

機能性RNAデータベースの紹介

産業技術総合研究所
生命情報工学研究センター

光山統泰

機能性RNAプロジェクト

- 期間: 2005年10月～2009年3月
- ① 機能性RNAの探索・解析のためのバイオインフォマティクス技術の開発
- ② 機能性RNA解析のための支援技術・ツールの開発
- ③ 機能性RNAの機能の解明



京都大学



生命情報工学研究センター



三菱総合研究所



協和発酵



バイオインフォマティクスの実施項目

ゲノム配列からの機能性RNAの網羅的予測

必要

必要

機能性RNAに特化したバイ
オインフォマティクス技術の
開発

統合

データベース開発

網羅的予測には、アルゴリズムの開発とデータベースの開発が必要不可欠

プロジェクト成果のデータベースや情報
ツールを公開するためのポータルサイト

- ナビゲーション
- ホーム
 - News
 - UCSC
GenomeBrowser
for Functional RNA
 - Project Specific
Custom Tracks
 - fRNAdb: A
comprehensive
non-coding RNA
sequence database
 - CentroidFold
 - Idiographica -
Create Your Own
Idiograms
On-demand!
 - Bioinformatics
Tools for RNAs
 - RNA Bibliography
 - The Contents of
Genomes
 - Rfam Secondary
Structure Gallery
 - ncRNA
KnowledgeBase
 - ジャーナルクラブ | 論文

Main Contents

Databases



Ver. 3.0

fRNAdb is a database for comprehensive non-coding RNA sequences.



Ver. 3.0

UCSC Genome Browser mirror with large inclusion of functional RNA related custom tracks and several enhancements for RNA secondary structure support.

[Project Specific Custom Tracks](#)

Table of custom tracks added by the functional RNA project.

[The Contents of Genomes](#)

The Contents of Genomes presents major components of a genome at a glance.

[RNA Bibliography](#)

RNA Bibliography hosts a collection of journal articles related to RNAs.

Bioinformatics tools

[software.ncrna.org](#)

Our bioinformatics tools for RNA sequence analysis with web servers and source codes.

[CentroidFold](#)

CentroidFold predicts an RNA secondary structure from an RNA sequence.

[Idiographica](#)

Idiographica is a web server for you to create your own idiograms on-demand.

Database

機能性RNAデータベース

Bioinformatics Tools

RNA二次構造予測

RNA配列のアラインメント

機能性RNAデータベース

①

**RNAdb
functional**

配列DB



[Top](#) [Statistics](#) [Blast](#) [Download](#) [ncrna.org](#) [Help](#)

**RNAdb
functional**

A comprehensive non-coding RNA sequence database ver. 3.0

Total 407,799 entries






[Statistics](#) [Blast](#) [Download](#) [Help](#)

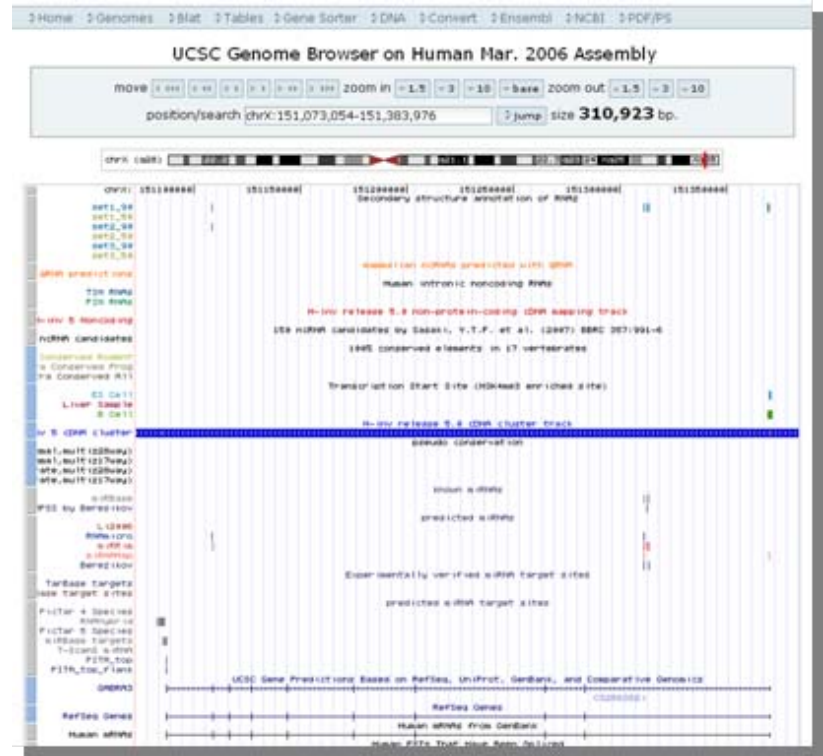
Please input some query keywords for retrieving RNAs via simple text search e.g. "miRNA" or "snoRNA".
 Try add disease/tissue name to make your search more specific e.g. "miRNA oncogene".
 Or can be more specific e.g. "miRNA oncogene human".

The icons used in this page are a part of Tango Desktop Project created by Tango Desktop Project which are available under CC BY-SA License.
 This site uses Yahoo! User Interface Library (YUI) which is available under BSD License.

