNPs in IncRNAs and analyzed their potential impacts on IncRNA structure and function. In total, we identified 495 729 and 777 095 SNPs in more than 30 000 IncRNA transcripts in human and mouse, respectively. A large number of SNPs were predicted with the potential to impact on the miRNA–IncRNA interaction. The experimental evidence and conservation of miRNA–IncRNA interaction, as well as miRNA expression from TCGA were also integrated to prioritize the miRNA–IncRNA interactions and SNPs on the binding sites. Furthermore, by mapping SNPs to GWAS results, we found that 142 human IncRNA SNPs are GWAS tagSNPs and 197 827 lncRNA SNPs are in the GWAS linkage disequilibrium regions. All these data for human and mouse IncRNAs were imported into IncRNASNP database (http://bioinfo.life.hust.edu.cn/IncRNASNP/), which includes two sub-databases IncRNASNP-human and IncRNASNP-mouse. The IncRNASNP database has a user-friendly interface for searching and browsing through the SNP, IncRNA and miRNA sections.

INTRODUCTION

The ENCODE project shows that most of the human genome nucleotides can be transcribed into primary transcripts in different cells (1), producing a range of non-coding RNAs (ncRNAs). Long non-coding RNAs (IncRNAs) are ncRNAs defined as transcripts >200 nt in length, which are typically transcribed by RNA polymerase II and are often multiexonic and polyadenylated (2). IncRNAs play critical roles in various biological processes and cellular contexts by diverse mechanisms to participate in the regulation of transcription/post-transcription/post-translation, organization of protein complexes, cell–cell signaling as well as recombination (3,4). The functions of IncRNAs are described in terms of their expressions, biological processes, diseases mechanisms, as well as their polymorphisms (5). By dysregulating target gene expression, IncRNAs involve in a number of complex human diseases, including coronary artery diseases, autoimmune diseases, neurological disorders and various cancers (6).

Single nucleotide polymorphisms (SNPs) represent the most frequent genetic variants among individuals and link to gene expression, function, phenotypes and diseases (7). Recently, SNPs in IncRNAs were found to be linked to their abnormal expressions and dysregulations, thus play important roles in disease association (5,8). For example, the rs2839698 TC genotype of IncRNA H19 was associated with a low risk of developing non-muscle-invasive bladder cancer (9). SNP rs6983267 affects the expression of lncRNA CCAT2 thus impacts tumor growth and metastasis in colorectal cancer (10). In addition, many genome-wide association studies (GWASs) have identified several trait-associated SNPs in or nearby IncRNAs, such as antisense noncoding RNA in the INK4 locus (ANRIL) (11). Therefore, it is important to identify SNPs in IncRNAs, especially link SNPs in IncRNAs with GWAS results for studying complex traits and diseases.

MicroRNAs (miRNAs) are another class of small ncRNAs. IncRNA can interact with miRNA by complementary sequence to act as a miRNA decoy or sponge, which are widely confirmed by experiments (12–15). For example,
lncRNA PTENP1 function as a decoy for PTEN-targeting miRNAs in tumor suppression (14). Currently, starBase has collected more than 10 000 miRNA–lncRNA interaction pairs from high-throughput studies (16). SNPs in miRNA target sites on lncRNAs may influence (destabilize or create) the miRNA–lncRNA interactions, thereby alter their functions. For protein-coding genes, SNP impacts on their mRNA target sites have been validated in many studies (17,18) and a number of databases for predicting functional SNPs in mRNA target sites were constructed, including MicroSNiPer, PolymiRTS, miRdSNP, MirSNP and our previous work miRNASNP (19–23). Thus, a database for SNPs in lncRNAs and their potential functions on lncRNA–miRNA binding will be very useful to accelerate the researches.

Currently, there are many databases about lncRNA sequence, expression and function, including NONCODE, LNCipedia, lncRNAdb, lncRNAtor, lncRNAMap, starBase, lncRNome, Linc2GO and LncRNADisease (16,24–31). There are very few studies and databases about the SNPs in lncRNAs, we identified 495729 SNPs and 777095 SNPs by comparing the genomic coordinates of SNP with those of lncRNAs, respectively. We refer to these SNPs as lncRNA SNPs in human and mouse.

