Help

The Database

The **lncRNAome** is a knowledgebase of long noncoding RNA, which contains detailed information of 17547 lncRNAs. The data for lncRNA has been compiled from GENCODE v12. The details about each RNA are tabulated in a manner, making it easy to use and interpret. The 'Browse' and 'Search' options provide a systematic approach for accessing the database.

Browse Options

The entire data on lncRNAs has been sorted based on the chromosome number to which each RNA belongs. The RNAs can be browsed chromosome wise by clicking on each chromosome button. Click on a chromosome opens a 'Search' tab which displays the details of lncRNA in a tabulated manner. Users can in addition browse based on the type of lncRNAs or even by genomic locations using the Genome Browser option.

Search

Search can be performed in following 3 ways.

1. The Data table -

The table obtained after chromosome search provides the following details:

- Local ID – A unique ID provided to each lncRNA.
- Gene ID - These are the Ensemble gene Ids. The Ids provide a link to the Ensemble Browser.
- Gene name
- Gene Type – The gene can be Antisense, lincRNA, Processed transcript, Sense intronic
- Transcript ID – The RNA transcript ID also provides a direct link to the Ensemble Browser.
- Transcript name
- Transcript Type – The transcripts can be Antisense, LincRNA, Noncoding, Processed Transcript, Retained Intron, Sense Intronic
- HGNC ID – Provides a direct link to the HUGO Gene Nomenclature Committee
- Chromosome number

Example - Search

Browsing for lncRNAs on chromosome number 1 opens a search tab as follows:

2. The Search box –
The box provided at the top of the page allows search by keywords, making the search more specific. The box looks up for gene name, transcript name, transcript ID, single nucleotide polymorphism, disease associations and other information.

3. RNA Advanced Search –
In order to obtain more specific outcomes this feature provides filters which can be set as per the users requirements. The options in each category appear as a drop down. The filters provided are : Chromosome number, Genomic loci start and end, Gene type, Transcript type, Disorder associations, Other associations.

Are lncRNAs functional ?
There is significant evidence that suggests lncRNAs are functional and evidence and molecular charecterisation of their roles in disease pathogenesis is established. A compilation of lncRNAs in disease processes is available below. If you use this data, please cite

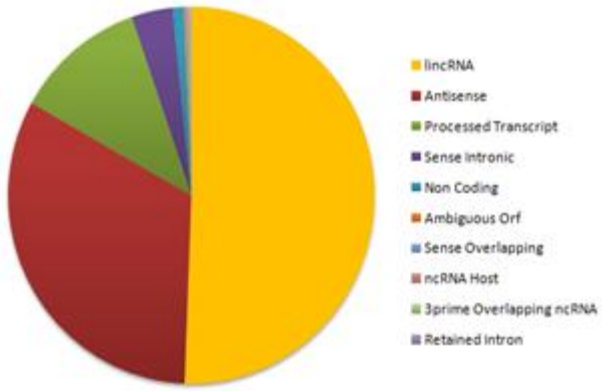
Bhartiya D, Kapoor S, Jalali S, Sati S, Kaushik K, Sachidanandan C, Sivasubbu S, Scaria V
Conceptual approaches for lncRNA drug discovery and future strategies
Expert Opinion on Drug Discovery (2012) In Press

Long noncoding RNA Name	**Biomolecular Interactions**	**Disease Associations**
H19	RNA-DNA(IGF2)	Lung, liver, breast, colorectal, prostate, ovary, cervix, esophagus and bladder cancer.
HOTAIR	RNA-Protein (Polycomb Repressive Complex 2 (PRC2)	Breast cancer, hepatocellular carcinoma
MALAT1 / NEAT2	RNA- protein (SR splicing Protein), RNA-DNA (motility-related genes)	Cancers of Breast, cervix, colon, lung, uterus, pancreas, liver, osteosarcoma, neuroblastoma