②

**UCSC GenomeBrowser
for Functional RNA**

ゲノムブラウザ



UCSC Genome Browser on Human Mar. 2006 Assembly

[Home](#) [Genomes](#) [Blat](#) [Tables](#) [Gene Sorter](#) [DNA](#) [Convert](#) [Ensembl](#) [NCBI](#) [PDF/PS](#)

position/search chrX:151,073,054-151,383,976 size 310,923 bp.

chrX:151,073,054-151,383,976

chrX: 151288884 151288888 151288892 151288896 151288900 151288904
 Secondary structure annotation of hsaR1
 hsaR1 release 5.8 non-protein-coding tRNA mapping track
 hsaR1 release 5.8 tRNA cluster track
 UCSC Gene Predictions Based on RefSeq, UniProt, GenBank, and Comparative Genomics
 RefSeq Genes
 Human tRNA
 Human tRNA

fRNAdb: 配列データベース

[Top](#) [Statistics](#) [Blast](#) [Download](#) [ncrna.org](#)

[Help](#)



A comprehensive non-coding RNA sequence database ver. 3.0

Total: 407,759 entries



[Statistics](#)



[Blast](#)



[Download](#)



[Help](#)

Please input some query keywords for retrieving RNAs via simple text search e.g. "miRNA" or "snoRNA".

Try add disease/tissue name to make your search more specific e.g. "miRNA oncogene".

Or can be more specific e.g. "miRNA oncogene human".

[submit](#)

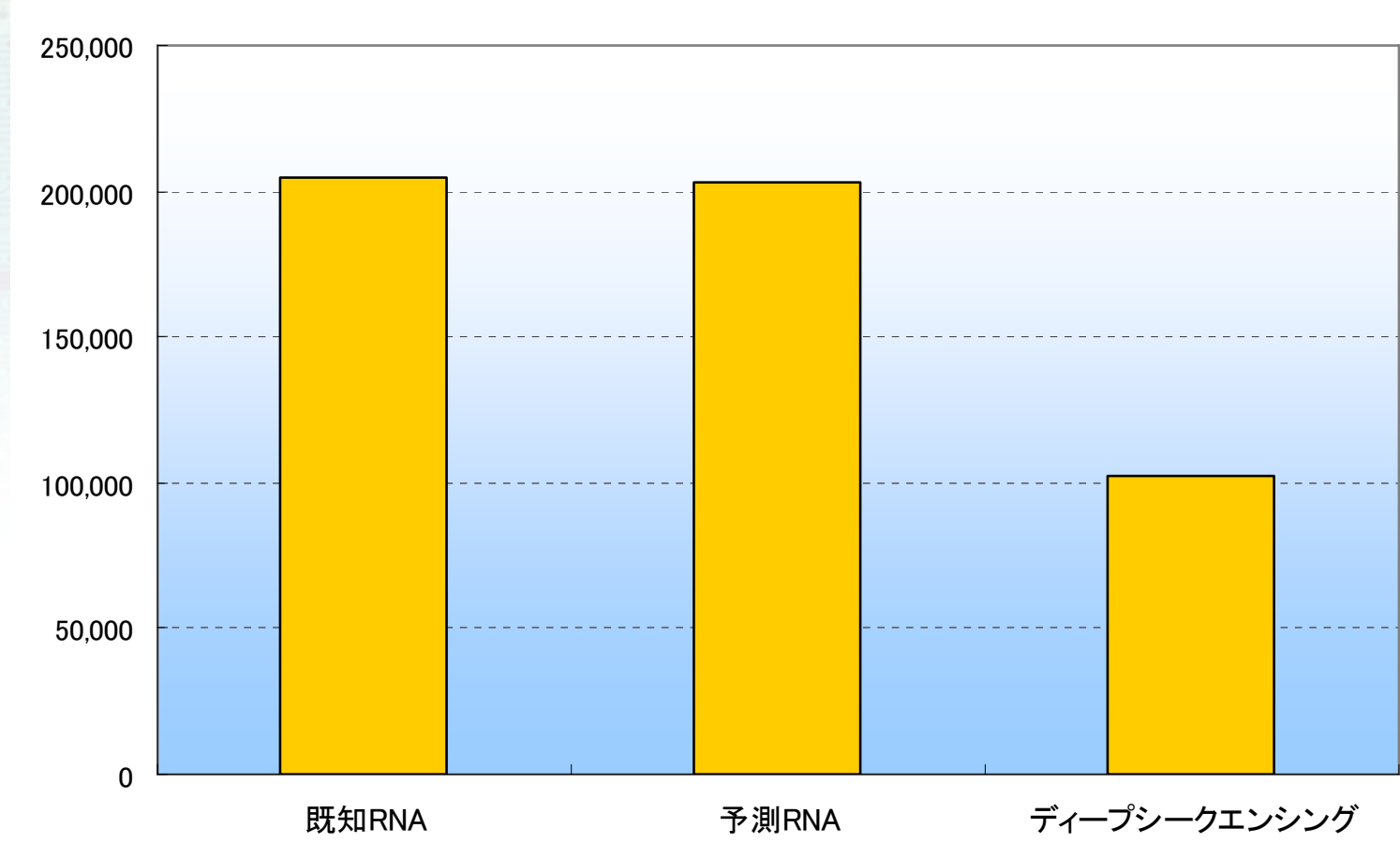


The icons used in this page are a part of Tango Desktop Project [created](#) by Tango Desktop Project which are available under [CC BY-SA License](#).

