

人类lncRNA数据资源的使用及数据挖掘分析

马利娜 博士 青年研究员
非编码RNA小组负责人

生命组学数据汇交管理及共享应用培训
2024.10.25



中国科学院北京基因组研究所（国家生物信息中心）

BEIJING INSTITUTE OF GENOMICS CHINESE ACADEMY OF SCIENCES / CHINA NATIONAL CENTER FOR BIOINFORMATION

人类lncRNA参考数据集 (LncBook)

人类lncRNA大规模整合与鉴定

268,848 transcripts



- BLAST comparison

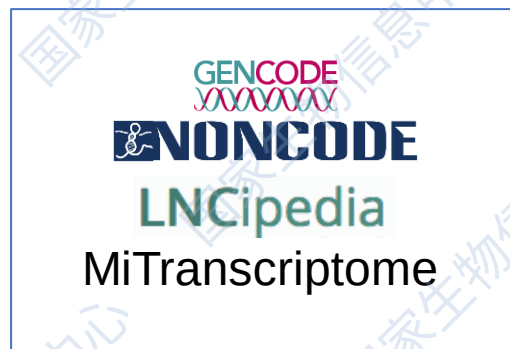
LncRNAWiki

2015

Nucleic Acids Res 2015, 2022

140,356 genes

268,848 transcripts



- Cuffcompare comparison
- Curation of high-confidence lncRNAs
- Gene assignment

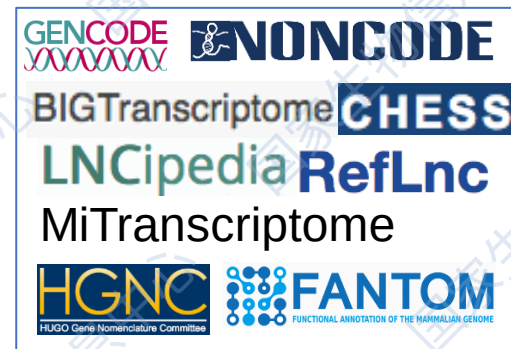
LncBook

2019

Nucleic Acids Res 2019

101,293 genes

331,244 transcripts



- Removal of pre-mRNAs
- GffCompare comparison
- Curation of high-confidence lncRNAs
- Gene assignment

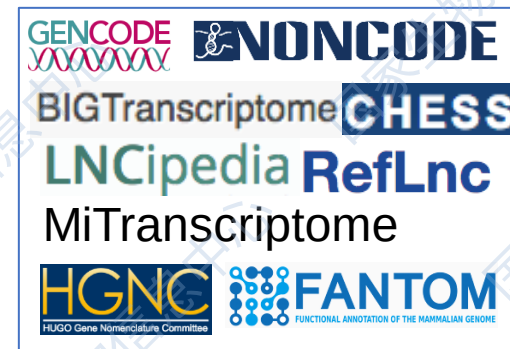
LncExpDB

2021

Nucleic Acids Res 2021

95,243 genes

323,950 transcripts



- Removal of miRNA presursors, small RNAs and pseudogenes
- Curation of high-confidence lncRNAs
- Gene assignment

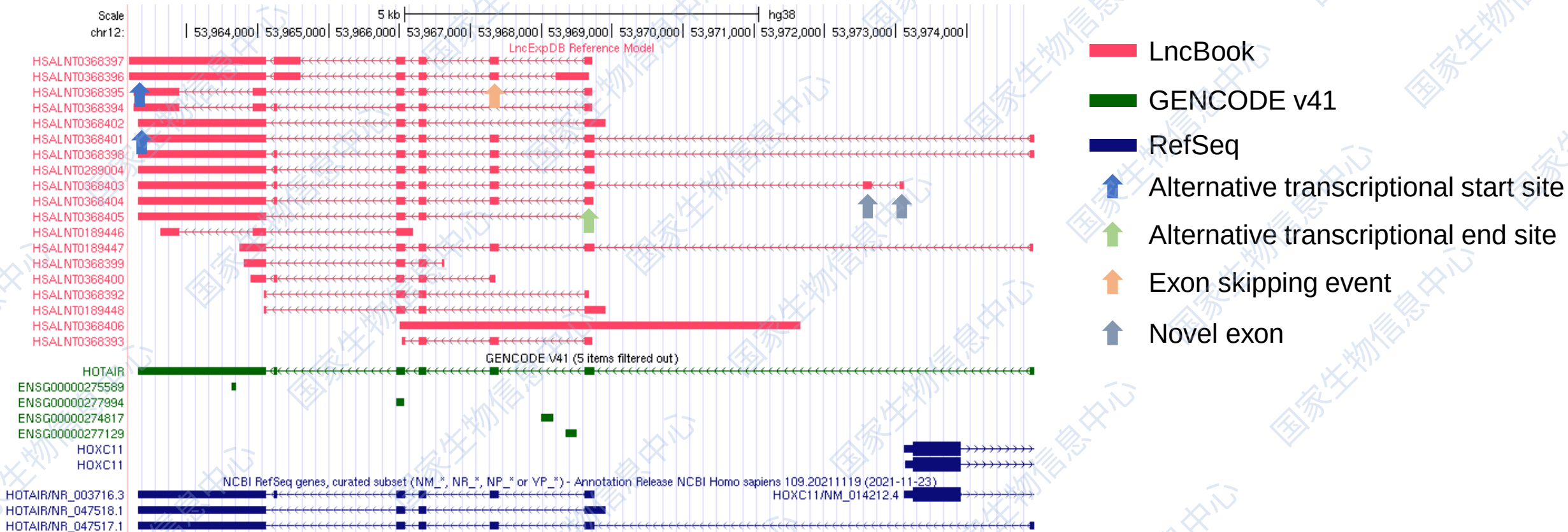
LncBook 2.0

2022

Nucleic Acids Res 2023

LncRNA基因的完整性

HOTAIR gene loci in LncBook, GENCODE and RefSeq



<https://ngdc.cncb.ac.cn/lncbook/gene?geneid=HSALNG0091318>

LncBook数据库：提供全面高质量的人类lncRNA参考数据集

[Home](#) > [Downloads](#)

LncRNA reference

Description	File Name	File Size
New! Description for LncRNA GTF format	README.txt	8.0 KB
New! Merged GTF file; 526,318 lncRNA transcripts of LncBook v2.1	lncRNA_LncBookv2.1_GRCh38.gtf.gz	28 MB
New! FASTA sequences of LncBook v2.1 lncRNA transcripts; 526,318 sequences	lncRNA_LncBookv2.1_GRCh38.fa.gz	156 MB
 Merged GTF file; 95,243 lncRNA genes of LncBook v2.0 and 43,770 other types of genes from GENCODE v34	LncBookv2.0_GENCODEv34_GRCh38.gtf.gz	57.84 MB
 Curated based on LncBook v1.9; 95,243 lncRNA genes	lncRNA_LncBookv2.0_GRCh38.gtf.gz	15.20 MB
 FASTA sequences of LncBook v2.0 lncRNA transcripts; 323,950 sequences	lncRNA_LncBookv2.0.fa.gz	107.47 MB
 Merged GTF file; 101,293 lncRNA genes of LncBook v1.9, and 19,957 protein coding genes from GENCODE v33; reference for LncExpDB analysis	LncBookv1.9_GENCODEv33_GRCh38.gtf.gz	54.11 MB
 101,293 lncRNA genes	lncRNA_LncBookv1.9_GRCh38.gtf.gz	14.90 MB
 FASTA sequences of LncBook v1.9 lncRNA transcripts; 331,420 sequences	lncRNA_LncBookv1.9.fa.gz	108.67 MB

<https://ngdc.cncb.ac.cn/lncbook/downloads>

目前研究内容：注释高准确度的lncRNA转录本

整合新的lncRNA并基于长读长测序数据注释lncRNA转录本

转录本来源

lncRNA 资源整合 转录本

来自10个数据库：
GENCODE, NONCODE,
LNCipedia,
MiTranscriptome beta,
RefLnc, CHESSE,
FANTOM-CAT,
BIGTranscriptome,
RNAAtlas, TAiC

长读长 数据识别 转录本

基于94套 PacBio 转录组测
序数据

lncRNA 审编与质控

转录本过滤

冗余转录本

- 外显子一致

背景噪音

- 比对错误
- Polymerase run-on

不完整 或不可信

- 5'/3'端不确定
(5'/3'外显子 ≤ 15 nt
或经 Bambu 修正)
- 无链信息

非 lncRNAs

- 长度 < 200 nt
- miRNA 前体
- Small RNAs
- 潜在 mRNA

LncBook v2.1

LncRNA 转录本

- 526,318 个 lncRNA 转录本

参考基因组

- GRCh38.p14

文件类型

- GTF 文件：人类基因组 GRCh38 的 lncRNA 注释
- FASTA 文件：lncRNA 转录本的序列信息
- README 文件：对以上文件的详细说明

人类lncRNA基因功能注释 (LncExpDB+LncBook)