DATA COLLECTION AND DATABASE CONTENT

SNPs in lncRNA transcripts

The human and mouse SNP data were downloaded from dbSNP of NCBI (human GRCh37 and mouse mm10). We obtained the human lncRNA data from LNCipedia database, which contains 32 108 human lncRNA transcripts of 17 436 lncRNA genes (GRCh37) (25). We extracted 36 471 confident mouse lncRNA transcripts from NONCODEv4 by excluding the unpublished and unconfident data. The genomic coordinates of mouse lncRNAs were converted from genome assembly mm9 to mm10 using the LiftOver utility in UCSC (http://genome.ucsc.edu). By comparing the genomic coordinates of SNPs with those of lncRNAs, we identified 495 729 SNPs and 777 095 SNPs in human and mouse lncRNA transcripts (Table 1), respectively. We refer to these SNPs as lncRNA SNPs in human and mouse.

Impacts of SNPs on lncRNA structures

A part of lncRNAs can exert their functions by scaffolding protein complexes (34), suggesting the structural features of lncRNAs play important roles in their functions. SNPs in lncRNA transcripts may influence the lncRNA secondary structure to impact its stability, expression and functions (35). To assess the impacts of SNPs on lncRNA secondary structures, we first extracted lncRNA transcript sequences from human/mouse reference genome according to the lncRNA transcript BED file as Ref-transcripts. For each SNP, we changed the corresponding allele in the given Ref-transcript to alternative allele as Alt-transcript. Then, RNAfold program (36) was used to illustrate the secondary structure and calculate the minimal free energy (MFE, ΔG). Energy change of RNA structures (ΔΔG) was calculated by the MFE differences using ΔΔG = |ΔGalt – ΔGref|, where ΔGref and ΔGalt are the MFEs of the Ref-transcript and Alt-transcript, respectively. In human and mouse lncRNAs, the average energy changes by SNPs are (1.30 ± 1.62) kcal/mol and (1.26 ± 1.27) kcal/mol, while the energy changes of the top 10% are 3.10 kcal/mol and 3.00 kcal/mol, respectively.

miRNA–lncRNA interaction

Increasing evidence shows that lncRNAs can be directly regulated by miRNAs (16), thereby indirectly regulate gene expression by competing with shared miRNAs (12). In lncRNASNP database, we not only predicted miRNA target sites on lncRNAs, but also integrated the interaction conservation, miRNA expression and experimental validated miRNA–lncRNA interaction for prioritization.

After downloading mature miRNA sequences from latest version of miRBase (37) (release 20), miRNA target sites on lncRNAs were predicted by combining the results of two popular tools, namely TargetScan (38) and miRanda (39). To further reduce false positives and prioritize the interactions, we also considered the conservations of target sites and the expression of miRNAs. The miRNA binding sites existing in the conserved exons across human, mouse, rat and dog were classified as conserved. The conservation analysis of lncRNA exons were analyzed by UCSC LiftOver tool and set the parameter of ‘minimum ratio of bases that must remap’ as 0.5. By analyzing the miRNA expression profiles of 9566 human small RNA samples across 28 diseases from The Cancer Genome Atlas (TCGA), we found that only 553 miRNAs (21%) have an expression level > 1 RPM (reads per million) and the remaining 79% of miRNAs are almost no expression. In a specific tissue, the predicted bindings of these unexpressed miRNAs may be very week and useless. Thus, we provided the miRNA expression levels to help users to assess the confidence of the miRNA–lncRNA interaction in specific tissues. Furthermore, we also included 8091 experimentally supported human miRNA–lncRNA interactions from starBase (16).

Impacts of SNPs on miRNA–lncRNA interactions

SNPs in the mRNA target sites may destroy or create miRNA binding sites on lncRNAs, which result in loss and/or gain function of miRNA–lncRNA interactions. To
TABLE 1. Data summary in lncRNASNP database

| Data content                                      | Human       | Mouse       |
|--------------------------------------------------|-------------|-------------|
| lncRNA gene/transcript                            | 17436/32108 | 25512/36471 |
| miRNA                                            | 2576        | 1900        |
| SNPs in lncRNA transcripts                       | 495729      | 777095      |
| SNPs affect MLP* (loss/gain)                      | 262154/280012| 366731/357246|
| All predicted MLP                                 | 6413273     | 7448200     |
| Conserved MLPb                                   | 69637       | 13780       |
| Experimental MLPc (all/loss by SNPs)             | 8091/2116   | NA          |
| MLP affected by SNPs (loss/gain)                 | 799694/830829| 898220/864767|
| MLP loss by SNPs with miRNA RPM>1d               | 158105      | NA          |
| MLP gain by SNPs with miRNA RPM>1d               | 165731      | NA          |
| lncRNA SNPs are GWAS tagSNPs                     | 142         | NA          |
| lncRNA SNPs in GWAS LD regions                   | 197827      | NA          |

