D192–D198 *Nucleic Acids Research, 2023, Vol. 51, Database issue* *Published online 9 November 2022*
https://doi.org/10.1093/nar/gkac981

lncRNASNP v3: an updated database for functional variants in long non-coding RNAs

Yanbo Yang[1,†], Dongyang Wang[1,†], Ya-Ru Miao[2], Xiaohong Wu[1], Haohui Luo[1], Wen Cao[1], Wenqian Yang[1], Jianye Yang[1], An-Yuan Guo[2,*] and Jing Gong[1,3,*]

[1]Hubei Key Laboratory of Agricultural Bioinformatics, College of Informatics, Huazhong Agricultural University, Wuhan 430070, China, [2]Hubei Bioinformatics and Molecular Imaging Key Laboratory, Key Laboratory of Molecular Biophysics of the Ministry of Education, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China and [3]College of Biomedicine and Health, Huazhong Agricultural University, Wuhan 430070, China

Received September 14, 2022; Revised October 12, 2022; Editorial Decision October 14, 2022; Accepted November 03, 2022

ABSTRACT

Long non-coding RNAs (lncRNAs) act as versatile regulators of many biological processes and play vital roles in various diseases. lncRNASNP is dedicated to providing a comprehensive repository of single nucleotide polymorphisms (SNPs) and somatic mutations in lncRNAs and their impacts on lncRNA structure and function. Since the last release in 2018, there has been a huge increase in the number of variants and lncRNAs. Thus, we updated the lncRNASNP to version 3 by expanding the species to eight eukaryotic species (human, chimpanzee, pig, mouse, rat, chicken, zebrafish, and fruitfly), updating the data and adding several new features. SNPs in lncRNASNP have increased from 11 181 387 to 67 513 785. The human mutations have increased from 1 174 768 to 2 387 685, including 1 031 639 TCGA mutations and 1 356 046 CosmicNCVs. Compared with the last release, updated and new features in lncRNASNP v3 include (i) SNPs in lncRNAs and their impacts on lncRNAs for eight species, (ii) SNP effects on miRNA–lncRNA interactions for eight species, (iii) lncRNA expression profiles for six species, (iv) disease & GWAS-associated lncRNAs and variants, (v) experimental & predicted lncRNAs and drug target associations and (vi) SNP effects on lncRNA expression (eQTL) across tumor & normal tissues. The lncRNASNP v3 is freely available at http://gong_lab.hzau.edu.cn/lncRNASNP3/.

INTRODUCTION

Long non-coding RNAs (lncRNAs) are defined as transcripts longer than 200 nucleotides (nt) and lack protein-coding capacity (1). In recent decades, a large number of lncRNAs have been identified in animals and plants (2). Accumulating studies have revealed numerous functional lncRNAs, which exert their functions through multiple approaches, including interacting with DNA, RNA and protein (3), thereby regulating diverse cellular functions, such as RNA processing, mRNA stability, translation, and post-translational modifications (4). In addition, many lncRNAs have been reported to be involved in critical biological processes and diseases. For example, several lncRNAs were reported as suppressive or oncogenic factors in different cancers (5,6).

Single nucleotide polymorphisms (SNPs) and somatic mutations in lncRNAs can alter the lncRNA structure and affect lncRNA function and are thus further involved in various biological processes and human diseases (7–11). For example, rs12982687 could affect the binding capacity of lncRNA *UCA1* with miR-873-5p and is involved in smoking-triggered colorectal cancer cell migration and invasion (9). Rs140618127 creates a binding site of miR-539-5p on lncRNA *LOC146880*, which causes the reduction of phosphorylation of *ENO1* and is further linked to non-small cell lung cancer progression (8). However, among millions of SNPs and hundreds of thousands of lncRNAs, the functions of SNPs in lncRNAs remain largely unknown. Thus, we developed lncRNASNP that is dedicated to annotating SNPs in lncRNAs and predicting their effects on lncRNA structures and functions (12,13). Based on the lncRNASNP, several functional SNPs in lncRNAs have been identified and experimentally validated (8,14). However, more challenges than victories still exist for the functional validation of lncRNAs and related variants for

*To whom correspondence should be addressed. Tel: +86 27 8728 5085; Fax: +86 27 8728 5085; Email: gong.jing@mail.hzau.edu.cn
Correspondence may also be addressed to An-Yuan Guo. Tel: +86 27 8779 3177; Fax: +86 27 8779 3177; Email: guoay@hust.edu.cn
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

Table 1. Data summary in lncRNASNP v3

Data content	Version 1.0	Version 2.0	Version 3.0
lncRNA genes/transcripts	42 948/68 579	170 002/258 758	265 602/438 104
All SNPs	1 272 824	11 181 387	67 513 785
lncRNASNP in GWAS[a]	142/197 827	602/2 859 147	14 222/42 830 177
SNP affected MLP[b]	628 885/637 258	7 169 172/5 872 466	19 692 736/18 088 802
All Predicted MLP[b]	13 861 473	16 942 990	45 774 338
TCGA cancer mutations	NA	315 234	1 031 639
TCGA mutations affected MLP[b]	NA	83 633/80 114	340 422/283 635
CosmicNCVs	NA	859 534	1 356 046
CosmicNCVs affected MLP[b]	NA	362 940/350 827	453 202/354 231
lncRNA-associated diseases[c]	NA	697	129 513
ClinVar SNPs	NA	NA	135 937
lncRNA-associated drug & compounds[d]	NA	NA	10 849/4688
lncRNA-associated eQTLs[e]	NA	NA	28 019/615 376