LncExpDB: 人类lncRNA基因表达图谱数据库

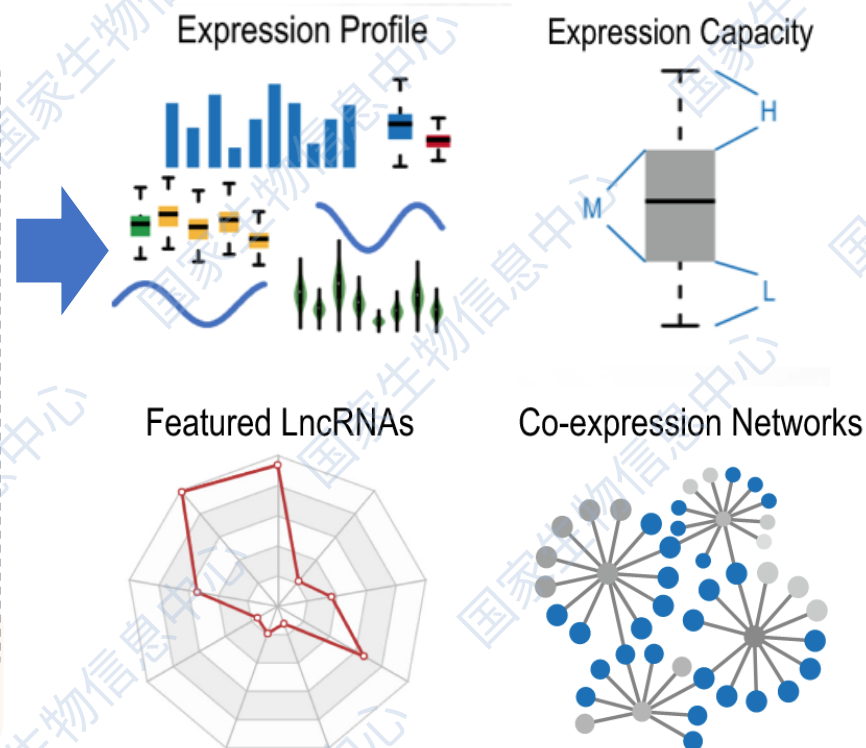
9 大生物学场景

- Normal tissue/cell line
- Cancer cell line
- Subcellular localization
- Exosome
- Cell differentiation
- Preimplantation embryo
- Organ development
- Circadian rhythm
- Virus infection

- 14 个项目
- 337 生物学状态
- 1,977个样本
- 24套RNA-seq数据集



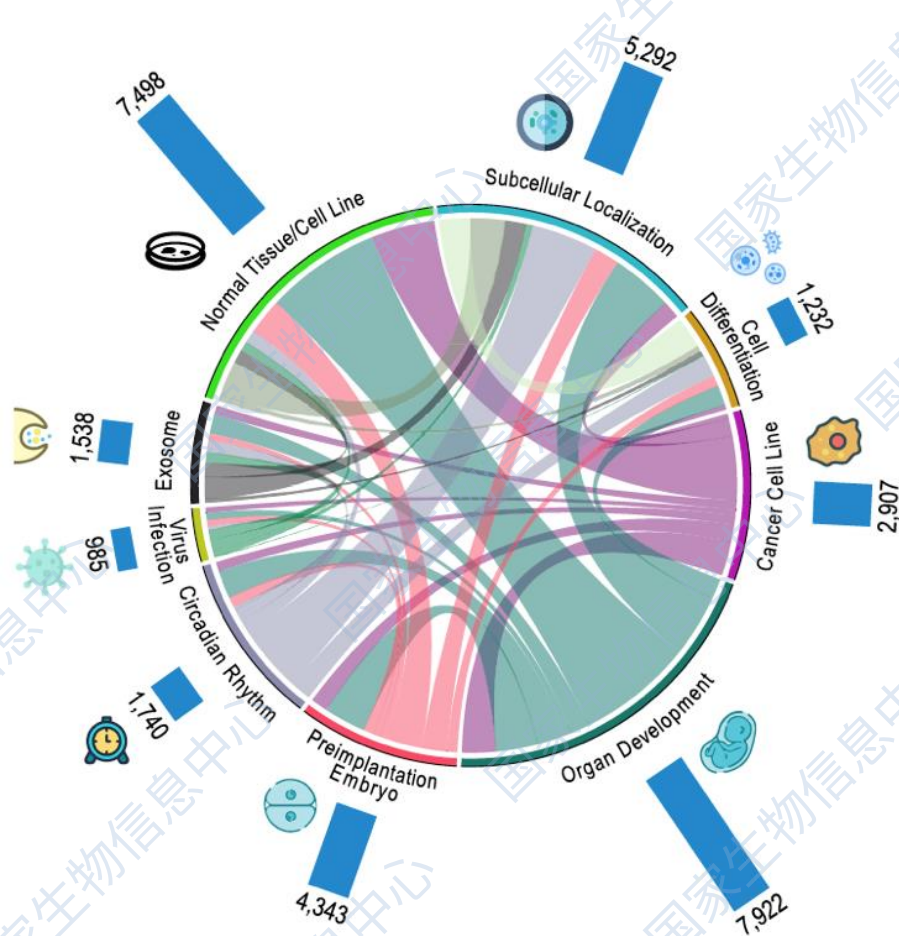
表达特征刻画



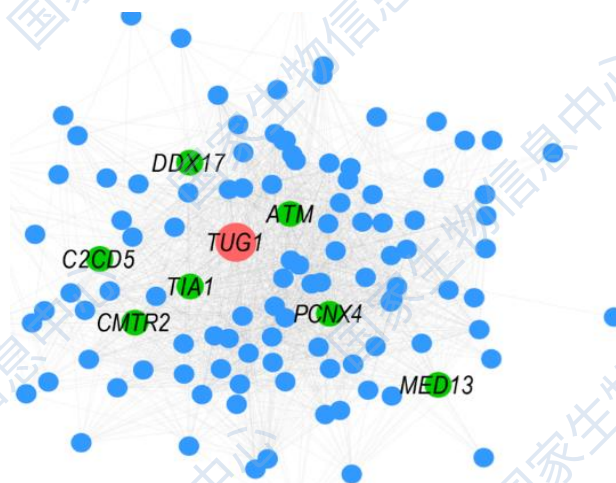
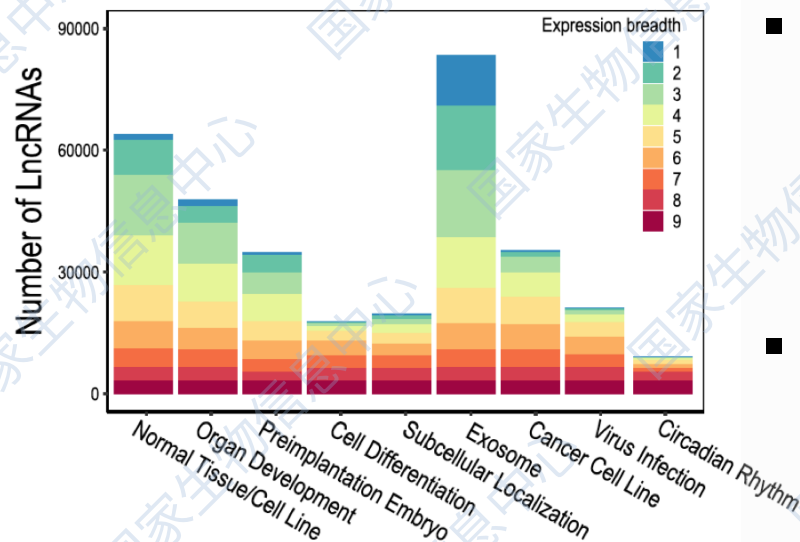
LncExpDB, Li et al. *Nucleic Acids Res* 2021

<https://ngdc.cncb.ac.cn/lncexpdb>

人类lncRNA基因的表达特征



<https://ngdc.cncb.ac.cn/lncexpdb>



Li et al. Nucleic Acids Res 2021

- 92,016 lncRNA genes (**90.8%**) with reliable evidence
 - ✓ **34%**: HCL
 - ✓ 55.7%: MCL
 - ✓ 10.3%: LCL
- **25,191** featured lncRNA genes, which are specifically/differentially/dynamically/periodically expressed
- Example: **TUG1**
 - ✓ featured in 5 biological contexts
 - ✓ co-expressed with 114 mRNAs in at least 4 biological contexts
 - ✓ highly expressed in 321 conditions

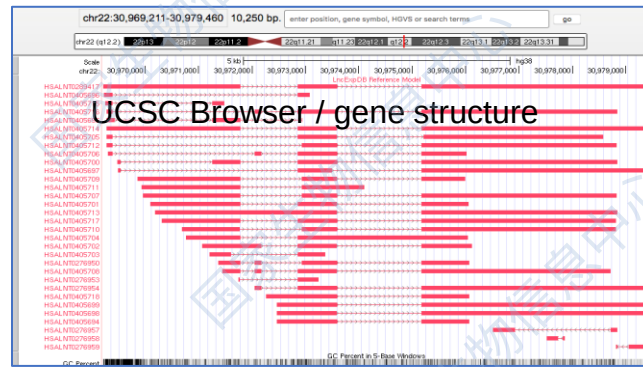
TUG1多生物学场景表达谱可视化

Summary

Gene information

Gene ID:	HSALNG0134930	Classification:	Intergenic
Gene Symbol:	TUG1	Gene Length (nt):	10250
Chromosome:	chr22	Exon Length (nt):	10250
Start Site:	30969211	UCSC Genome Browser:	View
End Site:	30979460	Assembly:	hg38
Strand:	+		