*MLP means miRNA-lncRNA target pair.
*The miRNA binding sites are existed in all the conserved exons across human, mouse, rat, and dog.
*The miRNA–lncRNA target pairs supported by CLIP experiment results from starBase.
*dThe miRNA RPM is the average value estimated from TCGA samples (9566 samples across 28 cancer types).

assess the impact of each SNP on miRNA–lncRNA interaction, we used TargetScan and miRanda to predict target sites on the SNP around sequences (±25 bp around SNP) of Ref-transcripts and Alt-transcripts. Thus, we obtained four groups of target datasets, which were recorded as RT (targets on Ref-transcripts by TargetScan), RM (targets on Ref-transcripts by miRanda), AT (targets on Alt-transcripts by TargetScan) and AM (targets on Alt-transcripts by miRanda). The miRNA–lncRNA pairs exist in both RT and RM, but in neither AT nor AM, were defined as interaction losses. On the contrary, if a miRNA–lncRNA pair was predicted in both AT and AM, but in neither RT nor RM, we defined the Alt-transcript gained a novel target site. We found a large number of SNPs with the potential to disturb original miRNA target sites or/and created new potential miRNA target sites. The detail results are shown in Table 1.

SNPs in GWAS identified trait associated regions

Currently, GWASs have linked more than 10,000 SNPs to human traits and diseases. However, most of them locate outside of protein-coding regions. The lncRNASNP database attempts to link SNPs in lncRNAs to phenotypes and human diseases. We utilized the GWAS results from the NHGRI GWAS Catalog (40). By intersecting the GWAS identified strongest risk SNPs (record as GWAS tagSNPs in our study) with SNPs in lncRNAs, we found 142 GWAS tagSNPs in human lncRNAs. Aim to better utilize the GWAS data, we also captured the LD blocks of each tagSNP for different populations using Haploview (setting $R^2 \geq 0.8$) (41) by analyzing the genotype information of ±500 kb around the tagSNP and defined these LD blocks as GWAS identified trait associated regions. The used populations include ASW, CHB, CEU, CHD, JPT, TSI, YRI, LWK, MEX, MKK and GIH. Our analysis shows that 197 827 lncRNA SNPs are in the GWAS LD regions of all populations.

DATABASE ORGANIZATION AND WEB INTERFACE

All data were organized into a set of relational MySQL tables. The lncRNAs transcript ID, miRNA ID and SNP ID were used as the primary keys to organize and link all tables. The lncRNASNP website (http://bioinfo.life.hust.edu.cn/lncRNASNP) was built with the Ruby on Rails (RoR) open-source web framework. We developed several user-friendly search boxes to make the data retrieval easily and efficiently. A quick search box was designed at the top-right of each page for searching by lncRNA transcript ID, miRNA ID, SNP ID and even a defined chromosome region (Figure 1A). The advanced search was provided on the home page with multiple ways for searching by different keywords or batch search by a list of IDs.

Since lncRNASNP database contains data from human and mouse, we divided it into two sub-databases, namely lncRNASNP-human and lncRNASNP-mouse. Each of them displays data mainly through three sections, which are SNP section, lncRNA section and miRNA section.

The SNP section stored all the SNPs in lncRNAs as well as their detailed information. Data in this section include (i) basic information of the SNP; (ii) SNP impacts on lncRNA secondary structure; (iii) SNP impacts on miRNA–lncRNA interaction. It also contains information of SNPs in GWAS LD regions and related GWAS information (Figure 1B) in the lncRNASNP-human database. We set four options to optimize the SNP selection because of the huge SNP list. The first is the ‘position’ option allowing users to query SNPs in a user-defined chromosome region. The second is the ‘GMAF’ option helping users to filter results by the global minor allele frequency of a SNP, which is an important consideration of the sample size in association studies. Users can also limit the queried SNPs in GWAS regions or/and impacting miRNA–lncRNA interaction. By combining these options, users can quickly access to the interested polymorphisms.