[a]lncRNASNPs are GWAS TagSNPs/lncRNA SNPs in GWAS LD regions.
[b]MLP represents the miRNA−lncRNA target pair, and variants (SNPs, TCGA mutations, CosmicNCVs) in lncRNAs induce potential MLP loss/gain.
[c]The number of experimentally supported lncRNA-associated disease pairs.
[d]The number of lncRNA transcripts associated with drug & compounds (predicted)/the number of experimentally supported lncRNA-associated drug & compound pairs.
[e]The numbers of eQTLs across normal/tumor tissues.

several possible reasons. First, prioritizing functional variants in lncRNAs is more difficult than protein-coding genes because variants in lncRNAs do not directly affect the codon sequence. In addition, previous studies have shown that most lncRNAs are less conservation and lower expressed than protein-coding genes (15), and abundant lncRNAs show tissue specificity in expression (16). Therefore, systematic prediction of functional variants in lncRNAs and collection of comprehensive information, such as lncRNA expression in multiple tissues and variant conservation in multiple species, are essential for the prioritization of functional lncRNAs and related variants.

With the wide application of next-generation sequencing in more species and tissues, the numbers of identified lncRNAs and SNPs have increased rapidly in the past few years. Since the last release of lncRNASNP in 2018 (13), the number of SNPs in human has increased sevenfold in the latest dbSNP (build 155) (17), and the lncRNAs provided by NONCODE (2) have also increased to 1.22-fold. In addition, the progress of high-throughput genomic technologies has also greatly benefitted animal researchers over the past decade. A series of animal SNP resources (14) and lncRNA databases (2) have been developed in recent years. Increasing studies have begun to focus on the impact of SNPs on lncRNAs in animals. For example, an analysis of pigs showed that several SNPs in lncRNA *MEG3* were associated with meat production (18). Another study in chicken also demonstrated that rs1914215137 in the lncRNA *pouBW*1 was associated with chicken growth and carcass (19). However, large-scale genome-wide analyses of animal lncRNA-SNPs and their impacts have rarely been reported.

In recent years, some new databases have been developed to display the sequence, expression, and functions of lncRNAs (2,20–30). There are also databases describing SNPs in lncRNAs, e.g. lincSNP (31) provides disease-associated SNPs on lncRNAs, and LncVar (32) describes the SNPs and structural variants on lncRNAs and assesses the function of these variants. However, few databases predict the effects of SNPs on lncRNA−miRNA binding and focus on lncRNA-related variants in animals. Thus, we updated the lncRNASNP with the latest data and added six commonly studied animals as well as several new features. With the abundant data, the latest release of the lncRNASNP database will be a helpful resource for functional studies of SNPs and lncRNAs. The lncRNASNP v3 is freely available at http://gong_lab.hzau.edu.cn/lncRNASNP3/.

DATA SOURCE AND SUMMARY

In addition to human and mouse, six other species (chimpanzee, pig, mouse, rat, chicken, zebrafish and fruitfly) were included in lncRNASNP v3, bringing the total number of species to eight. The current release of lncRNASNP contains 438 104 lncRNA transcripts of 265 602 lncRNA genes across eight species obtained from NONCODE v6 (2) (Table 1). SNP data for human were obtained from dbSNP (build 155) (17), and SNP data for other species were obtained from the European Variation Archive (EVA) database (33) (Table 1). After intersecting the genomic positions of SNPs and lncRNA transcripts, we identified 67 513 785 SNPs in all lncRNAs (Table 1), hereafter termed lncRNA-SNPs. For human, we also identified 1 356 046 COSMIC mutations (34) and 1 031 639 TCGA cancer mutations (23) in lncRNAs. In addition, we systematically analyzed the impacts of all variants (SNPs and mutations) on lncRNA secondary structure and miRNA−lncRNA interactions. Furthermore, additional resources were integrated, including lncRNA expression profiles for six species (except chimpanzee and pig)(35), experimentally supported disease-related lncRNAs for three species (human, mouse, and rat), ClinVar variants (36) on human lncRNAs, associations between lncRNA expression and drug targets, and lncRNA-eQTLs from ncRNA-eQTL (37) and GTEx (38). Compared with the previous release, lncRNASNP v3 provides more information on SNPs and lncRNAs in more species, and adds several new features (Table 1).

Table 2. Details of SNPs and lncRNAs for multiple species in lncRNASNP v3

Species	lncRNA-SNP	SNP affected lncRNA structure	SNP affected MLP[a]	lncRNA expression	lncRNA diseases
Homo sapiens	62 374 572	8 959 447	18 571 275/17 053 141	102 970	127 528
Pan troglodytes	5579	927	649/630	NA	NA
Mus musculus	4 353 914	677 302	1 069 549/988 382	130 051	1874
Rattus norvegicus	29 624	5 253	4 320/4 101	24 855	111
Sus scrofa	52 898	10 722	5 220/4 556	NA	NA
Gallus gallus	439	68	86/88	11 942	NA
Danio rerio	26 655	4 382	835/753	4 827	NA
Drosophila melanogaster	670 104	124 102	40 802/37 151	42 537	NA

[a]MLP represents the miRNA−lncRNA target pair, and variants (SNPs, TCGA mutations, CosmicNCVs) in lncRNAs induce potential MLP loss/gain.