This site uses Yahoo! User Interface Library (YUI) [which](#) is available under [BSD License](#).

配列DBの内容

全 509,795 配列



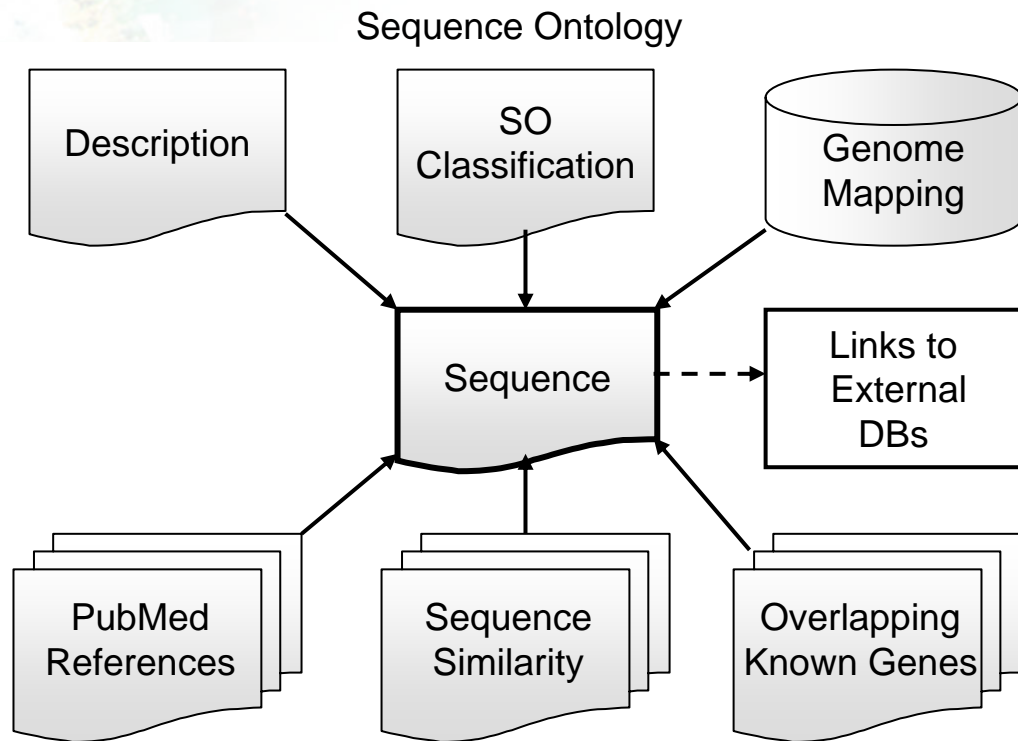
配列情報ソース

Database	IDs
<u>RNAdb</u>	<u>289,336</u>
<u>Rfam</u>	<u>108,107</u>
<u>Gene Expression Omnibus</u>	<u>102,130</u>
<u>miRBase</u>	<u>5,912</u>
<u>NONCODE</u>	<u>5,516</u>
<u>H-invitational</u>	<u>2,169</u>
<u>snoRNA-LBME-db</u>	<u>359</u>
<u>European ribosomal RNA database</u>	<u>44</u>

プロジェクト独自の成果

Category	Sequences
<p><u>150 mRNA-like ncRNA Candidates by Hirose Lab.</u> Sasaki YT, Sano M, Ideue T, Kin T, Asai K, Hirose T Identification and characterization of human non-coding RNAs with tissue-specific expression. <i>Biochem Biophys Res Commun</i>, 357(4):991-6, 2007. 産総研 バイオメディシナル情報研究センター 廣瀬研究室</p>	150
<p><u>Drosophila AGO2-associated small RNAs by Siomi Lab.</u> Kawamura Y, Saito K, Kin T, Ono Y, Asai K, Sunohara T, Okada TN, Siomi MC, Siomi H Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. <i>Nature</i>, 453(7196):793-7, 2008. 慶応義塾大学 医学部 塩見研究室</p>	472

配列関連情報



多様な情報が配列に関連付いていることで、様々なキーワード検索に対応している

配列相同性検索

[Top](#)
[Statistics](#)
[Blast](#)
[Download](#)
[ncrna.org](#)
[Help](#)

fRNAdb::Blast

Enter one or more query sequence(s)

The query sequence is filtered for low complexity regions by default.

