

2020年度「情報生命工学」

# 生命科学を支える塩基配列 データベース・ウェブツール

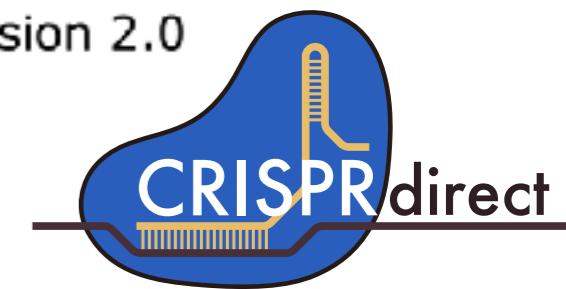
2020年 5月17日  
ライフサイエンス統合データベースセンター (DBCLS)  
**内藤雄樹**

- 内藤 雄樹 (ないとう ゆうき)

Twitter: @meso\_cacase



siDirect version 2.0



- ライフサイエンス統合データベースセンター

Database Center for Life Science =



生命科学データベース研究開発のための機関  
東大(柏)と三島の国立遺伝学研究所に間借り

- DNAやアミノ酸の配列
    - GenBank/ENA/DDBJ, RefSeq, …
  - タンパク質の立体構造 – PDB
  - 遺伝子発現 – GEO, ArrayExpress
  - パスウェイ – KEGG, Reactome
  - 遺伝病 – OMIM
  - 文献 – PubMed
- ……など多岐にわたる

科学において、データは共有されるべきもの

- **再現性** → 現象が真理か？
- **透明性** → データそのものが間違っていないか？
- **再利用** → 異なる視点からの解析
- **技術開発** → 例) 機械学習, AI等の技術の発展
- **公的研究費によるデータの公開**

# データベースを探すためのデータベース

- 生命科学系データベースを一覧から探す -  
**Integbio** データベースカタログ

English



[全条件をリセット](#)
  
[一覧内を検索する](#)

🔍

  
[一覧を絞り込む](#)

- 動物 (6)
- 植物 (2)
- 原生生物 (0)
- 菌類 (2)
- 真正細菌 (5)
- 古細菌 (3)
- ウィルス (27)

  
[タグ<対象>](#)

- ゲノム/遺伝子 (20)
- cDNA/EST (4)
- 遺伝的多様性 (0)
- + 続きを見る

  
[タグ<データの種類>](#)

- 表現型 (0)
- バイオリソース (2)
- 手法 (2)
- + 続きを見る

  
[稼動状況](#)

- 稼動中
- 休止
- 公開停止中
- 運用終了

## データベースのレコード一覧 (件数: 35)

並べ替え: [レコード公開順](#)

サムネイル	データベース名	概要
	<a href="#">pEVE: predicted endogenous viral elements in eukaryotic genomes</a>	運用機関: 東海大学 医学部 生物種: Eukaryota   Viruses 説明: 真核生物ゲノム内に存在するウイルスゲノム様配列 (endogenous viral element: EVE) の予測データベースです。GenomeSync databaseから得られた真核生物およびウィルスのゲノムの配列について... 詳細へ
	<a href="#">EDK: Editome disease knowledgebase</a>	運用機関: National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences 生物種: Homo sapiens   Viruses 説明: RNA編集と疾患の関連についてのデータベースです。mRNA、miRNA、lncRNA、ウイルスのRNA上でおこる実験的に検証された異常な編集イベントと疾患との関係や、RNA編集酵素に関連する異常な活動... 詳細へ
	<a href="#">ViralZone: ViralZone</a>	運用機関: Swiss Institute of Bioinformatics (SIB) 生物種: Viruses 説明: ウィルスの遺伝子や家族ファミリーのためのウェブリソースであり、詳細な分子や役割情報とビリオンやゲノムの情報を提供します。各ウィルス・ファミリーのページから、UniProtKB/Swiss-Prot ... 詳細へ
	<a href="#">Hepatitis C Virus Ires Variation Database: HCVIVdb</a>	運用機関: 生物種: Hepatitis C virus 説明: HCVIVdb は、C型肝炎ウィルスの内部リボソーム進入部位 (IRES) 内で観察される変異についての公開データを集めたデータベースです。 詳細へ
	<a href="#">Virus Pathogen Database and Analysis Resource: ViPR</a>	運用機関: J. Craig Venter Institute (JCVI), Northrop Grumman Health IT, Vencore Technologies 生物種: Arenaviridae   Bunyaviridae   Caliciviridae   Coronaviridae   Filoviridae   ... 説明: 複数のウィルスファミリーのデータと分析ツールを提供する統合リポジトリです。アレルギー・感染症研究所 (NIH) のバイオインフォマティクスリソースセンター (BRC) プログラムの支援を受... 詳細へ
	<a href="#">Ebola and Hemorrhagic Fever Virus Database: HFV Database</a>	運用機関: Theoretical Biology and Biophysics, T-10, Los Alamos National Laboratory