IMPROVED CONTENT AND NEW FEATURES

Effects of variants on lncRNA secondary structures

Following the strategy of the last release, we used RNAsnp v1.2 (39) to assess variant effects on lncRNA secondary structure for human, mouse, and other species. We chose mode 1 of RNAsnp for lncRNAs <1000 nt and mode 2 for lncRNAs ≥1000 nt, as the software recommended. With an empirical *P*-value <0.2 (40), we obtained 9 782 203 SNPs with effects on lncRNA structure across species. Compared with lncRNASNP2, this number for human was updated from 1 425 449 to 8 959 447, and that of the mouse was updated from 395 443 to 677 302. The number of SNPs affecting lncRNA structure for other animals is 145 454 (Table 2).

Impacts of variants on miRNA–lncRNA interactions

Studies have proven that lncRNAs can interact with miRNAs, and variants in lncRNAs may affect the miRNA−lncRNA interactions (8). Hence, we systematically predicted the potential binding sites of miRNAs on lncRNAs and the impact of SNPs and mutations on miRNA–lncRNA interactions. We first predicted wild-type miRNA binding sites on lncRNAs using the mature miRNA sequences obtained from miRBase (release 22.1) (41). For better reliability, we intersected the prediction results of two popular software, miRmap (42) and TargetScan (43). Only the binding sites identified by both software were included as final miRNA binding sites. In addition, we provided the conservation information of miRNA binding sites. The UCSC LiftOver tool (44) was used to obtain the conservation information with the parameter of 'minimum ratio of bases that must remap' as 0.5. The miRNA binding sites in the conserved exons of at least two species were classified as conserved. Finally, we identified 45 774 338 lncRNA−miRNA pairs in eight species.

To assess the impact of SNPs and mutations on miRNA–lncRNA interactions, the sequences around each variant (±25 bp) were first extracted. Then, for the sequence of each variant, the wild-type allele was replaced with the alternative allele, and miRmap (42) and TargetScan (43) were used to predict the miRNA binding sites on them. Similar to the former criteria used in lncRNA−miRNA interaction identification, we only kept the miRNA–lncRNA pairs identified by both software. miRNA–lncRNA pairs existing in wild-type transcripts but not in variant alternative transcripts were defined as interaction losses, and on the contrary, they were defined as gains of miRNA target sites. Finally, we found 27 157 121 SNPs and 1 032 392 mutations that potentially caused the gain/loss of original miRNA target sites (Table 1).

lncRNA expression profiles for multiple species

lncRNASNP2 included lncRNA expression profiles obtained from TCGA (45). As TCGA v32.0 updated expression data based on a newer gene annotation file (GENCODE v36) (20), we updated TCGA lncRNA expression profiles from 11 857 to 14 996 human lncRNA genes across 33 human cancer types. In addition to the lncRNA expression profiles from TCGA, we also integrated lncRNA expression profiles from the LncExpDB database (21), which added the number of expressed human lncRNA transcripts to 102 970 (Table 2). In addition, we collected lncRNA expression profiles for pig, mouse, rat, chicken, zebrafish, and fruitfly from LncRBase v2 (35). The number of lncRNAs ranged from 4827 in zebrafish to 130 051 in mouse (Table 2). Detailed information on lncRNA expression can be queried by searching for specific lncRNAs in lncRNASNP v3.

Mutations in lncRNAs

Mutations within lncRNAs have also been proven to play important roles in cancer (46). Using the latest somatic mutation data from TCGA v32.0 and COSMIC v95, we obtained 1 031 639 TCGA mutations and 1 356 046 CosmicNCVs on human lncRNA transcripts, which increased by 3.3-fold and 1.6-fold compared with the lncRNASNP2. We used RNAsnp v1.2 (39) to estimate the effects of TCGA and COSMIC mutations on lncRNA secondary structure. With the threshold of Empirical *P* value <0.2, we obtained 199 450 TCGA lncRNA-mutations and 250 733 COSMIC lncRNA-mutations that potentially affected the lncRNA secondary structure. We next utilized FATHMM (47) to assess whether TCGA mutations in lncRNAs are deleterious. With the threshold of score >0.7, we identified 419 523 (42.11%) pathogenic variants on lncRNA transcripts.

In addition, for each lncRNA gene, we compared the expression of TCGA samples with and without mutations using the Wilcoxon signed-rank test. Under the threshold of nominal *P* value <0.05, expressions of 895 lncRNA genes across 28 cancer types were identified as significantly affected by TCGA mutations.