Filter Low complexity


Search Category Search Length

execution	
Cost to open a gap [Integer] "auto": 5 for a query \geq 50 nt. Otherwise gapped alignment is disabled.	<input type="text" value="auto"/>
Cost to extend a gap [Integer] "auto": 2 for a query \geq 50 nt. Otherwise gapped alignment is disabled.	<input type="text" value="auto"/>
Penalty for a mismatch in the blast portion of run [Integer] "auto": -3	<input type="text" value="auto"/>
Reward for a match in the blast portion of run [Integer] "auto": 1	<input type="text" value="auto"/>
Expectation value [Real] "auto": 10.0 for a query \geq 50 bases, 0.001 for a query \geq 25 nt and $<$ 50 nt, and 0.0001 for q query $<$ 25 nt.	<input type="text" value="auto"/>
Word size [Integer] "auto": 11 for a query \geq 50 nt. Otherwise 7.	<input type="text" value="auto"/>
Number of one-line descriptions [Integer] "auto": 20	<input type="text" value="auto"/>
Number of alignments to show [Integer] "auto": 20	<input type="text" value="auto"/>

パラメータは入力配列に基づいて自動調整される(割り込み可)

ゲノムブラウザ UCSC GenomeBrowser for Functional RNA

- UCSC Genome Browserの拡張版
 - ヒト、マウス、ラット、ショウジョウバエ、線虫
- 追加トラック(ヒト)
 - 予測RNA: 8トラック
 - 既知RNA: 21トラック(種別ごとに分類)
 - ゲノム因子: 8トラック
 - miRNA関連トラック: 16トラック



UCSC GenomeBrowser
for Functional RNA

予測RNAトラック

Description	Items
二次構造保存領域予測①(RNAz)	35,984
アンチセンス転写産物	295
snoRNA予測	322
二次構造保存領域予測②(QRNA)	3,377
イントロン内のnon-coding EST	67,729

マイクロRNA関連

Description	Items
既知のmiRNA (miRBase他)	573
予測miRNA (miRRim他)	11,616
検証されたmiRNAターゲット	233
予測miRNAターゲット領域	2,312,386
miRNA発現アトラス	992

miRRim: a novel system to find conserved miRNAs with high sensitivity and specificity.
Terai G, Komori T, Asai K, Kin T
RNA 13(12):2081-90, 2007 Dec, Epub 2007 Oct

ゲノムブラウザ画面の例

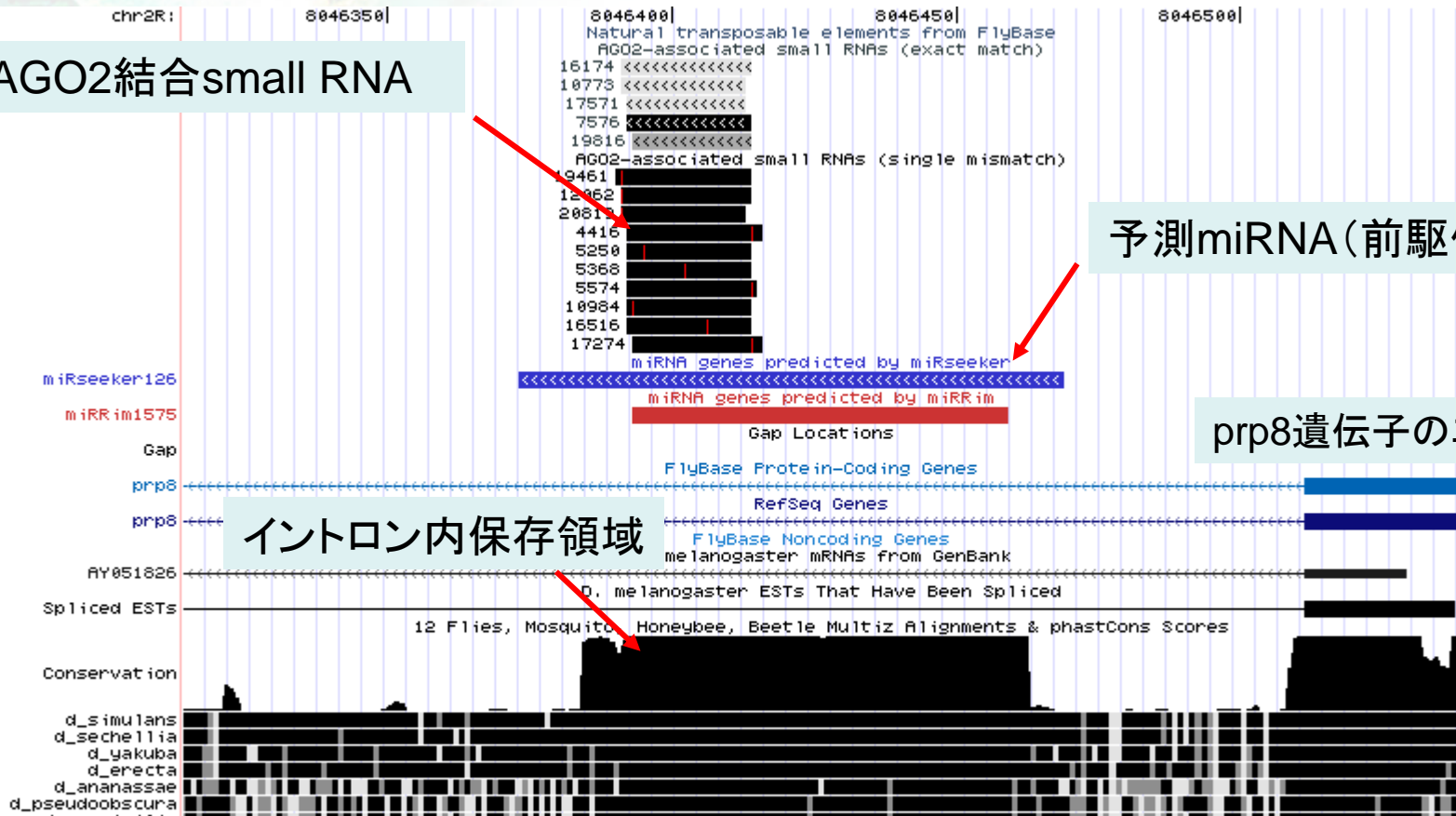
ショウジョウバエの予測miRNAが塩見研のAGO2結合small RNAによって検出された例