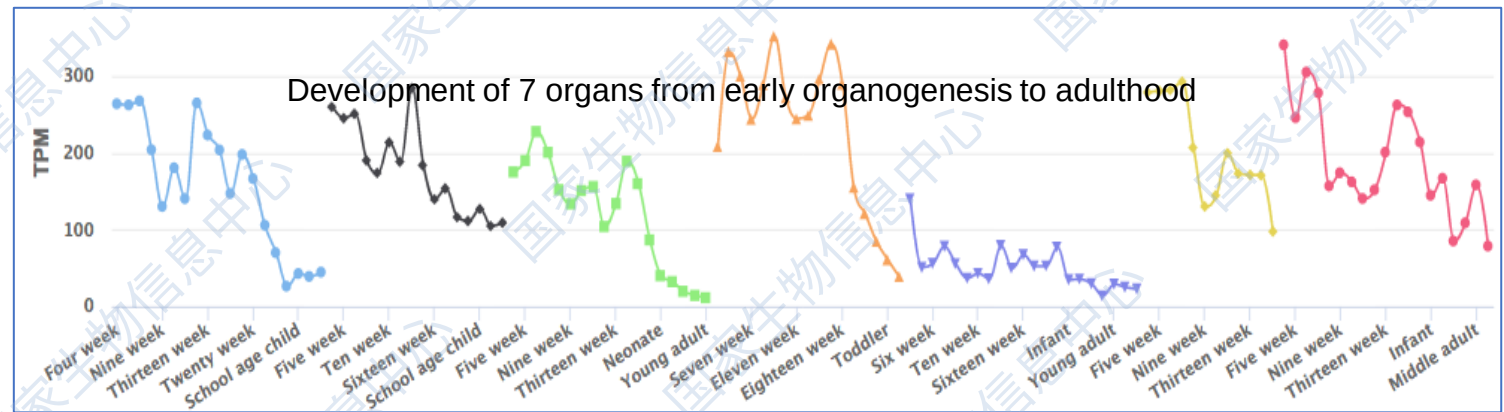
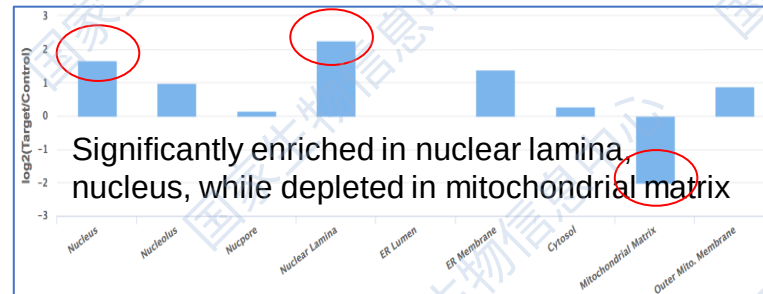
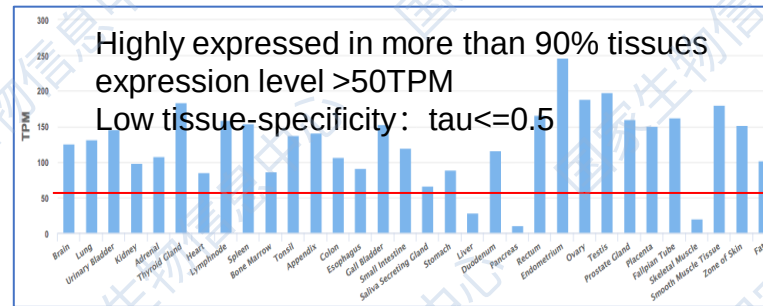
External Links: [ENSG00000253352.9 \(GENCODE\); TUG1 \(LNCipedia\); NONHSAG033691.3 \(NONCODE\); NR_002323.2 \(RefSeq\); CHS.35608 \(CHES\); FANTOM_ENSG00000253352.4 \(FANTOM\);](#)



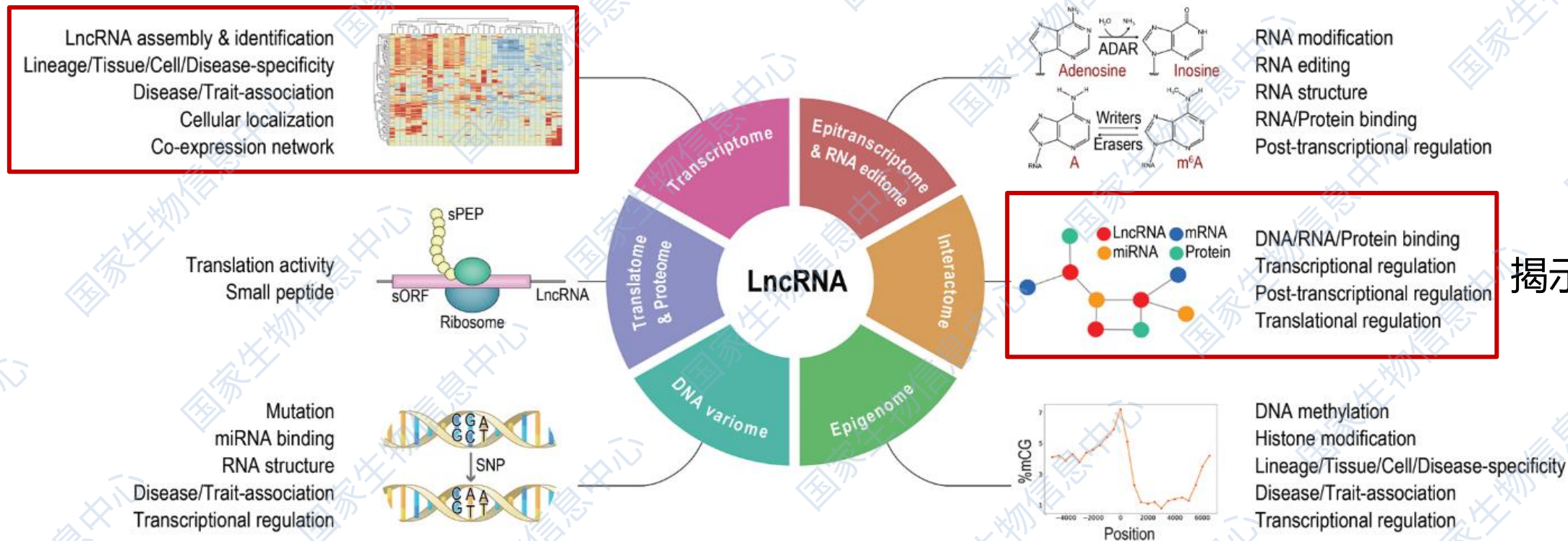
Transcript information

Transcript ID	Chromosome	Strand	Start Site	End Site	Transcript Length	Exon Length	Exon Number
HSALNT0276950	chr22	+	30971367	30976063	4697	2110	4
HSALNT0276953	chr22	+	30971745	30973245	1501	416	2
HSALNT0276954	chr22	+	30972045	30973351	7307	5057	3
HSALNT0286417	chr22	+	30969211	30979395	10185	7542	3
HSALNT0405694	chr22	+	30972466	30978063	3598	2030	2
HSALNT0405695	chr22	+	30969269	30972938	3670	695	3
HSALNT0405696	chr22	+	30969223	30973063	3861	393	2
HSALNT0405697	chr22	+	30969490	30979450	9961	4968	3
HSALNT0405698	chr22	+	30972466	30979460	6995	5413	2
HSALNT0405699	chr22	+	30972466	30979460	6995	5427	2

Showing 1 to 10 of 29 entries



基于多组学数据整合分析注释lncRNA



Li, Q. et al. *Biochem Soc Trans*, 2020

LncRNA多组学注释可系统全面地刻画lncRNA功能；**转录组**具有丰富的场景信息；**互作组**揭示调控关系

LncBook多组学特征注释与比较分析

40
species for
conservation

24,157
featured
expression

19,543
featured
methylation

34,012
small proteins
(5743)

959,138
trait variants

146,092,274
lncRNA-
miRNA
interactions

772,745
lncRNA-
protein
interactions

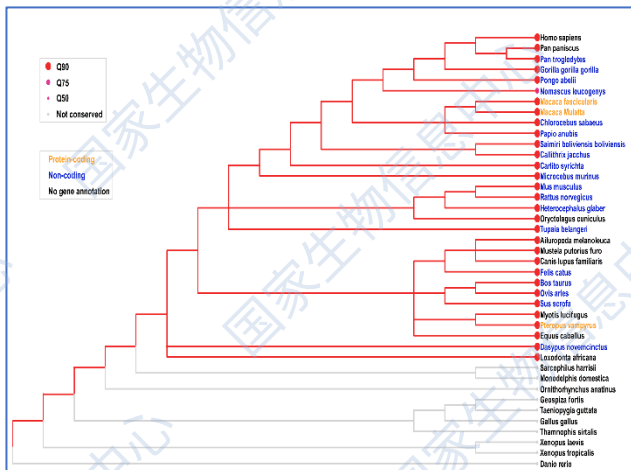
Gene ID	Symbol	Conservation (age)	Expression Capacity	Featured Expression	Featured Methylation	Variation (disease/ trait-association)	Interaction	Small Protein
HSALNG0012977	ODC1-DT	14. Mammalia	HC	✓	✓	✓	✗	✓
HSALNG0084892	NEAT1	14. Mammalia	HC	✓	✓	✓	✓	✓
HSALNG0088950	MIR200CHG	14. Mammalia	HC	✓	✓	✓	✗	✓
HSALNG0075219	LINC02846	14. Mammalia	HC	✓	✓	✗	✗	✓
HSALNG0100193	LINC02306	14. Mammalia	HC	✓	✓	✓	✗	✓
HSALNG0119134	LINC01977	14. Mammalia	HC	✓	✓	✓	✗	✗
HSALNG0000162	LINC01770	14. Mammalia	HC	✓	✓	✓	✓	✓
HSALNG0002236	LINC01226	14. Mammalia	HC	✓	✓	✓	✓	✓
HSALNG0099129	LINC00343	14. Mammalia	HC	✓	✓	✓	✗	✗
HSALNG0143634	KRT7-AS	14. Mammalia	HC	✓	✓	✓	✗	✗