Disease & GWAS-associated lncRNAs and variants

In recent years, massive research has focused on the roles of lncRNAs in diseases (3,48). Due to the growing need for disease-related lncRNAs, we collected the latest disease-associated lncRNA information from continuously updated databases (Lnc2Cancer 3.0 (49), LncRNADisease 2.0 (50), LncRNAWiki 2.0 (51) and MNDR v3.1 (52)). After data integration and deduplication, we obtained a total of 127 528 disease-lncRNA pairs in human. Since MNDR v3.1 (52) is a comprehensive database providing lncRNA information on multiple species, we also collected 1874 disease-lncRNA pairs in mouse and 111 disease-lncRNA pairs in rat. In addition to lncRNA transcripts, we also matched 135 937 disease-associated variants obtained from ClinVar v4.1 (36) on lncRNAs. Furthermore, to identify the lncRNA-SNPs related to human diseases or complex traits, we first collected 182 272 GWAS tagSNPs (14 222 tagSNPs were located in lncRNA regions) from the NHGRI GWAS Catalog (53) and obtained the LD regions of each GWAS SNP for different populations using plink v1.90 with the '–block' parameter. The phased human SNV files (1000 Genomes 30x) and population information were downloaded from IGSR (54). Finally, 19 populations with more than 100 individuals were included in the above analysis. Our results show that 42 830 177 lncRNA-SNPs are in the GWAS LD regions of all populations (Table 1).

Experimental & predicted lncRNA and drug target associations

Increasing evidence indicates that lncRNAs could be potential therapeutic targets for cancer and disease, and linked to drug resistance (55). The NoncoRNA database (56) is a comprehensive database providing experimentally supported ncRNAs and drug target associations. We therefore integrated NoncoRNA datasets with lncRNAs in lncRNASNP v3 and obtained 4688 drug-associated lncRNA transcripts (Table 1). As associations between lncRNAs and drugs remain largely unknown, we further calculated correlations between lncRNA expression profiles and drug sensitivity data. To do so, we first downloaded the NCI-60 dataset containing half-cell growth inhibition concentrations (GI50) of 24 360 drugs/compounds from the CellMiner Cross Database (57) and then collected the corresponding lncRNA expression profiles of NCI-60 from GSE80332 (58). The correlations were calculated using the Spearman correlation, and the P value of the correlation coefficient was corrected by FDR. We finally kept 10 849 lncRNA-drug pairs with FDR < 0.05 and $|r| > 0.5$ (Table 1).

SNP effects on lncRNA expression (eQTL) across tumor and normal tissues

Expression quantitative trait locus (eQTL) analysis, which links variants in gene expression to genotypes, has been widely used in genetic studies to decipher target genes of functional SNPs (59). Recent studies have shown that SNPs could also exert their roles by regulating the expression of lncRNAs, thereby increasing the risk of cancer (60). In previous years, we have systematically identified eQTLs on ncRNAs across tumor samples and developed the ncRNA-eQTL database (37). To maximize the use of this resource, we matched lncRNA genes in lncRNASNP v3 with the eQTL genes in ncRNA-eQTL (lncRNA genes with the same position were considered to be the same lncRNAs) and extracted all the eQTL SNPs regulating these lncRNAs. We linked 564 157 *cis*-eQTLs, 51 219 *trans*-eQTLs, 7009 GWAS-eQTLs and 2831 survival-eQTLs into lncRNASNP v3 (Table 1). In addition, the GTEx project has also systematically analyzed the associations between SNPs and gene expressions for 'normal' and 'non-disease' tissues (38). Hence, we downloaded GTEx-eQTLs across 49 human tissues and extracted eQTLs with q-values ≤ 0.05. After matching lncRNA genes by position, we obtained 28,019 cis-eQTLs regulating lncRNA genes across normal tissues (Table 1).

DATABASE ORGANIZATION AND WEB INTERFACE

The lncRNASNP v3 database was built with the Flask framework (http://flask.pocoo.org/) as the backend server, and all data mentioned above were organized into MongoDB. The database is freely available at http://gong_lab.hzau.edu.cn/lncRNASNP3. lncRNASNP v3 comprises eight sections: lncRNA, SNP, Mutation, miRNA, eQTL, Disease, Drug and Tool (Figure 1A). On the homepage, users can browse the resources by species or modules (Figure 1B). In lncRNASNP v3, most pages developed in the last release were redesigned for the convenience of searching and browsing information for multiple species (Figure 1C). Users can obtain more information on SNP, miRNA, and lncRNA pages by clicking the 'Detail' button at each retrieved record (Figure 1D). On the 'eQTL' and 'Drug' pages, we displayed results according to different types or different sources. Users can browse distinct information by switching the button above the module (Figure 1E). On the 'Help' page, we provided more detailed information about the data sources, analysis methods, and usage instructions.