塩見研AGO2結合small RNA

予測miRNA(前駆体)

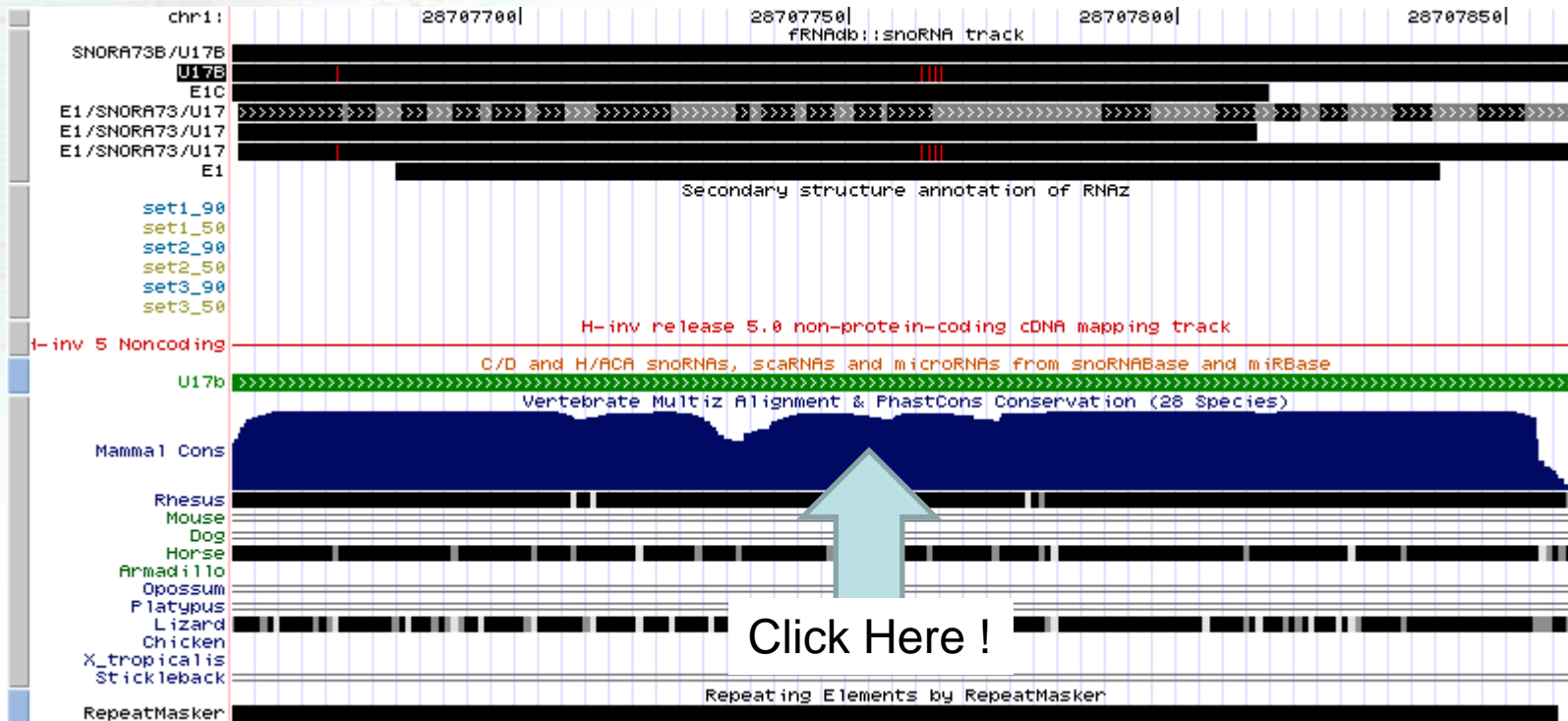
prp8遺伝子のエクソン

イントロン内保存領域



画面上で予測結果と実験によって検出された転写産物の関係が視覚化されて非常に解りやすく示されている。

Dynamic Conserved Secondary Structure Prediction at Arbitrary Genome Region



Track for Genome Comparison

Conserved Secondary Structure Prediction

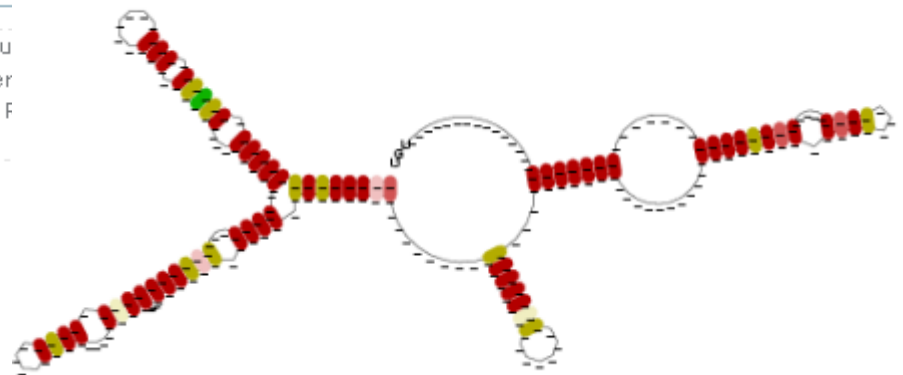
Vertebrate Multiz Alignment & PhastCons Conservation (28 Species)

RNA Secondary Structure Prediction

✓ This view presents a result of evolutionary conserved RNA secondary structure. The result is computed and presented automatically on any region shorter structure. However, this does not mean that there is certainly conserved functional RNA analysis.

Highlighting compensatory mutations in RNA multiple alignments

- Gray: Not part of annotated pair, no substitution.
- Lt. purple: Not part of annotated pair, substitution.
- Black: Compatible with annotated pair, no substitutions.
- Blue: Compatible with annotated pair, single substitution.
- Green: Compatible with annotated pair, double substitution.