<https://ngdc.cncb.ac.cn/lncbook/genes>

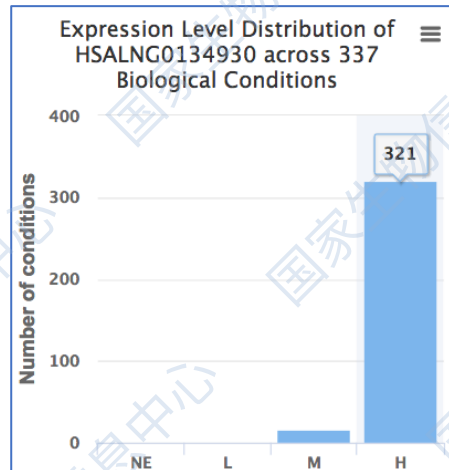
TUG1 在LncBook中的多组学分子特征

Classification ?	Intergenic
Gene Length (nt)	10,250
Exon Length (nt)	10,250
UCSC Genome Browser	
Assembly	hg38
Localization ?	Nuclear Lamina,Nucleus
Tissue/Cell specificity ?	-

基因间区，核定位，特异性低



在30个物种中高度保守



较高的表达能力

Transcript ID	SmProt ID
HSALNT0276950	SPROHSA100072
HSALNT0276950	SPROHSA137538
HSALNT0276950	SPROHSA137539
HSALNT0276950	SPROHSA137540

编码94个小蛋白

与12个蛋白互作

其它功能特性:

在器官发育和细胞分化过程持续高表达，无特异性；食管鳞状细胞癌早期表达显著下调；表达呈昼夜节律特征；癌症中甲基化丰度无变化等

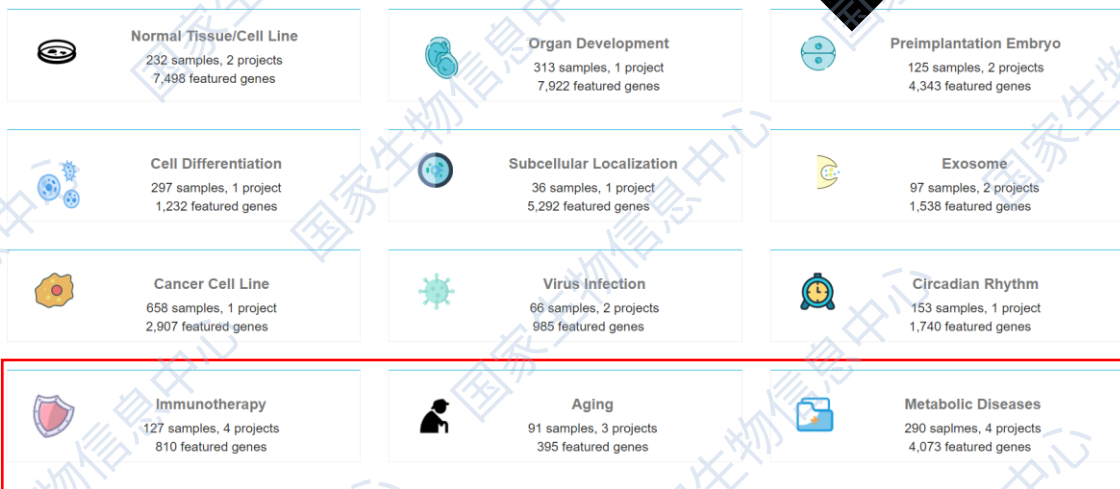
Gene ID ?	Symbol ?	Conservation (age) ?	Expression Capacity ?	Featured Expression ?	Featured Methylation ?	Variation (disease/trait-association) ?	Interaction ?	Small Protein ?
HSALNG0134930	TUG1	12. Eutheria	HC	✓	✗	✓	✓	✓

2024年更新内容：新增衰老、代谢疾病、肿瘤免疫治疗场景

LncExpDB

Expression Database of Human Long non-coding RNAs

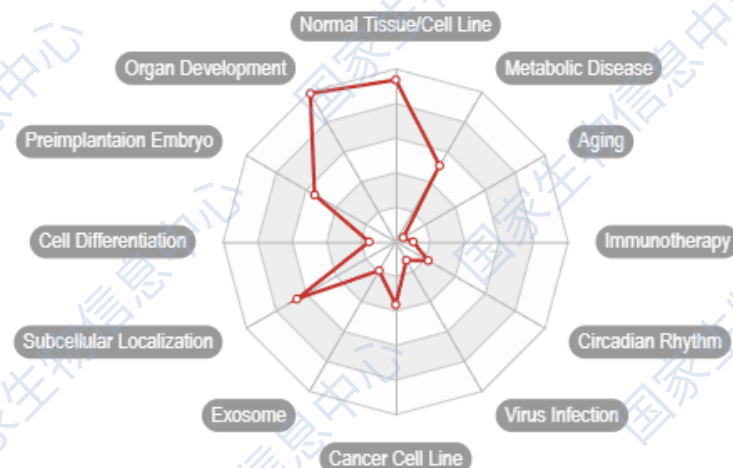
Home Genes Featured Genes Contexts Expression Capacity Interactions Downloads Statistics Help



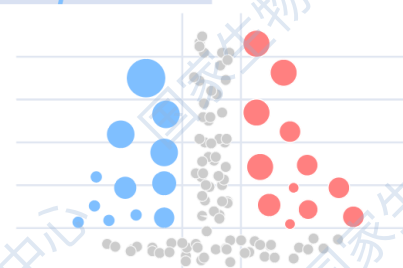
New Update

12个生物学场景，涉及在 361 个生物学条件下的
2,485 个样本，筛选出26,565个特征基因，并提供了
28,443,865对lncRNA-mRNA互作关系

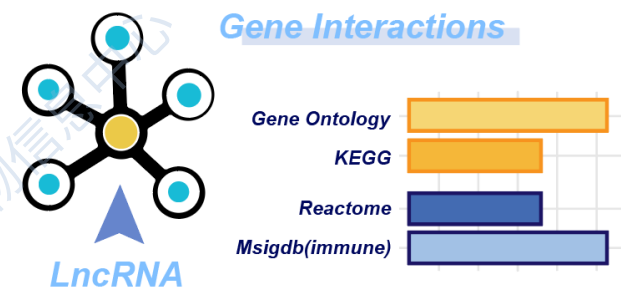
Overview of Featured Genes



Gene Expression



Gene Interactions



基因表达实例 - PRDM16-DT

Up / Down:
Up-regulated in old groups

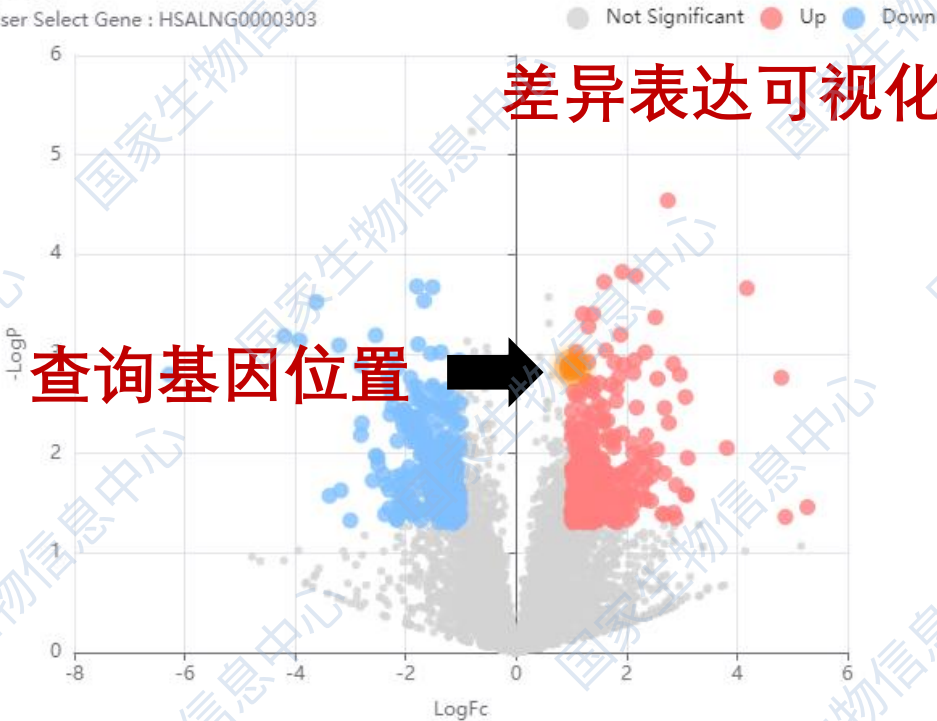
Organ / Tissue:
Brain

Gene Symbol	Brain		
	P-value	LogFC	Up / Down
OLMALINC	0.022	-1.135	Down
PACRG-AS1	0.039	-1.387	Down
PACRG-AS3	0.026	-1.244	Down
PICSAR	1.000	0.000	Not Significant
PPP1R26-AS1	0.018	1.240	Up
PRAL	1.000	0.000	Not Significant
PRDM16-DT	0.001	1.025	Up
PRKCQ-AS1,LINC02648	0.006	-1.435	Down
PWRN4,PWRN1,PWRN3	0.655	0.110	Not Significant
RNF157-AS1	0.013	1.380	Up

Gene Symbol	Brain	
	Old	Young
PRDM16-DT	H	M

Brain

User Select Gene : HSALNG0000303



PRDM16-DT is a novel lncRNA that regulates astrocyte function in Alzheimer's disease

Sophie Schröder¹, Ulrike Fuchs¹, Verena Gisa¹, Tonaŧiuh Pena^{1 2}, Dennis M Krüger^{1 2}, Nina Hempel¹, Susanne Burkhardt¹, Gabriela Salinas³, Anna-Lena Schütz⁴, Ivana Delalle⁵, Farahnaz Sananbenesi⁴, Andre Fischer^{6 7 8}