SUMMARY AND FUTURE PERSPECTIVES

Since the last release of lncRNASNP in 2018, the numbers of the lncRNAs and SNPs/mutations, especially the number of lncRNAs and SNPs in animals, have increased significantly, attracting increasing attention from animal researchers. Therefore, we updated the lncRNASNP using the latest data, expanded the species to eight commonly studied animals and added several new features. To meet the demand for more comprehensive data, in the current release of lncRNASNP, we expanded the lncRNA expression and disease information for more species. We also developed new features, such as disease-associated lncRNAs and variants, experimental & predicted lncRNA-drug target associations, and TCGA & GTEx eQTLs on lncRNAs, to provide comprehensive insights for SNP and lncRNA-related research. In addition, we redesigned the web interface to make it more convenient for users to obtain the information. In the future, with the cost of sequencing technologies continually decreasing, lncRNAs and SNPs in more species are expected to be identified. We will update the lncRNASNP database regularly and maintain lncRNASNP as a

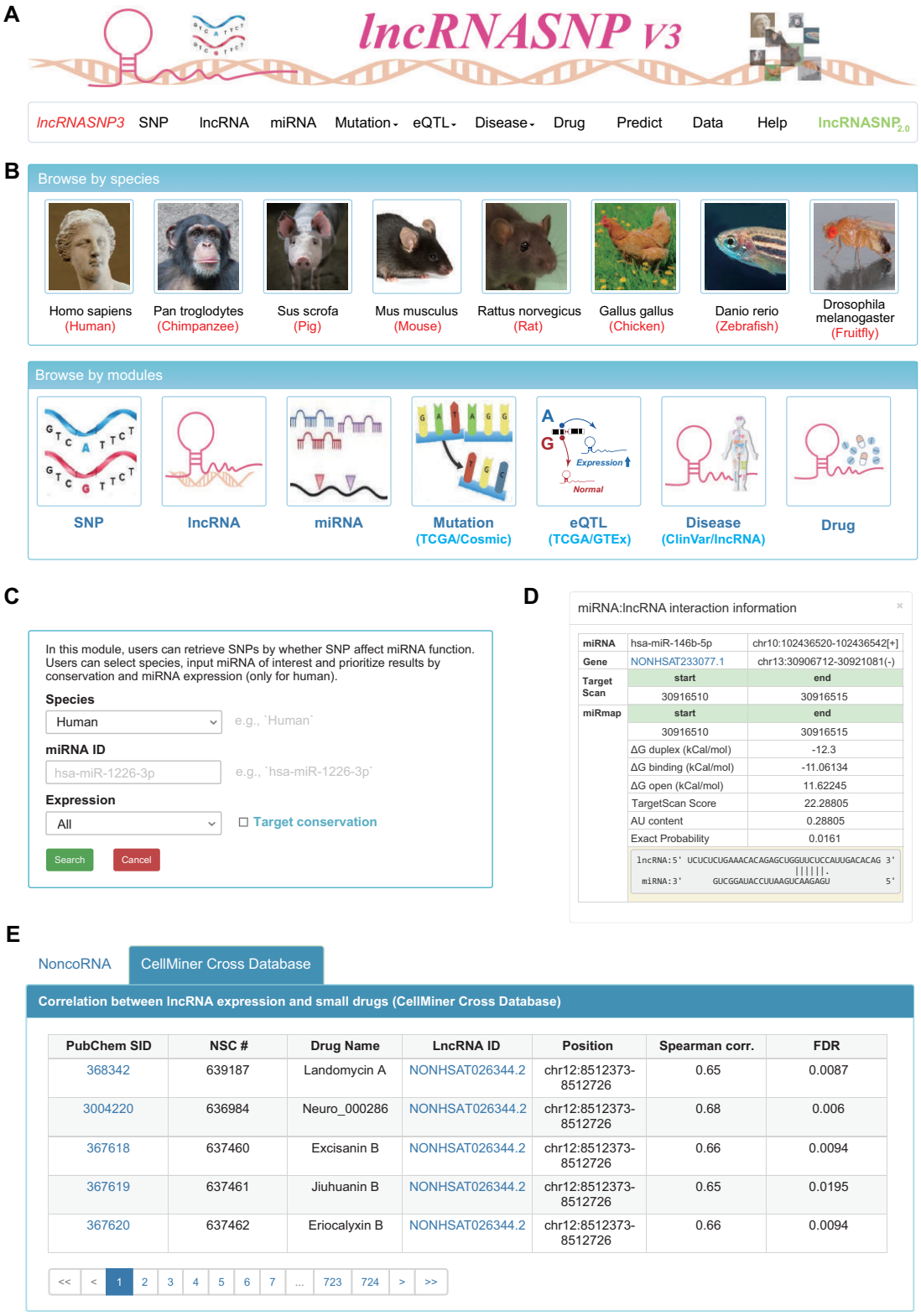

Figure 1. The interface of lncRNASNP v3. (**A**) Navigation in lncRNASNP v3. (**B**) Browse by species or by modules on the homepage. (**C**) The search box on the 'miRNA' page. (**D**) The details of the miRNA:lncRNA interaction after clicking the 'Detail' button on the 'miRNA' page. (**E**) The module for lncRNA expression-drugs correlation results (NCI-60). Users can browse the experimental lncRNAs and drug target associations by clicking the 'NoncoRNA' button.

useful repository for the functional study of lncRNAs and lncRNAs-variants.

DATA AVAILABILITY

lncRNASNP is freely available to the public without registration or login requirements (http://gong_lab.hzau.edu.cn/lncRNASNP3).

FUNDING

National Key R&D Program of China [2021YFF0703700 to J.G. and A.Y.G.]; National Natural Science Foundation of China [31970644 to J.G.]; Huazhong Agricultural University Scientific & Technological Self-innovation Foundation [11041810351 to J.G.]. Funding for open access charge: National Key R&D Program of China [2021YFF0703700].
Conflict of interest statement. None declared.