- Red: Not compatible with annotated pair.



strand	minimum free energy (kcal/mol)	# of pairs, no substitutions
forward	-69.36	62
reverse	-57.72	55

Multiple alignment and RNA Secondary structure (forward strand)

%ID	offset	0	1	2	3	4	5	6
Human		tccaacgtggataaccctgg-gaggtcactctcccaggctctgtccaagtgccataggggagccttagggc						
97.6 Rhesus		tccaacgtggataaccctgg-gaggtcactctcccaggctctgtccaagtgccataggggagccttagggc						
92.1 Horse		tccaacgtggataaccctgg-gaggtcactctcccaggctctgtccaagtgccataggggagccttagggc						
77.6 Lizard		-CCAATGCGGATATCTCGCAGAGGTCCGCTCTCTCAGCGCTGTCCCGCTGGCATGTGGGAGCCTAGCC						
SS anno		((
pair symbol		abcdefghijklmnopqrs tuv vut srqpo nmlki iklm nopqrstuvwxyzA						
offset		7	8	9	0	1	2	3
Human		tct-gccccatgatgtacagtcaccttccacaacggtgaaagatgaagctgggcctc-----gtgtctgc-						
Rhesus		tct-gccccatgatgtacagtcaccttccacaacggtgaaagatgaagctgggcgtt-----gtgtctgc-						
Horse		tcc-gccccgtgatgtacagtcaccttccacaacggtggaagatgaagcggggcctt-----gagctctgc-						
Lizard		TCTGCGCCGGTCA TGTACAGTCCCCTTCCACACCCATTGGAGAGCAAGACCGGGCCTCCTCGGAGGCTGCA						