**メニュー**

- [ホーム](#)
- [本カタログについて](#)
- [更新履歴](#)
- [データベース関係マップ](#)
- [ダウンロード](#)
- [ご利用にあたって](#)
- [類似サイトリンク集](#)

**新着情報**

- 2020/05/12: 2019 NAR database issueを基に5件のレコードを追加しました。
- 2020/04/28: 2件のレコードを追加しました。
- 2020/04/22: 2019 NAR database issueを基に3件のレコードを追加しました。
- 2020/04/14: 2019 NAR database issueを基に7件のレコードを追加しました。
- 2020/03/31: 5件のレコードを追加しました。

**本カタログの使い方**

[データベースカタログの使い方](#) 2017

# データベースやツールの紹介動画 (統合TV)



生命科学系DB・ツール使い倒し系チャンネル

[はじめての方へ](#)

[再生数ランキング](#)

[お問い合わせ・番組をリクエスト](#)

『統合TV』は、生命科学分野の有用なデータベースやツールの使い方を動画で紹介するウェブサイトです。

## 目的別に検索

[講習会 実習資料 \(AJACS\)](#)

[ゲノム・核酸 配列解析](#)

[タンパク質 配列・構造解析](#)

[発現制御解析・可視化](#)

[文献・辞書・プログラミング](#)

[著名データベース](#)

[その他講演・講習会](#)

[自由に使える画像を探す](#)

## 関連するタグから検索

ゲノム (363)

遺伝子 (539)

タンパク質 (251)

配列解析 (287)

発現解析 (416)

NGS (302)

## Q 全番組のリストから、調べたいDBやウェブツールに関するキーワードで検索！ (全 20 件)

番組のタイトルや画像をクリックすると番組の再生ページへ移動します。番組リクエストやお問い合わせは[こちらからどうぞ！](#)

表示件数を選ぶ ▾

SNP



### jMorpを使って代謝物の血中濃度とSNPの関係や変異部位の立体構造を調べる

jMorp: Japanese Multi Omics Reference Panel (日本人多層オミックス参照パネル) は、ToMMo (東北メディカル・メガバンク機構)が作成、運用しているデータベースで、大規模な人数の血漿中の代謝物を調べるメタボローム解析や、タンパク質を調べるプロテオーム解析を行った結果を総合し、各代謝物やタンパク質の分布や頻度情報などがまとめられています。[\(原著論文: jMorp: Japanese Multi Omics Reference Panel\)](#) また、数千人分の全ゲノム解析によって得られたSNVの位置とアレル頻度情報をまとめたデータベース（全ゲノムリファレンスパネル）も利用することができます。

TOGO TV

jMorpを使って  
代謝物の血中濃度とSNPの関係や  
変異部位の立体構造を調べる

200414版

NCBI

### NCBI dbSNPとPubMed・OMIMを連携させて遺伝子の多型情報を調べる

NCBI dbSNPは、NCBIが管理・運営しているヒトの一塩基多型 (single nucleotide polymorphism: SNP)に関するデータベースです。今回は、基本的な検索例と

TOGO TV

- 遺伝子名などキーワードで探す
- 遺伝子のさまざまなIDとは？
- 塩基配列から遺伝子を探す

## ● NCBI トップページ

NCBI Resources How To Sign in to NCBI

NCBI National Center for Biotechnology Information

NCBI Home Resource List (A-Z) All Resources Chemicals & Bioassays Data & Software DNA & RNA Domains & Structures Genes & Expression Genetics & Medicine Genomes & Maps Homology Literature Proteins Sequence Analysis Taxonomy Training & Tutorials Variation

All Databases ✓ All Databases Assembly BioProject BioSample BioSystems Books ClinVar Clone Conserved Domains dbGaP dbVar EST Gene Genome GEO DataSets GEO Profiles GSS GTR HomoloGene MedGen MeSH NCBI Web Site NLM Catalog Nucleotide OMIM PMC PopSet Probe Protein Protein Clusters PubChem BioAssay PubChem Compound

Welcome to NCBI

NCBI advances science and providing access to biomedical and genomic information.