REFERENCES

1. Hangauer,M.J., Vaughn,I.W. and McManus,M.T. (2013) Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. *PLoS Genet.*, **9**, e1003569.
2. Zhao,L., Wang,J., Li,Y., Song,T., Wu,Y., Fang,S., Bu,D., Li,H., Sun,L., Pei,D. *et al.* (2021) NONCODEV6: an updated database dedicated to long non-coding RNA annotation in both animals and plants. *Nucleic Acids Res.*, **49**, D165–D171.
3. Statello,L., Guo,C.J., Chen,L.L. and Huarte,M. (2021) Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.*, **22**, 96–118.
4. Yao,R.W., Wang,Y. and Chen,L.L. (2019) Cellular functions of long noncoding RNAs. *Nat. Cell Biol.*, **21**, 542–551.
5. Fan,H., Li,J., Wang,J. and Hu,Z. (2019) Long non-coding RNAs (lncRNAs) tumor-suppressive role of lncRNA on chromosome 8p12 (TSLNC8) inhibits tumor metastasis and promotes apoptosis by regulating interleukin 6 (IL-6)/Signal transducer and activator of transcription 3 (STAT3)/Hypoxia-Inducible factor 1-alpha (HIF-1α) signaling pathway in non-small cell lung cancer. *Med. Sci. Monit.*, **25**, 7624–7633.
6. Do,H. and Kim,W. (2018) Roles of oncogenic long Non-coding RNAs in cancer development. *Genomics Inform.*, **16**, e18.
7. Huang,Z., Zhou,J.K., Peng,Y., He,W. and Huang,C. (2020) The role of long noncoding RNAs in hepatocellular carcinoma. *Mol. Cancer*, **19**, 77.
8. Feng,T., Feng,N., Zhu,T., Li,Q., Zhang,Q., Wang,Y., Gao,M., Zhou,B., Yu,H., Zheng,M. *et al.* (2020) A SNP-mediated lncRNA (LOC146880) and microRNA (miR-539-5p) interaction and its potential impact on the NSCLC risk. *J. Exp. Clin. Cancer Res.*, **39**, 157.
9. Fu,Y., Zhang,Y., Cui,J., Yang,G., Peng,S., Mi,W., Yin,X., Yu,Y., Jiang,J., Liu,Q. *et al.* (2020) SNP rs12982687 affects binding capacity of lncRNA UCA1 with miR-873-5p: involvement in smoking-triggered colorectal cancer progression. *Cell Commun. Signal.*, **18**, 37.
10. Khurana,E., Fu,Y., Chakravarty,D., Demichelis,F., Rubin,M.A. and Gerstein,M. (2016) Role of non-coding sequence variants in cancer. *Nat. Rev. Genet.*, **17**, 93–108.
11. Hrdlickova,B., de Almeida,R.C., Borek,Z. and Withoff,S. (2014) Genetic variation in the non-coding genome: involvement of micro-RNAs and long non-coding RNAs in disease. *Biochim. Biophys. Acta*, **1842**, 1910–1922.
12. Gong,J., Liu,W., Zhang,J., Miao,X. and Guo,A.Y. (2015) lncRNASNP: a database of SNPs in lncRNAs and their potential functions in human and mouse. *Nucleic Acids Res.*, **43**, D181–D186.
13. Miao,Y.R., Liu,W., Zhang,Q. and Guo,A.Y. (2018) lncRNASNP2: an updated database of functional SNPs and mutations in human and mouse lncRNAs. *Nucleic Acids Res.*, **46**, D276–D280.
14. Chen,Y.X., Chen,J., Yin,J.Y., Zhou,H.H., He,B.M. and Liu,Z.Q. (2021) Non-Coding RNA polymorphisms (rs2910164 and rs1333049) associated with prognosis of lung cancer under platinum-based chemotherapy. *Front. Pharmacol.*, **12**, 709528.
15. Tsagakis,I., Douka,K., Birds,I. and Aspden,J.L. (2020) Long non-coding RNAs in development and disease: conservation to mechanisms. *J. Pathol.*, **250**, 480–495.
16. Sun,Q., Hao,Q. and Prasanth,K.V. (2018) Nuclear long noncoding RNAs: key regulators of gene expression. *Trends Genet.*, **34**, 142–157.