Idiogram Generation via Table Browser

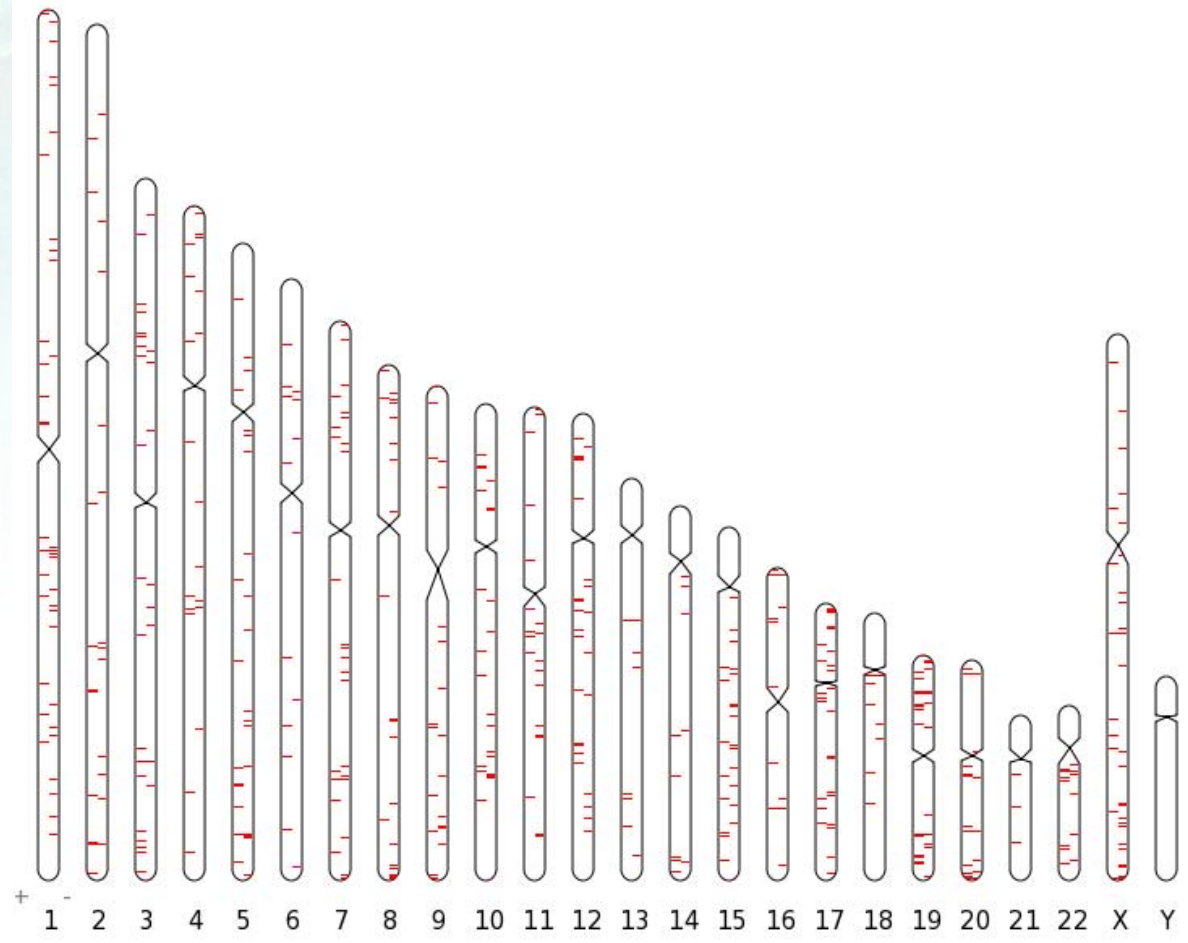
Table Browser

✔ Use this program to retrieve the data associated with a track in text format, to calculate intersections between tracks, and to retrieve DNA sequence covered by a track. For help in using this application see [Using the Table Browser](#) for a description of the controls in this form, the [User's Guide](#) for general information and sample queries, and the [OpenHelix Table Browser tutorial](#) for a narrated presentation of the software features and usage. For more complex queries, you may want to use [Galaxy](#) or our [public MySQL server](#). Refer to the [Credits](#) page for the list of contributors and usage restrictions associated with these data.

clade: Vertebrate
genome: Human
assembly: Mar. 2006
group: Genes and Gene Prediction Tracks
track: UCSC Genes
table: knownGene describe table schema
region: genome position chr1:28707657-28707861
identifiers (names/accessions):
filter:
intersection:
correlation:
output format: all fields from selected table Send output to Galaxy
output file: (leave blank to keep output in browser)
file type returned: plain text gzip compressed

To reset **all** user cart settings (including custom tracks), [click here](#).





Idiogram of known miRNAs



使用例


- ヒトとマウスに保存されているマイクロRNAの検索
- <http://www.ncrna.org/frnadb/>

Total: 401,109 entries

Statistics Blast Download Help

Please input some query keywords for retrieving RNAs via simple text search e.g. "miRNA" or "snoRNA".
Try add disease/tissue name to make your search more specific e.g. "miRNA oncogene".
Or can be more specific e.g. "miRNA oncogene human".



検索結果

Top Statistics Blast Download ncrna.org

Help

microRNA human[org] mouse[org]

submit ?

<< < 1 / 11 > >>

Display: 20

Summary

Order by

Download

Total hits: 201

1:FR000692 [acc:AJ459739](#) [MIMAT0000140](#) [miRBase v9.2](#) [\(more\)](#)

mature micro RNA (miRNA) miR-128-3p / miR-128a-3p

SO name:miRNA [\(more\)](#) Organism:*Ateles geoffroyi* [\(more\)](#) 2 citations 22nt [SEQ](#)

Mapped:human(hg17) human(hg18) rat(rn3)

2:FR004174 [acc:AJ459757](#) [MIMAT0000157](#) [miRBase v9.2](#) [\(more\)](#)

mature micro RNA (miRNA) miR-145-5p

SO name:miRNA [\(more\)](#) Organism:*Gorilla gorilla* [\(more\)](#) 2 citations 24nt [SEQ](#)

Mapped:human(hg17) human(hg18) rat(rn3)

3:FR007012 [acc:AJ550419](#) [MIMAT0000273](#) [miRBase v9.2](#) [\(more\)](#)

mature micro RNA (miRNA) miR-216-5p / miR-216a-5p

SO name:miRNA [\(more\)](#) Organism:*Gallus gallus* [\(more\)](#) 2 citations 21nt [SEQ](#)

Mapped:human(hg17) human(hg18) rat(rn3)

4:FR010708 [acc:AJ421750](#) [MIMAT0000085](#) [miRBase v9.2](#) [\(more\)](#)

mature micro RNA (miRNA) miR-28-5p

SO name:miRNA [\(more\)](#) Organism:*Ateles geoffroyi* [\(more\)](#) 2 citations 22nt [SEQ](#)