Affiliations + expand

PMID: 39207536 PMCID: PMC11362476 DOI: 10.1007/s00401-024-02787-x

Acta Neuropathol If:9.3 Q1

互作网络与功能预测- PRDM16-DT

Interactions

Search By: LncRNA Symbol

PRDM16-DT

Search By: All Protein-Codings

eg. ENSG00000152078.10

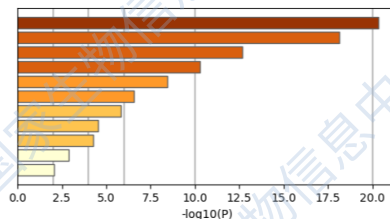
Download

LncRNA ID	LncRNA Symbol	LncRNA Location	PCG ID	PCG Symbol	PCG Location	Distance	Contexts	Breadth
HSALNG0000303	PRDM16-DT	chr1:3055439-3070489(-)	ENSG00000008083.14	JARID2	chr6:15246069-15522042(+)	Trans	OrganDevelopment	1
HSALNG0000303	PRDM16-DT	chr1:3055439-3070489(-)	ENSG00000008405.12	CRY1	chr12:106991364-107093549(-)	Trans	OrganDevelopment	1
HSALNG0000303	PRDM16-DT	chr1:3055439-3070489(-)	ENSG00000009954.11	BAZ1B	chr7:73440406-73522293(-)	Trans	OrganDevelopment	1
HSALNG0000303	PRDM16-DT	chr1:3055439-3070489(-)	ENSG00000010072.16	SPRTN	chr1:231337104-231355023(+)	228266608	OrganDevelopment	1
HSALNG0000303	PRDM16-DT	chr1:3055439-3070489(-)	ENSG00000011007.12	ELOA	chr1:23743155-23762059(+)	20672666	OrganDevelopment	1
HSALNG0000303	PRDM16-DT	chr1:3055439-3070489(-)	ENSG00000016391.11	CHDH	chr3:53812335-53846419(-)	Trans	Normal	1
HSALNG0000303	PRDM16-DT	chr1:3055439-3070489(-)	ENSG00000021776.11	AQR	chr15:34851782-34969742(-)	Trans	OrganDevelopment	1
HSALNG0000303	PRDM16-DT	chr1:3055439-3070489(-)	ENSG00000024526.17	DEPDC1	chr1:68474152-68497221(-)	65403664	OrganDevelopment	1
HSALNG0000303	PRDM16-DT	chr1:3055439-3070489(-)	ENSG00000025800.14	KPNA6	chr1:32108056-32176563(+)	29037568	OrganDevelopment	1
HSALNG0000303	PRDM16-DT	chr1:3055439-3070489(-)	ENSG00000028310.18	BRD9	chr5:850291-892801(-)	Trans	OrganDevelopment	1

Showing 11 to 20 of 1622 rows

1 2 3 4 5 ... 163

功能富集



GO:0008152: metabolic process
GO:0051179: localization
GO:0032501: multicellular organismal process
GO:0032502: developmental process
GO:0042592: homeostatic process
GO:0009987: cellular process
GO:0050896: response to stimulus
GO:0065007: biological regulation
GO:0098754: detoxification
GO:0048519: negative regulation of biological process
GO:0044119: process involved in interspecies interaction between organisms

LncRNA mRNA

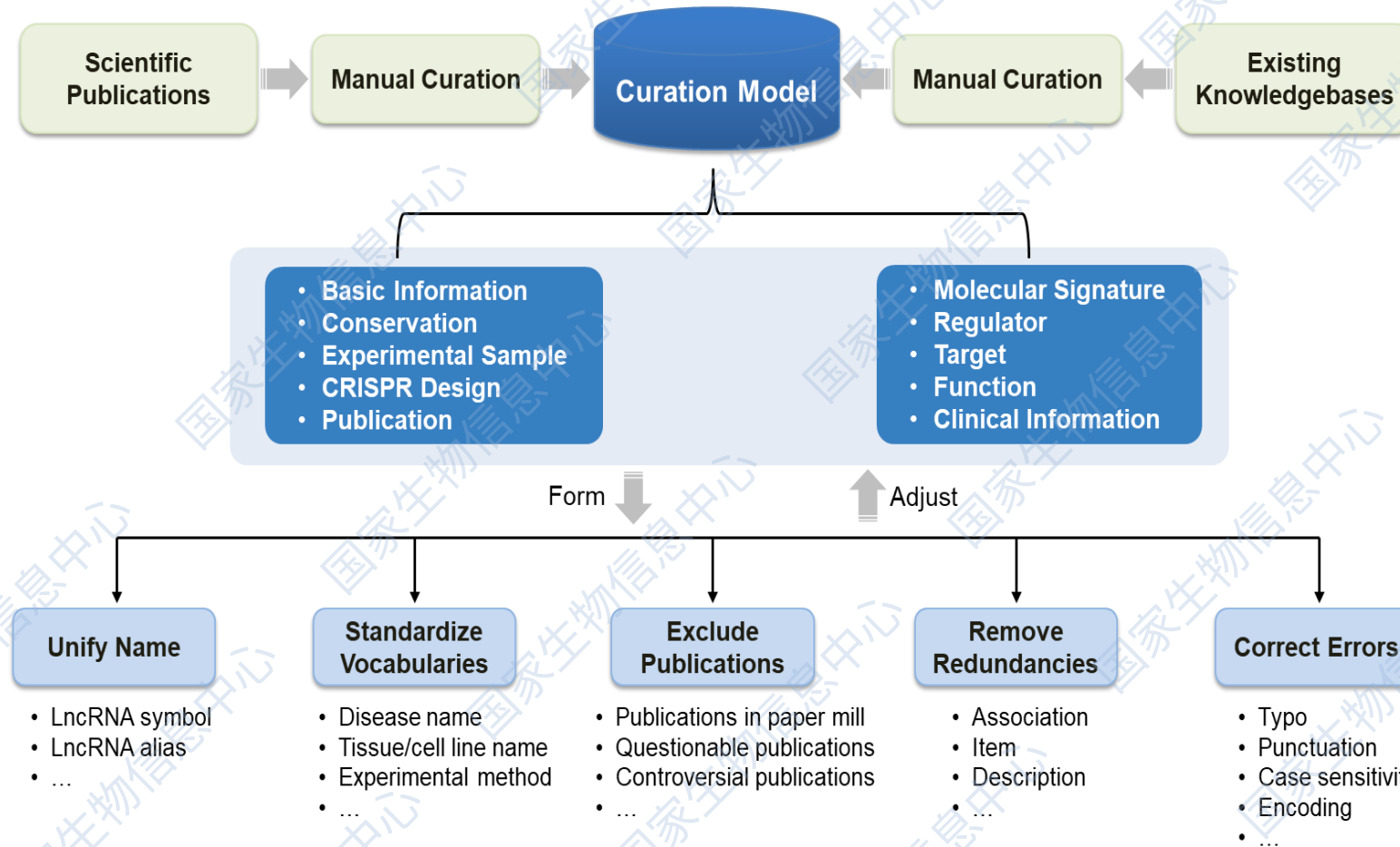
网络可视化

Different Contexts

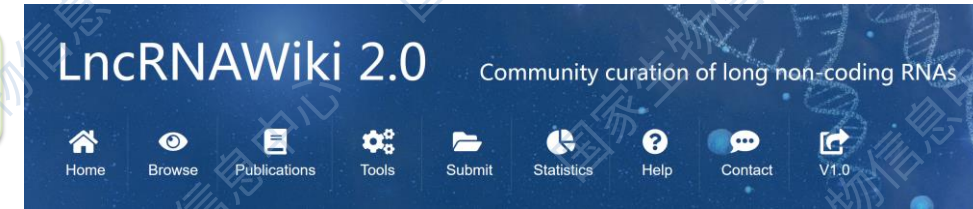


人类lncRNA知识审编与共享 (LncRNAWiki)

LncRNAWiki: lncRNA知识审编模型与平台



人类lncRNA审编模型



Data Statistics

2,609 LncRNAs	111,216 Associations	8,238 Publications
693 Diseases	336 Functions	136 Drugs
1,507 Targets	123 Regulators	587 SgRNAs
1,789 Signatures	3,200 Samples	11 Contexts