The lncRNA section provides the information of lncRNAs. By clicking on a specific lncRNA, users can access its basic information, the SNPs on it, miRNAs binding to it and potential miRNA target sites gain/loss by SNPs (Figure 1C and E). In the ‘Basic information’ part, we provide the conservation information of lncRNA gene by integrating the results of Necsulea et al. (42) and lncRNAtor database (27). In the ‘miRNAs binding to lncRNA’ part,
**Figure 1.** Overview of the lncRNASNP database. (A) The navigation bar of the database, including three main sections. (B) The SNP page with GWAS information and SNP effects on miRNA–lncRNA interactions. (C) An example of miRNA target loss by a SNP in lncRNA. (D) The miRNA section to search all targeted lncRNAs of miRNAs in four categories based on their expression. (E) The miRNA binding sites on a lncRNA.
the potential binding miRNAs were shown by small bars on the binding location with different colors representing different target types (grouped by conservation, experimental evidence and miRNA expression level) (Figure 1E).

The miRNA section enables users to select interested miRNAs by expression to obtain their targeting lncRNAs and possible gain/loss target of lncRNAs by SNPs. In this section, the miRNAs have been classified into four categories according to their expressions. They were high expressed miRNA group (RPM ≥ 1000), middle expressed miRNA group (10 ≤ RPM < 1000), low expressed miRNA group (1 ≤ RPM < 10) and unexpressed miRNA group (RPM < 1). For each human miRNA, we also provide the expression levels in 28 cancer types in the ‘basic information of miRNA’ part (Figure 1D).

SUMMARY AND FUTURE DIRECTIONS

With the rapid advances in high-throughput technologies, the lncRNA identification increases greatly in the past few years. It is necessary to analyze the lncRNA related SNPs and construct a database to help elucidating the biological functions of lncRNAs. In this study, we systematically integrated lncRNA, miRNA and SNP information as well as GWAS results and miRNA expression profiles to seek potential functional SNPs. The lncRNASNP database can be used for either small or large-scale SNP selection from lncRNAs and is particularly useful for association and functional studies. In lncRNASNP database, we found 142 of GWAS tagSNPs in lncRNAs and lots of lncRNA SNPs in GWAS LD regions. These results will help to find the real causative SNPs and genes for GWAS studies. Furthermore, we have predicted the miRNA target sites on lncRNAs and integrated experimentally validated miRNA–lncRNA pairs in the database. In combination with other miRNA/gene databases, users can further explore the relationship between miRNA, lncRNA and protein-coding genes.

We can envision that in the next few years the number of lncRNAs will accumulate rapidly not only in human and mouse, but also in other species. In the future, lncRNASNP will update the database frequently and extend to other species. For other species without GWAS data, the eQTL data will be utilized for SNP selection. As the data accumulation of lncRNA expression and miRNA–lncRNA binding, we will update them into the database to reduce false positive predictions. We aim to develop and maintain the lncRNASNP database as a useful resource for the research community.

ACKNOWLEDGEMENTS

The authors thank XueQin Chen, Yin Lin, Qiong Zhang and Shuyun Bo for the debugging and design of the website. We are also grateful to members of the Guo Lab and Miao lab for helpful suggestions.

FUNDING

National Natural Science Foundation of China [31171271, 31270885, 31471247, 81402744, 31201000 and 81222038]; Program for New Century Excellent Talents in University; Ministry of Education of China [to A.Y.G. and X.M.]; General Financial Grant from the China Postdoctoral Science Foundation [2014M552049 and 2012M511605]; National Program for Support of Top-notch Young Professionals [to X.M.]; Funding for open access charge: National Natural Science Foundation of China [NSFC-31270885].

Conflict of interest statement. None declared.