NCBI | Mission | Organization | NCBI News | Blog

Submit Download Learn

Transfer NCBI data to your computer

Find help documents, attend a class or watch a tutorial

Analyze Research

Identify an NCBI tool for your data analysis task

Explore NCBI research and collaborative projects

Popular Resources

PubMed Bookshelf PubMed Central PubMed Health BLAST Nucleotide Genome SNP Gene Protein PubChem

NCBI Announcements

HTTPS at NCBI: Guidance for NCBI web API users

As originally announced on June 10, NCBI will be moving all web services to

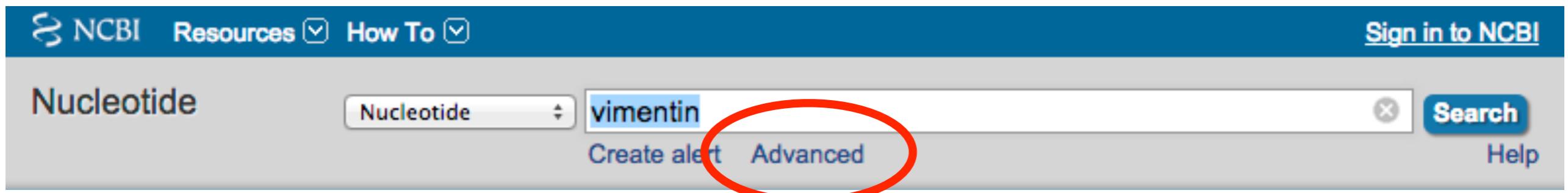
Search

# 絞り込み

## ● 検索窓にキーワードを追加していく

- ... AND “Bombyx mori”[Organism]
- ... AND AGO1 [Gene Name]
- ... AND patent[Title]

## ● または、Advanced searchに行く



- Accession Number
- RefSeq ID
- Gene ID
- Symbol (遺伝子名)

- GenBank/ENA/DDBJ 国際塩基配列データベースに登録された塩基配列のID  
俗にGenBankのAccession番号と呼ばれることがあるが、正確でない。
- A12345 や AB123456 の形をしている  
参考：アルファベットの割り当て  
→ DDBJ の Prefix Letter List
- A12345.1 のようにバージョンを表示  
UTRが延長されたりエラーが修正されて A12345.2 のようにアップデートされる。

- 国際塩基配列データベースに登録された配列をもとに transcriptごとに1個登録 → RefSeqデータベース  
遺伝子の百科事典のようなもの
- 選択的スプライシングで生じるvariantには別々の  
IDが付与されている
- NM\_012345.6 の形式をしている  
実用上はAccession番号の一種として扱うことができる。

- 遺伝子ごとに付与される遺伝子名と番号

慣用名	Symbol	Gene ID
ヒト Argonaute 1	AGO1	26523
ショウジョウバエ Argonaute 1	AGO1	36544
カイコ Argonaute 1	Ago1	100124421
シロイヌナズナ Argonaute 1	AGO1	841262

- Symbolは慣用名と一致しないこともある（ヒトp53→TP53）
- 別の生物種で同一のSymbolがついていることもある
- Gene IDは生物種と遺伝子を特定できる

ヒト Chr22 (q11)