詳細画面

[Top](#)
[Statistics](#)
[Blast](#)
[Download](#)
[ncrna.org](#)

[Help](#)

FR000692

[\[Return\]](#)

Summary

Sequence

Map

Summary

ID	FR000692
Description	mature micro RNA (miRNA) miR-128-3p / miR-128a-3p
Accession	AJ459739
Sequence Ontology	miRNA
Organism	
<i>Ateles geoffroyi</i>	<i>Ateles geoffroyi</i> , <i>Ateles jeffroyi</i> , <i>Central American spider monkey</i> , <i>black-handed spider monkey</i>
<i>Bos taurus</i>	<i>Bos Taurus</i> , <i>Bos bovis</i> , <i>Bos primigenius taurus</i> , <i>bovine</i> , <i>cattle</i> , <i>cow</i> , <i>domestic cattle</i> , <i>domestic cow</i>
<i>Danio rerio</i>	<i>Brachidanio rerio</i> , <i>Brachydanio rerio</i> , <i>Cyprinus rerio</i> , <i>Cyprinus rerio Hamilton, 1822</i> , <i>Danio rerio (Hamilton, 1822)</i> , <i>leopard danio</i> , <i>zebra danio</i> , <i>zebra fish</i> , <i>zebrafish</i>
<i>Homo sapiens</i>	<i>human</i> , <i>man</i>
<i>Macaca mulatta</i>	<i>rhesus macaque</i> , <i>rhesus macaques</i> , <i>rhesus monkey</i> , <i>rhesus monkeys</i>
<i>Mus musculus</i>	<i>LK3 transgenic mice</i> , <i>Mus muscaris</i> , <i>Mus sp. 129SV</i> , <i>house mouse</i> , <i>mice C57BL/6xCBA/CaJ hybrid</i> , <i>mouse</i> , <i>nude mice</i> , <i>transgenic mice</i>

ゲノムへのマッピング結果

[Top](#)
[Statistics](#)
[Blast](#)
[Download](#)
[ncrna.org](#)







FR000692

[Return]

Summary

Sequence

Map

Genome	Cytoband	Locus	Strand	Mapping Evidence
human(hg17) 2 regions	2q21.3	chr2:136256748-136256769	+	BLAT 
	3p22.3	chr3:35761023-35761043	+	BLAT 
human(hg18) 2 regions	2q21.3	chr2:136139486-136139507	+	BLAT 
	3p22.3	chr3:35761023-35761043	+	BLAT 
rat(m3) 2 regions	8q32	chr8:116746704-116746724	-	BLAT 
	13q12	chr13:40921667-40921688	+	BLAT 

[Return]

ゲノムブラウザの画面

[Home](#)
[Genomes](#)
[Blat](#)
[Tables](#)
[Gene Sorter](#)
[PCR](#)
[DNA](#)
[Convert](#)
[Ensembl](#)
[NCBI](#)
[PDF/PS](#)

 Mituyama Toutai
  logout

UCSC Genome Browser on Human Mar. 2006 Assembly

move
 zoom in base zoom out

position/search size **22** bp.

chr2 (q21.3) 

chr2:	T	C	A	136139490	G	T	G	136139495	C	C	G	136139500	C	T	C	136139505	T	T	
---->				C	A			A	A	C	C	G	G	T			T	T	
	Secondary structure annotation of RNAz																		
set1_90	>	>	>	>	>	>	>	.	.	.	>	>	>	>	>	.	.	.	>
set1_50	>	>	>	>	>	>	>	.	.	.	>	>	>	>	>	.	.	.	>
set2_90	>	>	>	>	>	>	>	.	.	.	>	>	>	>	>	.	.	.	>
set2_50	>	>	>	>	>	>	>	.	.	.	>	>	>	>	>	.	.	.	>
set3_90	>	>	>	>	>	>	>	.	.	.	>	>	>	>	>	.	.	.	>
set3_50	>	>	>	>	>	>	>	.	.	.	>	>	>	>	>	.	.	.	>
QRNA predictions	mammalian ncRNAs predicted with QRNA																		
TIN RNAs	Human intronic noncoding RNAs																		
FIN RNAs																			
t-inv 5 Noncoding	H-inv release 5.0 non-protein-coding cDNA mapping track																		
ncRNA candidates	150 ncRNA candidates by Sasaki, Y.T.F. et al. (2007) BBRC 357:991-6																		
hsa-mir-128-1	MicroRNAs from miRBase																		
	Novel miRNAs detected by Berezikov et al. (2006) predicted miRNAs																		
Li2006																			
RNAmicro																			
miRim																			
miRNAMap																			
Berezikov																			
mir-128a	>	>	>	>	>	>	>	.	.	.	>	>	>	>	>	.	.	.	>
mir-128a	>	>	>	>	>	>	>	.	.	.	>	>	>	>	>	.	.	.	>
28-3p miR-128a-3p																			
miR-128b-3p																			
28-3p miR-128b-3p																			
UCSC Gene Predictions Based on RefSeq, UniProt, GenBank, and Comparative Genomics																			
R3HDM1																			
R3HDM1																			
R3HDM1																			
RefSeq Gene																			

ズームアウト