REFERENCES

1. Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E. et al. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature, 447, 799–816.
2. Rinn, J.L. and Chang, H.Y. (2012) Genome regulation by long noncoding RNAs. Annu. Rev. Biochem., 81, 145–166.
3. Geisler, S. and Coller, J. (2013) RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. Nat. Rev. Mol. Cell Biol., 14, 699–712.
4. Hu, W., Alvarez-Dominguez, J.R. and Lodish, H.F. (2012) Regulation of mammalian cell differentiation by long non-coding RNAs. EMBO Rep., 13, 971–983.
5. Tang, J.Y., Lee, J.C., Chang, Y.T., Hou, M.F., Huang, H.W., Liaw, C.C. and Chang, H.W. (2013) Long noncoding RNAs-related diseases, cancers, and drugs. The Scientific World Journal, 2013, 943539.
6. Li, J., Yuan, Z. and Liu, C. (2013) Long non-coding RNAs and complex human diseases. Int. J. Mol. Sci., 14, 18790–18808.
7. Shastry, B.S. (2009) SNPs: impact on gene function and phenotype. Methods Mol. Biol., 578, 3–22.
8. Zhang, X., Zhou, L., Fu, G., Sun, F., Shi, J., Wei, J., Lu, C., Zhou, C., Yuan, Q. and Yang, M. (2014) The identification of an ESCC susceptibility SNP rs920778 that regulates the expression of lncRNA HOTAIR via a novel intronic enhancer. Carcinogenesis, 35, 2062–2067.
9. Verhaegh, G.W., Verkleij, L., Vermeulen, S.H., Heijer, M., Wittjes, J.A. and Kiemeneij, L.A. (2008) Polymorphisms in the H19 gene and the risk of bladder cancer. Eur. Urol., 54, 1118–1126.
10. Ling, H., Spizzo, R., Atlasi, Y., Nicoloso, M., Shimizu, M., Redis, R.S., Nishida, N., Gafa, R., Song, J., Guo, Z. et al. (2013) CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res., 23, 1446–1461.
11. Pasman, E., Sabbagh, A., Vidaud, M. and Bieche, I. (2011) ANRIL, a long, non-coding RNA, is an unexpected major hotspot in GWAS. FASEB J., 25, 444–448.
12. Wang, Y., Xu, Z., Jiang, J., Xie, C., Wang, D., Wang, M., Xu, X., Xu, J., Wu, M., Yang, M., Chen, J. and Zhang, J. (2010) A SNP in a pseudogene mRNA regulates tumour biology. Eur. J. Cancer, 46, 358–369.
13. Bak, R.O. and Mikkelson, J.G. (2014) miRNA sponges: soaking up miRNAs for regulation of gene expression. Wiley Interdiscip. Rev. RNA, 5, 317–333.
14. Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W.J. and Pandolfi, P.P. (2010) A coding-independent function of gene and pseudogene miRNAs regulates tumour biology. Nature, 465, 1033–1038.
15. Cesana, M., Cacchiarelli, D., Regnini, I., Santini, T., Shandier, O., Chiappini, M., Tramontano, A. and Bozzoni, I. (2011) A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell, 147, 358–369.
16. Li, J.H., Liu, S., Zhou, H., Qu, L.H. and Yang, J.H. (2014) starBase v2.0: decoding miRNA-ceRNA, miRNA-mRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res., 42, D92–D97.
17. Chin, L.J., Rainier, E., Leng, S., Zhai, R., Nallur, S., Babar, I., Muller, R.U., Straka, E., Su, L., Burki, E.A. et al. (2008) A SNP in a let-7 microRNA complementary site in the KRAS 3’ untranslated region increases non-small cell lung cancer risk. Cancer Res., 68, 8535–8540.
18. Kapeller, J., Houghton, L.A., Monnikes, H., Walstab, J., Moller, D., Bonisch, H., Burwinkel, B., Autschbach, F., Funke, B., Lasitschka, F.
et al. (2008) First evidence for an association of a functional variant in the microRNA-510 target site of the serotonin receptor-type 3E gene with diarrhea predominant irritable bowel syndrome. *Hum. Mol. Genet.*, 17, 2967–2977.

19. Barenboim, M., Zoltick, B.J., Guo, Y. and Weinberger, D.R. (2010) MicroSNiPer: a web tool for prediction of SNP effects on putative microRNA targets. *Hum. Mutat.*, 31, 1223–1232.

20. Bhattacharya, A., Ziebarth, J.D. and Cui, Y. (2014) PolymiRTS Database 3.0: linking polymorphisms in microRNAs and their target sites with human diseases and biological pathways. *Nucleic Acids Res.*, 42, D86–D91.

21. Bruno, A.E., Li, L., Kalabus, J.L., Pan, Y., Yu, A. and Hu, Z. (2012) miRdSNP: a database of disease-associated SNPs and microRNA target sites on 3′UTRs of human genes. *BMC Genomics*, 13, 44.

22. Liu, C., Zhang, F., Li, T., Wang, L., Yue, W. and Zhang, D. (2012) MirSNP, a database of polymorphisms altering miRNA target sites, identifies miRNA-related SNPs in GWAS SNPs and eQTLs. *BMC Genomics*, 13, 661.

23. Gong, J., Tong, Y., Zhang, H.M., Wang, K., Hu, T., Shan, G., Sun, J. and Guo, A.Y. (2012) Genome-wide identification of SNPs in microRNA genes and the SNP effects on microRNA target binding and biogenesis. *Human mutation*, 33, 254–263.