HULC	RNA-DNA(CREB)	Hepatocellular carcinomas and hepatic colorectal metastasis
PCGEM1	RNA-Protein(PARP cleavage inhibition), RNA-DNA(p53,p21)	Prostate cancer
PRNCR1		Prostate cancer
BIC / MIRHG2		B cell lymphoma
DD3		Prostate cancer
MEG3	RNA-DNA (p53)	Brain tumor
PTENP1		Prostate cancer
LSINCT5		Ovarian and breast cancer
BACE1AS	RNA-RNA (BACE1 RNA stabilization)	Alzheimer's Disease
FMR4	RNA-DNA(antiapoptotic genes)	Fragile X syndrome
PRINS	RNA-DNA(G1P3)	Psoriasis
AK082072		Autism
BC200	RNA-RNA(translational repressor)	Breast cancer, Aging, Alzheimer's disease
LOC285194 / BC040587		Osteosarcoma
UCA1	RNA-DNA (WNT6, CYP1A1, AURKC)	Bladder Carcinoma, Pancreaticobiliary maljunction
anti-NOS2A		Brain cancer like meningiomas and glioblastomas
TUC338		Hepatocellular carcinoma
Anril	RNA-Protein(SUZ12: a component of PRC2)	Ankle-Brachial Index, Neurofibromatosis, coronary disease, intracranial aneurysm and also type 2 diabetes
Gas5	RNA-Protein(GR transcription factor)	Controls Apopotosis/ Growth Arrest, Downregulated in breast cancer
XIST	RNA-DNA (X chr inactivation)	
SPRY4-IT1		Melanoma
Linc-MD1		Muscle Differentiation
NEAT1/VINC	RNA-protein(P54nrb: paraspeckle protein)	
TUG1	RNA-Protein (PRC2)	Huntington
TSIX	RNA-DNA (CTCF), RNA-RNA (XIST)	
TelRNAs/TERRA	RNA-DNA (chromatin modification at telomeres)	
SRA	RNA-DNA (AF-1 activity of	Breast cancer, Prostrate tumorigenesis

	ER alpha.	
SCA8		Spinocerebellar ataxia type8
Nkx2.2AS	RNA-DNA (Nk2.2,Nk2.4, Nlkx2.9,Nlkx2.5, Hmx3)	
Nespas	Gnas	Pseudohypoparathyroidism type Ib
KCNQ1OT1	RNA-DNA (ICR2 hypomethylation)	Beckwith-Wiedemann syndrome and Silver-Russell syndrome
NDM29	RNA-DNA (cell differentiation genes)	Neuroblastoma
LOC285194	RNA-DNA (apoptotic gene expression)	Osteosarcoma
Gadd7		ER stress leading to lipotoxicity
Gomafu/ MIAT	RNA-Protein (nuclear protein)	Myocardial infarction

Database Statistics
10840 of lncRNA genes
17547 Alternate splice variants
Statistics of subtypes/subclasses



4491 RNA-Protein interactions
310010 Variations mapping to lncRNAs
More than 40,000 of epigeentic markers mapping around lncRNA TSS

How was the Expression data derived ?
The Affy ID mappings for Affymetrix chipsets u133a and u133a_plus_2 for each of the lncRNAs were retrieved from Ensembl using BioMart. The expression of the transcripts in different tissues for each of the affymetrix probeID was retrieved from the Barcode 2.0 Project Catalog http://rafalab.jhsph.edu/barcode/index.php?page=catalog The datasets were parsed using custom scripts in Perl.

What Databases , Resources and Tools were used for creating lncRNome ?
The lncRNAs were derived from Gencode Project as well as literature evidence
Annotations were derived from literature.
Annotations and IDs were derived for relevant candidates from HGNC Mappings, Ensembl IDs and Vega IDs were retrieved from Ensembl.

Genomic variations are derived from dbSNP.

Expression data for the Affymetrix probesets were derived from Barcode 2.0 Project Catalog
http://rafalab.jhsph.edu/barcode/index.php?page=catalog.

RNA secondary structures were computed using Vienna RNA Package.

RNA Protein interactions were mapped based on CLIP-seq datasets.

microRNA binding sites were mapped using Ago-CLIP-seq datasets.

Cite for CLIP-seq datasets used Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M Jr, Jungkamp AC, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T (2010), "Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP", *Cell* **141** (1): 129–141

Epigenetic marks were mapped from raw reads and computed using MACS. The raw datasets were derived from the Roadmap Epigenomics Project.

SmallRNA processing was derived from DeepBase
Cite: Jian-Hua Yang, Peng Shao, Hui Zhou, Yue-Qin Chen, and Liang-Hu Qu.
deepBase: a database for deeply annotating and mining deep sequencing data.
Nucleic Acids Res. 2010 Jan;38(Database issue):D123-D130. Epub 2009 Dec 4.