A curated knowledgebase of functionally annotated lncRNAs

Liu et al. *Nucleic Acids Res* 2022

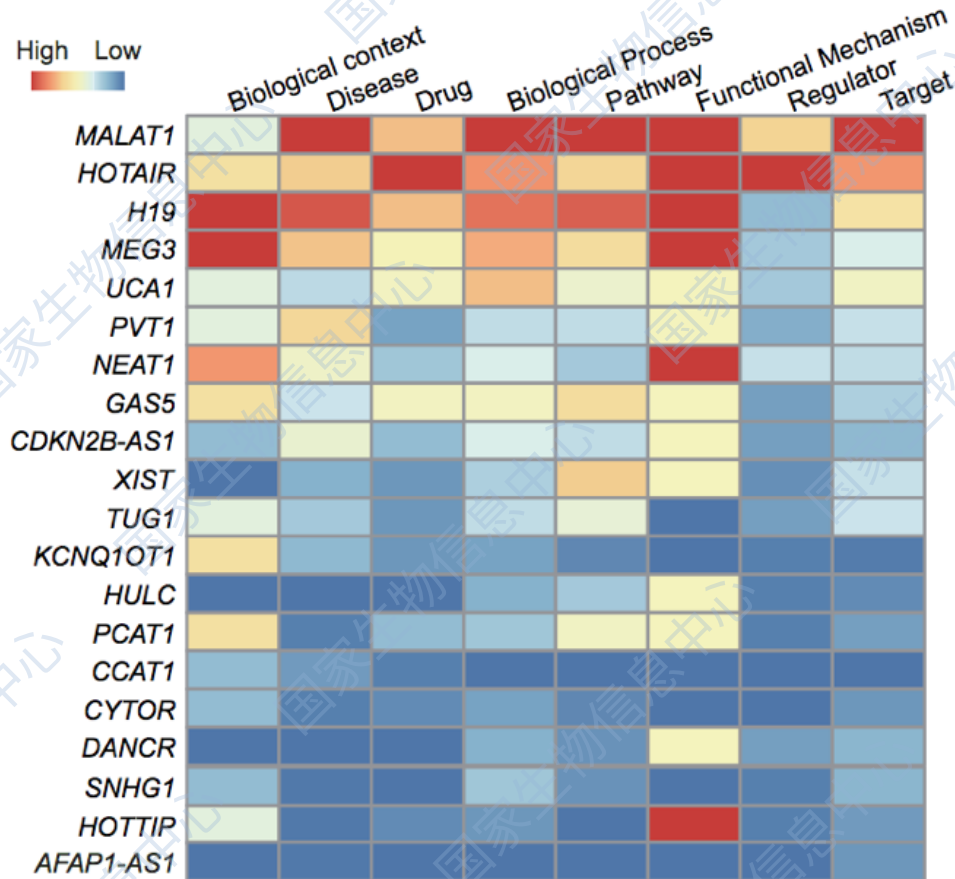
LncRNA研究成果递交与共享

Publications PMID: * e.g., 123456 Check Availability	Clinical Information Drug: ? e.g., abacavir Clinical Detail: ? e.g., metastasis	Target Target Type: ? e.g., miRNA Target: ? e.g., mir155HG Target Interaction: ? e.g., RNA-RNA Target Effect: ? e.g., Promote Experimental Method: ? e.g., RIC-seq + Add new target
Basic Information Symbol: ? * e.g., XIST Synonyms: ? e.g., X (inactive)-specific transcript Gene Locus: ? e.g., Xq13.2 Gene ID: ? e.g., ENSG00000221 Transcript ID: ? e.g., ENST00000425 LncRNA Sequence: ? Paste lncRNA sequence in the box	Molecular Signature Genome Variation: ? e.g., SNP Variation Detail: ? e.g., rs2395029 Epigenetic Modification: ? e.g., Histone Modification Modification Detail: ? e.g., H3K27me3, cg17092624 or cg17092624 Expression: ? e.g., Expression Expression Detail: ? e.g., Up-regulated	Biological Function Molecular Function: ? e.g., Protein Binding Biological Process: ? e.g., axon development Pathway: ? e.g., Signaling by WNT Functional Mechanism: ? e.g., Imprinting Tag: ? e.g., gender
Conservation Conservation: ? NA Species: ? e.g., Mus musculus Ortholog: ? e.g., Xist	Regulator of LncRNA Regulator Type: ? e.g., TF Regulator: ? e.g., P53 Regulator Interaction: ? e.g., Protein-DNA Regulator Effect: ? e.g., Promote Experimental Method: ? e.g., ChIP-seq + Add new regulator	CRISPR Design sgRNA Sequence: ? Paste sgRNA sequence in the box sgRNA Position: ? e.g., 65497676-654 sgRNA Pami: ? e.g., AGG sgRNA Type: ? e.g., CRISPRa
Sample Biological Context: ? e.g., Disease Context Detail: ? e.g., Glioma Tissue/Cell Line: ? e.g., Brain	Save	

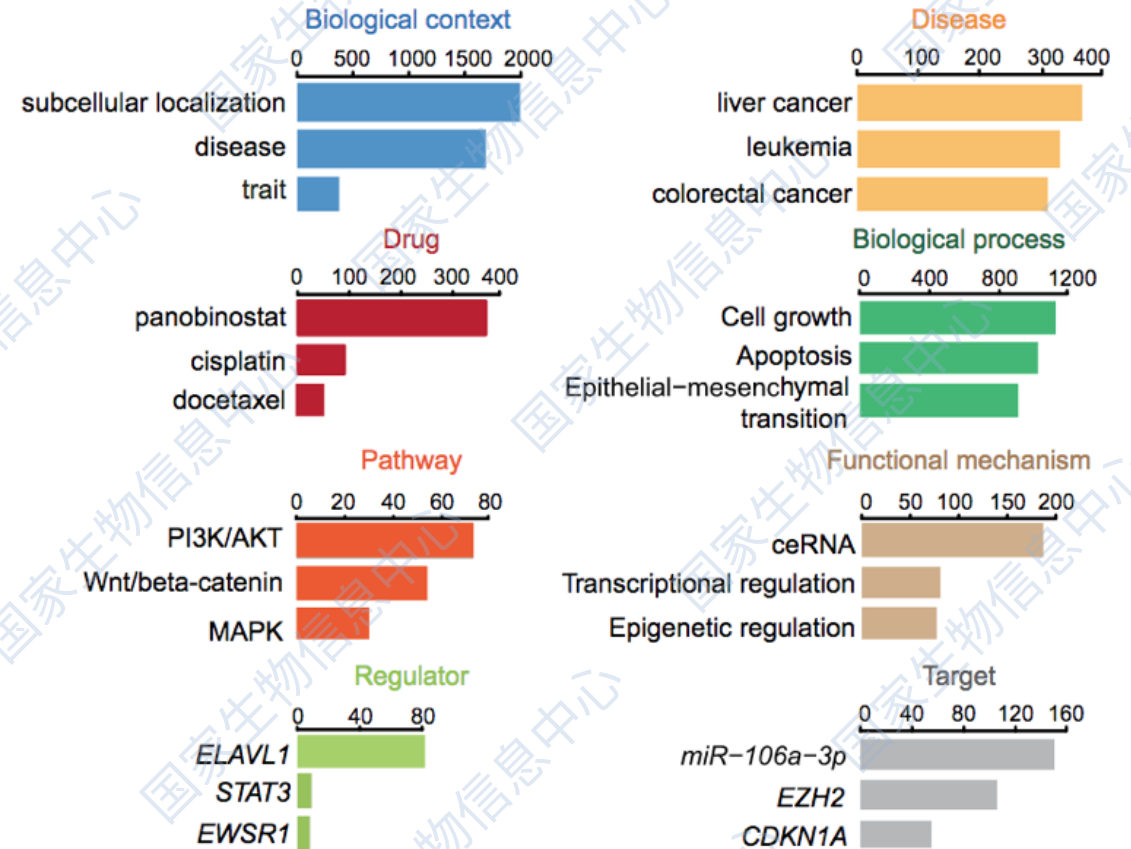
10个模块, 40个条目

LncRNA知识条目信息统计

Top 20 extensively studied lncRNAs by publication count



Top 3 frequently surveyed items by lncRNA count

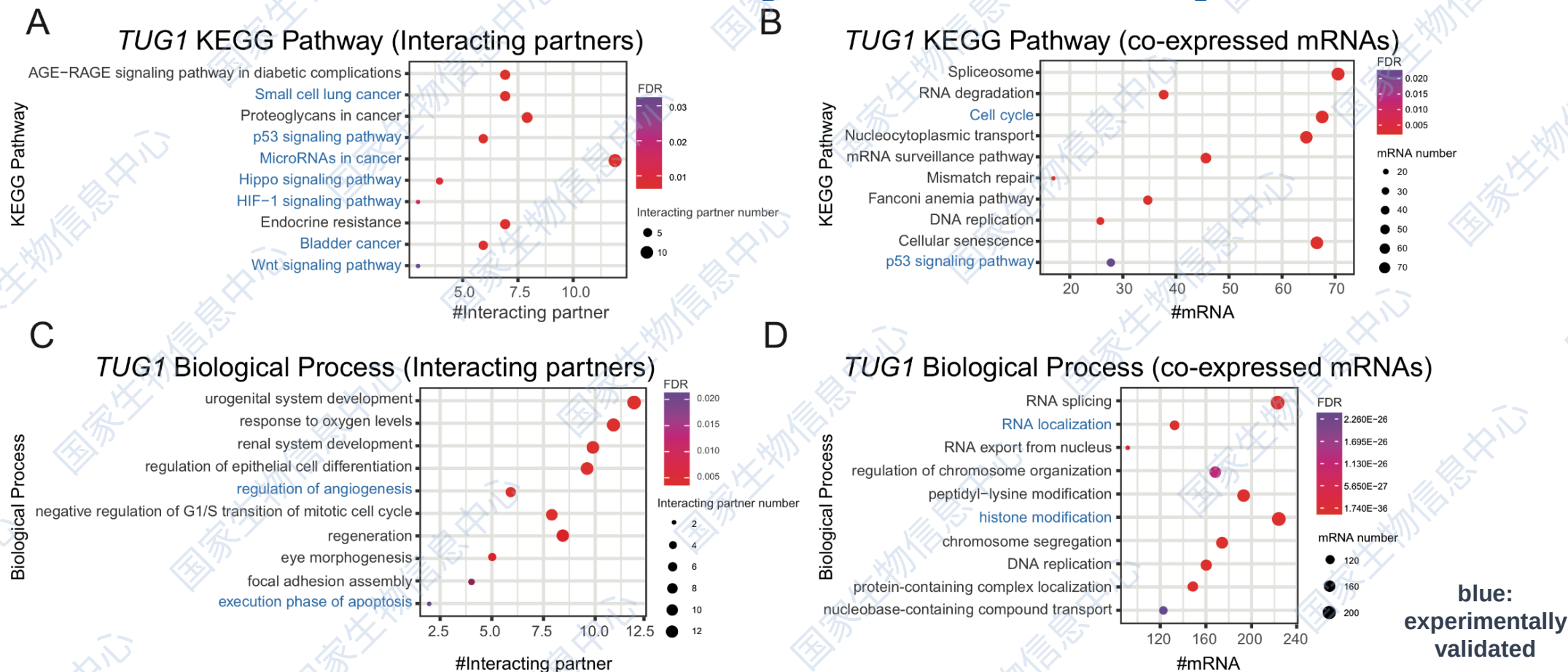


2,674 (35%) publications focus on the top 20 lncRNAs

<https://ngdc.cncb.ac.cn/lncrnawiki/statistics>

Liu et al. Nucleic Acids Res 2022

功能预测工具 (*TUG1*为例)



Functional prediction based on curated interacting partners (A and C) and co-expressed mRNAs (B and D) indicates that *TUG1* may be involved in many pathways and biological processes.