24. Bu, D., Yu, K., Sun, S., Xie, C., Skogerbo, G., Miao, R., Xia, H., Liao, Q., Luo, H., Zhao, G. et al. (2012) NONCODE v3.0: integrative annotation of long noncoding RNAs. *Nucleic Acids Res.*, 40, D210–D215.

25. Volders, P.J., Helsen, K., Wang, X., Menten, B., Martens, L., Gevaert, K., Vandesompele, J. and Mestdagh, P. (2013) LNCipedia: a database for annotated human lncRNA transcript sequences and structures. *Nucleic Acids Res.*, 41, D246–D251.

26. Amaral, P.P., Clark, M.B., Gascoigne, D.K., Dinger, M.E. and Mattick, J.S. (2011) lncRNAdb: a reference database for long noncoding RNAs. *Nucleic Acids Res.*, 39, D146–D151.

27. Park, C., Yu, N., Choi, I., Kim, W. and Lee, S. (2014) lncRNAtor: a comprehensive resource for functional investigation of long non-coding RNAs. *Bioinformatics*, 30, 2480–2485.

28. Chan, W.L., Huang, H.D. and Chang, J.G. (2014) lncRNAMap: a map of putative regulatory functions in the long non-coding transcriptome. *Comput. Biol. Chem.*, 50, 41–49.

29. Bhartiya, D., Pal, K., Ghosh, S., Kapoor, S., Jalali, S., Panwar, B., Jain, S., Sati, S., Sengupta, S., Sachidanandan, C. et al. (2013) lncRNome: a comprehensive knowledgebase of human long noncoding RNAs. *Database*, 2013, bat034.

30. Liu, K., Yan, Z., Li, Y. and Sun, Z. (2013) Linc2GO: a human LincRNA function annotation resource based on ceRNA hypothesis. *Bioinformatics*, 29, 2221–2222.

31. Chen, G., Wang, Z., Wang, D., Qiu, C., Liu, M., Chen, X., Zhang, Q., Yan, G. and Cui, Q. (2013) LncRNA Disease: a database for long-non-coding RNA-associated diseases. *Nucleic Acids Res.*, 41, D983–D986.

32. Chen, G., Qiu, C., Zhang, Q., Liu, B. and Cui, Q. (2013) Genome-wide analysis of human SNPs at long intergenic noncoding RNAs. *Hum. Mutat.*, 34, 338–344.

33. Ning, S., Zhao, Z., Ye, J., Wang, P., Zhi, H., Li, R., Wang, T. and Li, X. (2014) LincSNP: a database of linking disease-associated SNPs to human large intergenic non-coding RNAs. *BMC Bioinformatics*, 15, 152.

34. Mercer, T.R., Dinger, M.E. and Mattick, J.S. (2009) Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.*, 10, 155–159.

35. Ding, J., Lu, Q., Ouyang, Y., Mao, H., Zhang, P., Yao, J., Xu, C., Li, X., Xiao, J. and Zhang, Q. (2012) A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proc. Natl. Acad. Sci. U.S.A.*, 109, 2654–2659.

36. Hofacker, I.L. (2003) Vienna RNA secondary structure server.

37. Kozomara, A. and Griffiths-Jones, S. (2011) miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.*, 39, D152–D157.

38. Friedlander, R.C., Farh, K.K., Burge, C.B. and Bartel, D.P. (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.*, 19, 92–105.

39. Betel, D., Wilson, M., Gabow, A., Marks, D.S. and Sander, C. (2008) The microRNA.org resource: targets and expression. *Nucleic Acids Res.*, 36, D149–D153.

40. Klemm, A., Flicek, P., Manolio, T. and Hindorff, L. (2011) lncRNAdb: a reference database for long noncoding RNAs. *Nucleic Acids Res.*, 39, D146–D151.

41. Park, C., Yu, N., Choi, I., Kim, W. and Lee, S. (2014) lncRNAtor: a comprehensive resource for functional investigation of long non-coding RNAs. *Bioinformatics*, 30, 2480–2485.

42. Chan, W.L., Huang, H.D. and Chang, J.G. (2014) lncRNAMap: a map of putative regulatory functions in the long non-coding transcriptome. *Comput. Biol. Chem.*, 50, 41–49.

43. Bhartiya, D., Pal, K., Ghosh, S., Kapoor, S., Jalali, S., Panwar, B., Jain, S., Sati, S., Sengupta, S., Sachidanandan, C. et al. (2013) lncRNome: a comprehensive knowledgebase of human long noncoding RNAs. *Database*, 2013, bat034.