17. Sherry,S.T., Ward,M.H., Kholodov,M., Baker,J., Phan,L., Smigielski,E.M. and Sirotkin,K. (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.*, **29**, 308–311.
18. Yu,X., Wang,Z., Sun,H., Yang,Y., Li,K. and Tang,Z. (2018) Long non-coding MEG3 is a marker for skeletal muscle development and meat production traits in pigs. *Anim. Genet.*, **49**, 571–578.
19. Mei,X., Kang,X., Liu,X., Jia,L., Li,H., Li,Z. and Jiang,R. (2016) Identification and SNP association analysis of a novel gene in chicken. *Anim. Genet.*, **47**, 125–127.
20. Frankish,A., Diekhans,M., Jungreis,I., Lagarde,J., Loveland,J.E., Mudge,J.M., Sisu,C., Wright,J.C., Armstrong,J., Barnes,I. *et al.* (2021) gencode 2021. *Nucleic Acids Res.*, **49**, D916–D923.
21. Li,Z., Liu,L., Jiang,S., Li,Q., Feng,C., Du,Q., Zou,D., Xiao,J., Zhang,Z. and Ma,L. (2021) LncExpDB: an expression database of human long non-coding RNAs. *Nucleic Acids Res.*, **49**, D962–D968.
22. Das,T., Deb,A., Parida,S., Mondal,S., Khatua,S. and Ghosh,Z. (2021) LncRBase V.2: an updated resource for multispecies lncRNAs and ClinicLSNP hosting genetic variants in lncRNAs for cancer patients. *RNA Biol.*, **18**, 1136–1151.
23. Hutter,C. and Zenklusen,J.C. (2018) The cancer genome atlas: creating lasting value beyond its data. *Cell*, **173**, 283–285.
24. Volders,P.J., Anckaert,J., Verheggen,K., Nuytens,J., Martens,L., Mestdagh,P. and Vandesompele,J. (2019) LNCipedia 5: towards a reference set of human long non-coding RNAs. *Nucleic Acids Res.*, **47**, D135–D139.
25. Quek,X.C., Thomson,D.W., Maag,J.L.V., Bartonicek,N., Signal,B., Clark,M.B., Gloss,B.S. and Dinger,M.E. (2015) lncRNAdb v2.0: expanding the reference database for functional long noncoding RNAs. *Nucleic Acids Res.*, **43**, D168–D173.
26. 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.
27. Li,J.H., Liu,S., Zhou,H., Qu,L.H. and Yang,J.H. (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.
28. 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.
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 (Oxford)*, **2013**, bat034.
30. Bao,Z., Yang,Z., Huang,Z., Zhou,Y., Cui,Q. and Dong,D. (2019) LncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases. *Nucleic Acids Res.*, **47**, D1034–D1037.
31. Gao,Y., Li,X., Shang,S., Guo,S., Wang,P., Sun,D., Gan,J., Sun,J., Zhang,Y., Wang,J. *et al.* (2021) LincSNP 3.0: an updated database for linking functional variants to human long non-coding RNAs, circular RNAs and their regulatory elements. *Nucleic Acids Res.*, **49**, D1244–D1250.
32. Chen,X., Hao,Y., Cui,Y., Fan,Z., He,S., Luo,J. and Chen,R. (2017) LncVar: a database of genetic variation associated with long non-coding genes. *Bioinformatics*, **33**, 112–118.
33. Cezard,T., Cunningham,F., Hunt,S.E., Koylass,B., Kumar,N., Saunders,G., Shen,A., Silva,A.F., Tsukanov,K., Venkataraman,S. *et al.* (2022) The european variation archive: a FAIR resource of genomic variation for all species. *Nucleic Acids Res.*, **50**, D1216–D1220.
34. Tate,J.G., Bamford,S., Jubb,H.C., Sondka,Z., Beare,D.M., Bindal,N., Boutselakis,H., Cole,C.G., Creatore,C., Dawson,E. *et al.* (2019) COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res.*, **47**, D941–D947.
35. Das,T., Deb,A., Parida,S., Mondal,S., Khatua,S. and Ghosh,Z. (2021) LncRBase V.2: an updated resource for multispecies lncRNAs