<https://ngdc.cncb.ac.cn/lncrnawiki/tool>

Liu et al. Nucleic Acids Res 2022

2024年开发了智能审编工具LncBot（举例：*LINC00857*）

 User email:

 submit a PDF

Drag and drop file here
Limit 2GB per file • PDF

 sample files

Note - Please press Restart first if
uploading pdf files that have already been
chat...

Firefox is recommended

 Upload PDF first...

Select one meta-info to ask ...

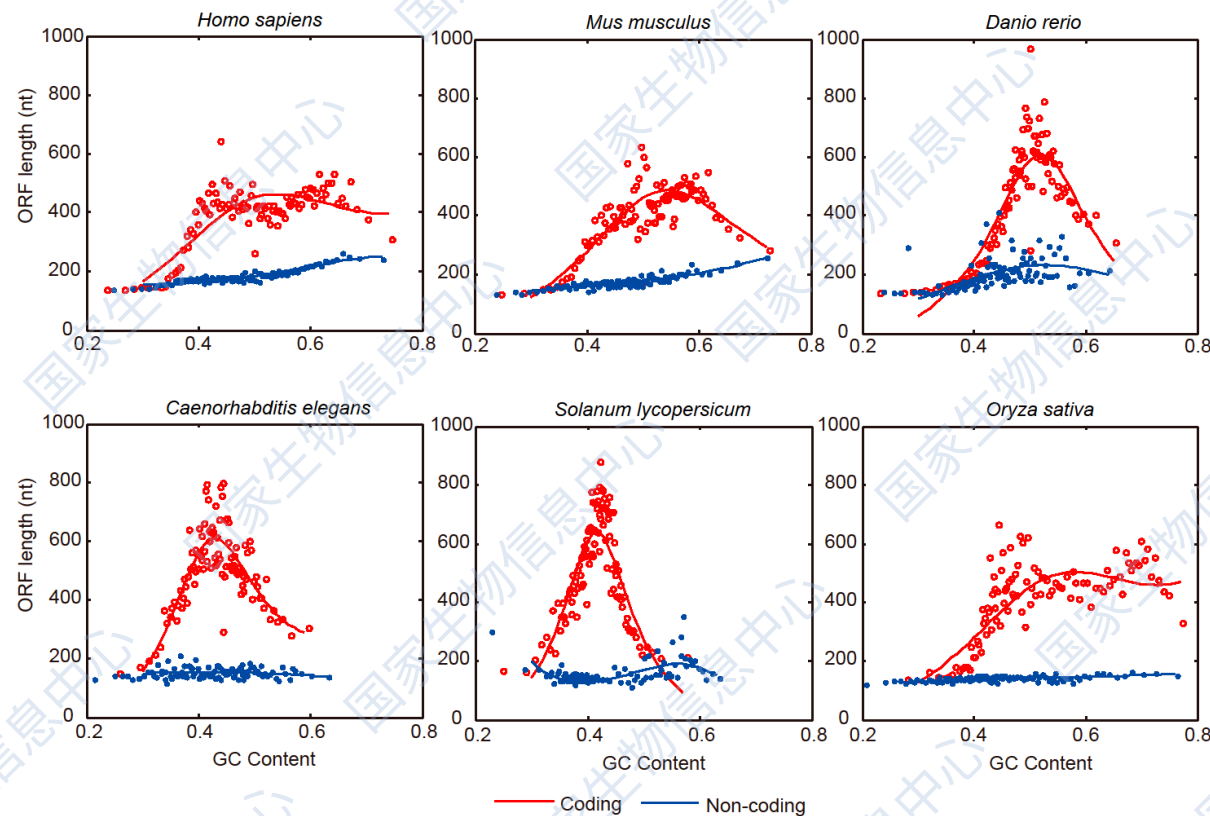
 Q&A based on PDFs...

 You are an efficient LncRNA literature reader. Based on the user's query lijzn@163.com, extract and summarize the most relevant and precise answers from the literature based on the document of None, focusing on the main points and avoiding unnecessary information.

<https://ngdc.cncb.ac.cn/lncrnawiki/lncbot/>

多物种lncRNA预测 (LGC算法与在线服务)

LGC: 基于特征关系预测编码潜能



期望的ORF 长度随GC含量增加而增加

$$f = a_0 + a_1 P_{GC} + a_2 P_{GC}^2 + a_3 P_{GC}^3$$

Oliver J.L. and Marín A. **A relationship between GC content and coding-sequence length.** *J Mol Evol* 1996

$$L = \log_2 \frac{p_c}{p_{nc}} = \log_2 \frac{(1-f_c)^{n-1} f_c}{(1-f_{nc})^{n-1} f_{nc}}$$