and ClinicLSNP hosting genetic variants in lncRNAs for cancer patients. *RNA Biol.*, **18**, 1136–1151.
36. Landrum,M.J., Chitipiralla,S., Brown,G.R., Chen,C., Gu,B., Hart,J., Hoffman,D., Jang,W., Kaur,K., Liu,C. *et al.* (2020) ClinVar: improvements to accessing data. *Nucleic Acids Res.*, **48**, D835–D844.
37. Li,J., Xue,Y., Amin,M.T., Yang,Y., Yang,J., Zhang,W., Yang,W., Niu,X., Zhang,H.Y. and Gong,J. (2020) ncRNA-eQTL: a database to systematically evaluate the effects of SNPs on non-coding RNA expression across cancer types. *Nucleic Acids Res.*, **48**, D956–D963.
38. The genotype-tissue expression (GTEx) project (2013) *Nat. Genet.*, **45**, 580–585.
39. Sabarinathan,R., Tafer,H., Seemann,S.E., Hofacker,I.L., Stadler,P.F. and Gorodkin,J. (2013) The RNAsnp web server: predicting SNP effects on local RNA secondary structure. *Nucleic Acids Res.*, **41**, W475–W479.
40. Ghaedi,H., Bastami,M., Zare-Abdollahi,D., Alipoor,B., Movafagh,A., Mirfakhraie,R., Omrani,M.D. and Masotti,A. (2015) Bioinformatics prioritization of SNPs perturbing microRNA regulation of hematological malignancy-implicated genes. *Genomics*, **106**, 360–366.
41. Kozomara,A., Birgaoanu,M. and Griffiths-Jones,S. (2019) miRBase: from microRNA sequences to function. *Nucleic Acids Res.*, **47**, D155–D162.
42. Vejnar,C.E. and Zdobnov,E.M. (2012) MiRmap: comprehensive prediction of microRNA target repression strength. *Nucleic Acids Res.*, **40**, 11673–11683.
43. Agarwal,V., Bell,G.W., Nam,J.-W. and Bartel,D.P. (2015) Predicting effective microRNA target sites in mammalian mRNAs. *Elife*, **4**,e05005 .
44. Karolchik,D., Hinrichs,A.S., Furey,T.S., Roskin,K.M., Sugnet,C.W., Haussler,D. and Kent,W.J. (2004) The UCSC table browser data retrieval tool. *Nucleic Acids Res.*, **32**, D493–D496.
45. Cancer Genome Atlas Network (2012) Comprehensive molecular portraits of human breast tumours. *Nature*, **490**, 61–70.
46. Fujimoto,A., Furuta,M., Totoki,Y., Tsunoda,T., Kato,M., Shiraishi,Y., Tanaka,H., Taniguchi,H., Kawakami,Y., Ueno,M. *et al.* (2016) Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. *Nat. Genet.*, **48**, 500–509.
47. Rogers,M.F., Shihab,H.A., Mort,M., Cooper,D.N., Gaunt,T.R. and Campbell,C. (2018) FATHMM-XF: accurate prediction of pathogenic point mutations via extended features. *Bioinformatics*, **34**, 511–513.
48. Ni,W., Yao,S., Zhou,Y., Liu,Y., Huang,P., Zhou,A., Liu,J., Che,L. and Li,J. (2019) Long noncoding RNA GAS5 inhibits progression of colorectal cancer by interacting with and triggering YAP phosphorylation and degradation and is negatively regulated by the m6A reader YTHDF3. *Mol. Cancer*, **18**, 143.
49. Gao,Y., Shang,S., Guo,S., Li,X., Zhou,H., Liu,H., Sun,Y., Wang,J., Wang,P., Zhi,H. *et al.* (2021) Lnc2Cancer 3.0: an updated resource for experimentally supported lncRNA/circRNA cancer associations and web tools based on RNA-seq and scRNA-seq data. *Nucleic Acids Res.*, **49**, D1251–D1258.
50. Bao,Z., Yang,Z., Huang,Z., Zhou,Y., Cui,Q. and Dong,D. (2019) LncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases. *Nucleic Acids Res.*, **47**, D1034–D1037.
51. Liu,L., Li,Z., Liu,C., Zou,D., Li,Q., Feng,C., Jing,W., Luo,S., Zhang,Z. and Ma,L. (2022) LncRNAWiki 2.0: a knowledgebase of human long non-coding RNAs with enhanced curation model and database system. *Nucleic Acids Res.*, **50**, D190–D195.
52. Ning,L., Cui,T., Zheng,B., Wang,N., Luo,J., Yang,B., Du,M., Cheng,J., Dou,Y. and Wang,D. (2021) MNDR v3.0: mammal ncRNA-disease repository with increased coverage and annotation. *Nucleic Acids Res.*, **49**, D160–D164.
53. Buniello,A., MacArthur,J.A.L., Cerezo,M., Harris,L.W., Hayhurst,J., Malangone,C., McMahon,A., Morales,J., Mountjoy,E., Sollis,E. *et al.* (2019) The NHGRI-EBI GWAS catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.*, **47**, D1005–D1012.
54. Fairley,S., Lowy-Gallego,E., Perry,E. and Flicek,P. (2020) The international genome sample resource (IGSR) collection of open human genomic variation resources. *Nucleic Acids Res.*, **48**, D941–D947.
55. Smallegan,M.J. and Rinn,J.L. (2019) Linking long noncoding RNA to drug resistance. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 21963–21965.
56. Li,L., Wu,P., Wang,Z., Meng,X., Zha,C., Li,Z., Qi,T., Zhang,Y., Han,B., Li,S. *et al.* (2020) NoncoRNA: a database of experimentally supported non-coding RNAs and drug targets in cancer. *J. Hematol. Oncol.*, **13**, 15.
57. Luna,A., Elloumi,F., Varma,S., Wang,Y., Rajapakse,V.N., Aladjem,M.I., Robert,J., Sander,C., Pommier,Y. and Reinhold,W.C. (2021) CellMiner cross-database (CellMinerCDB) version 1.2: exploration of patient-derived cancer cell line pharmacogenomics. *Nucleic Acids Res.*, **49**, D1083–D1093.
58. Mestdagh,P., Lefever,S., Volders,P.J., Derveaux,S., Hellemans,J. and Vandesompele,J. (2016) Long non-coding RNA expression profiling in the NCI60 cancer cell line panel using high-throughput RT-qPCR. *Sci. Data*, **3**, 160052.
59. Gong,J., Mei,S., Liu,C., Xiang,Y., Ye,Y., Zhang,Z., Feng,J., Liu,R., Diao,L., Guo,A.Y. *et al.* (2018) PancanQTL: systematic identification of cis-eQTLs and trans-eQTLs in 33 cancer types. *Nucleic Acids Res.*, **46**, D971–D976.
60. Li,D., Song,L., Wen,Z., Li,X., Jie,J., Wang,Y. and Peng,L. (2016) Strong evidence for LncRNA ZNRD1-AS1, and its functional Cis-eQTL locus contributing more to the susceptibility of lung cancer. *Oncotarget*, **7**, 35813–35817.