$L > 0$: Coding

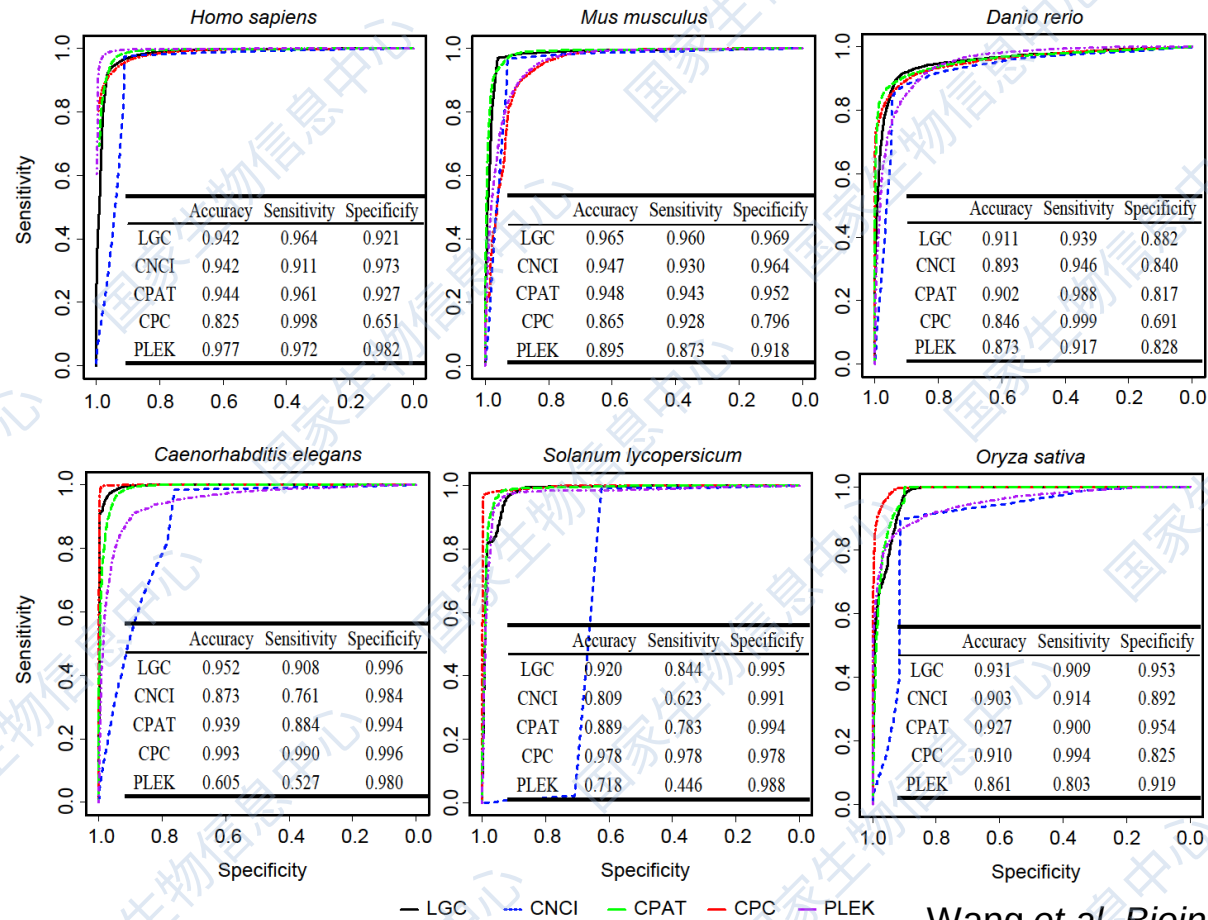
$L < 0$: Non-coding

Wang et al. *Bioinformatics* 2019

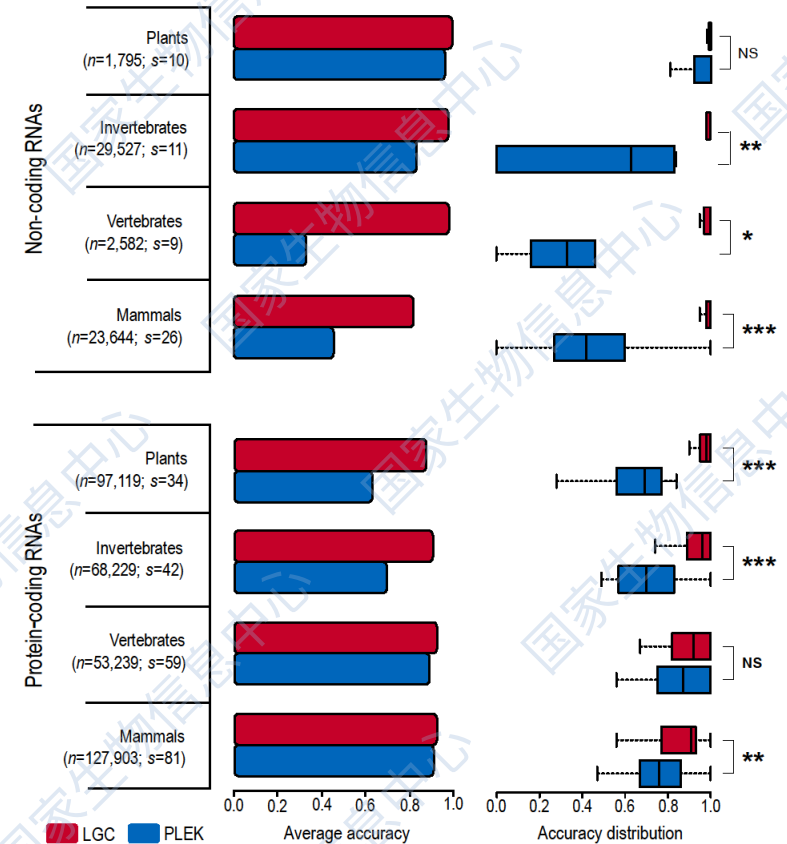
基于Oliver公式，利用“ORF长度和GC含量之间的关系在lncRNAs和protein-coding RNAs之间呈现明显差异，在多个物种中均有较一致的表现”的特征，构建最大似然算法模型。

LGC和其它算法的比较

Performances of **LGC**, CNCI, CPAT, CPC and PLEK



Accuracy of LGC and PLEK



Wang et al. Bioinformatics 2019

多物种（动物+植物） lncRNA在线鉴定平台

LGC



Characterization and identification of long non-coding RNAs based on feature relationship

☒ FASTA Format

☐ FASTA File: 未选择任何文件

☐ FASTA Data:

FASTA

e.g., *Example Sequence*

☐ GTF Format

☒ GTF File: 未选择任何文件

☐ GTF Data:

GTF

e.g., *Example Sequence*

Select Species assembly

- ☐ Human (hg38, GRCh38)
- ☐ Human (hg19, GRCh37)
- ☐ Mouse (GRCm38/mm10)
- ☐ Mouse (NCBI Build 37/mm9)
- ☐ Fly (dm3, BDGP Release 5)
- ☐ Zebrafish (Zv9/danRer7)

☐ BED Format

☐ BED File: 未选择任何文件

☐ BED Data:

BED

e.g., *Example Sequence*

Select Species assembly

- ☐ Human (hg38, GRCh38)
- ☐ Human (hg19, GRCh37)
- ☐ Mouse (GRCm38/mm10)
- ☐ Mouse (NCBI Build 37/mm9)

Web server: <https://bigd.big.ac.cn/lgc/calculator>

Wang et al., Bioinformatics, 2019

其他在线分析工具

其它lncRNA在线分析工具

BLAST

Fasta File: 未选择文件。

Fasta Sequence:

```
>ENST00000469225.1_1
GGAAAAGACCCAGTACAGTGTAGCCGTGATG
TGGTCATTTGTCCTGATGCCAGCCTGAAGAT
GCAAAAAAAGAGGGACCATATGATGTGGTGG
TTCTACCAGGAGGTAATCTGGGCGCACAGAA
```

e.g., [Example](#) [Clear](#)

Classification

☐ GTF File: 未选择文件。

☒ GTF List:

```
chr1 LncBook gene 16858
17751 . . . gene_id
"HSLNG0000003"; gene_type "lncRNA";
chr1 LncBook transcript 16858
17751 . . . gene_id
```

e.g., [Example](#) [Clear](#)

Select Species Assembly:

☒ Human (hg19, GRCh37)

☐ Human (hg38, GRCh38)

ID Conversion

☐ Gene ID File: 未选择文件。

☒ Gene ID List:

```
HSLNG0134930
NONHSAG008675.3
CHS.7801
HSLNT0189447
HSLNT0368398
```

e.g., [Example](#) [Clear](#)

Query ID	Subject ID	Query Start	Query End	Subject Start	Subject End	Identity	Evalue	Score (bit)
>ENST00000469225.1_1	HSALNT0001239	1	711	1	711	100.000	0.0	1283
>ENST00000469225.1_1	HSALNT0188749	1	709	696	1	85.675	0.0	865
>ENST00000469225.1_1	HSALNT0152586	1	483	119	570	80.372	5.34e-127	457
>ENST00000469225.1_1	HSALNT0187711	434	711	2658	2380	68.198	8.51e-17	91.5
>ENST00000469225.1_1	HSALNT0187710	434	711	3021	2743	68.198	8.51e-17	91.5

序列比对检索

Transcript ID	Classification
HSALNT0000005	Intergenic

基因组位置分类

Gene ID	Symbol	NONCODE	GENCODE	lncipedia	BiGTranscript	CHES	RefLnc	miTranscript	FANTOM	HGNC	Entrez	UCSC	RefSeq	lncRNAdb
HSALNG006	PVT1	NONHSAG05	ENSG000002	PVT1	BiG-gene-ch	CHS.54468	XR_928641.1	-	ENSG000002	HGNC:9709	5820	uc003ysl.4	NR_003367	pvt1
HSALNG008	H19	NONHSAG0C	ENSG000001	H19	-	CHS.7801	-	-	ENSG000001	HGNC:4713	283120	uc057xvv.1	NR_002196	H19
HSALNG008	NEAT1	NONHSAG0C	ENSG000002	NEAT1	BiG-gene-ch	CHS.9123	NR_028272.1	-	ENSG000002	HGNC:30815	283131	uc010rog.3	NR_028272	NEAT1
HSALNG008	MALAT1	NONHSAG0C	ENSG000002	MALAT1	BiG-gene-ch	CHS.9136	-	G014182	ENSG000002	HGNC:29665	378938	uc010roh.5	NR_002819	Malat1
HSALNG013	TUG1	NONHSAG05	ENSG000002	TUG1	-	CHS.35608	NR_002323.1	-	ENSG000002	-	-	-	-	-
HSALNT0189	-	-	ENST000004	HOTAIR.3	-	-	-	-	-	-	-	-	-	-
HSALNT0368	-	-	ENST000004	-	-	CHS.11761.4	-	-	-	-	-	-	-	-

ID转换

ID 转换举例

[Home](#) > [Tools](#) > [ID Conversion](#)

Upload a file containing gene IDs or enter gene IDs to obtain accession ID across different databases, including LncBook, NONCODE, GENCODE, LNCipedia, BIGtranscriptome, CHESS, RefLnc, MiTranscriptome, FANTOM, HGNC, Entrez, UCSC, RefSeq and lncRNAdb.

ID Conversion

☐ Gene ID
File:

选择文件 gene_ids.txt

☒ Gene ID
List:

HSALNG00000002
HSALNG00000003
HSALNG00000004
HSALNG00000005

e.g., [Example](#) [Clear](#)

Run

Reset

CSV

Gene ID	Symbol	NONCODE	GENCODE	LNCipedia
HSALNG00000002	-	NONHSAG000002.2	-	-
HSALNG00000003	-	NONHSAG000002.2	-	-
HSALNG00000004	-	NONHSAG000002.2	-	-
HSALNG00000005	MIR1302-2HG	NONHSAG000003.2	ENSG00000243485.5	lnc-OR4F5-1
HSALNG00000006	FAM138A	NONHSAG000004.2	ENSG00000237613.2	FAM138A
HSALNG00000007	-	NONHSAG058641.2	-	-
HSALNG00000009	-	-	-	-
HSALNG00000010	-	NONHSAG057396.1	-	LINC00115
HSALNG0141639	-	-	-	-
HSALNG0142075	-	-	ENSG00000238009.6	-

Showing 1 to 10 of 10 entries

总结：人类lncRNA数据资源系统

三个核心数据库，一个关键算法，五个分析工具，主要应用于：

- 人类lncRNA高质量参考数据集
- 多组学特征比较分析
- 多场景表达特征解析
- 研究成果在线提交共享
- 文献知识一键获取
- 在线功能预测
- 跨物种鉴定
- ID转换、序列检索、分类

LncBook
Nucleic Acids Res 2023, 2019



LncExpDB
Nucleic Acids Res 2021



LncRNAWiki
Nucleic Acids Res 2022, 2015



LGC
Bioinformatics 2019



课题参与人员



李昭



刘琳



冯昶瑞



周新羽



郑兴



刘畅



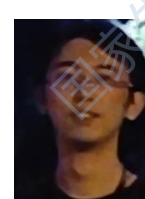
齐跃



覃煜鑫



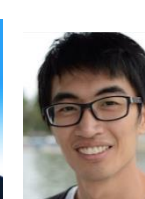
荆蔚



罗思成



曹佳宝



王光煜



QQ群：253662314
lncwiki@big.ac.cn

谢 谢

QQ群：253662314
lncwiki@big.ac.cn