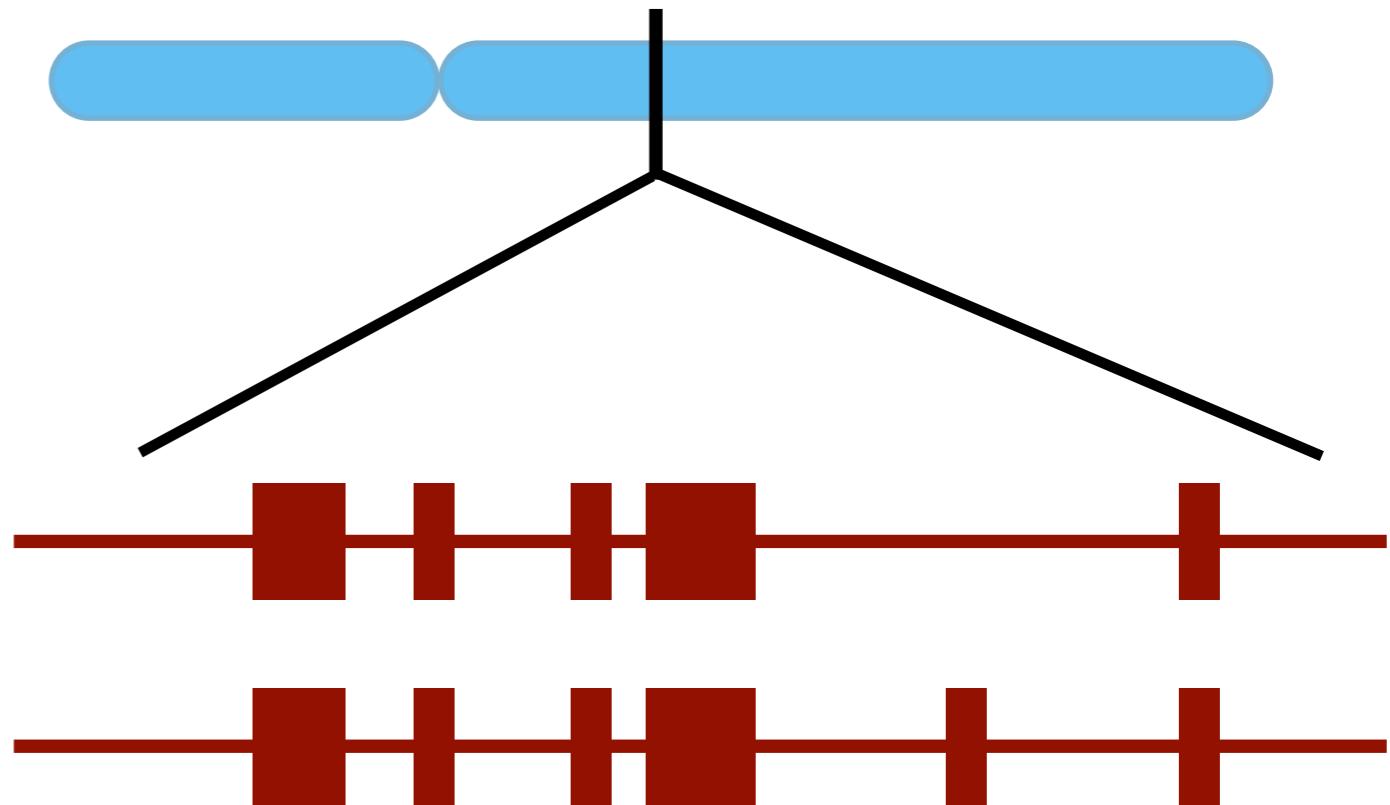
RefSeq ID:

NM\_001190326

NM\_022720

transcriptごと

(塩基配列ごと)



Symbol: DGCR8

Gene ID: 54487

遺伝子 (locus) ごと

## ● NCBI BLAST

<http://www.ncbi.nlm.nih.gov/blast/>

(または **BLAST** をググる)

## ● UCSC BLAT

<http://genome.ucsc.edu/> → Tools / Blat へ

(または **BLAT** をググる)

- 検索ワードが多様（遺伝子名, ID, 遺伝子機能, タンパクのドメイン名, 疾患, 塩基配列, アミノ酸配列, ....）

→ 入口が異なる。知らないと探せない

- BLAST 等による塩基配列の検索

→ 検索が遅い

# 遺伝子をGoogleのように探せれば....

16

Google 検索結果

月が綺麗ですね

すべて 画像 ショッピング 動画 ニュース もっと見る 設定 ツール

約 12,800,000 件 (0.30 秒)

**月が綺麗ですねの意味や返しとは?死んでもいいわ以外の返事や断る時は?**  
<https://ikedanaoya.com> › 恋愛 ▾  
月が綺麗ですね』という一見、何の変哲もない短い表現は、. ここ最近ドラマなどでよく取り上げられたこともあり、. 裏に隠された意味をご存知の方も少なくないかもしれませんね。ロマンチックな雰囲気が漂う静かな夜に、ちょっと気になる相手と一緒に夜空に ...

**「月が綺麗ですね(夏目漱石)」、「死んでもいいわ(二葉亭四迷)」について**  
<https://www.rikaboo.com> › 雑記 ▾  
2017/08/16 - 暫にいる彼が、満月を見上げて「月が綺麗ですね」なんて言い出したらドキッとする?キュンとしたら「死んでもいいわ」って答えちゃう? こんにちは。りかです。今日は、都市伝説のように広がる、二つのことを調べてみました。「愛している」 ...

**「月が綺麗ですね」に隠された意味とは? うまい返し方も | 恋学[Koi-Ga...]**  
<https://koigaku.machicon.jp> › 恋愛, 片思い, 告白 ▾  
2017/12/14 - 男性と二人っきりで夜の町を歩いていて、「月が綺麗ですね」と言わされたら、どうしますか? 「ああ、たしかに綺麗ですよね」と普通に返す、「えっ、突然何言っちゃってんですかw」とちょっと引き気味に返す.....。 実はこの「月が綺麗ですね」には ...

**月が綺麗ですね (つきがきれいですね)とは 【ピクシブ百科事典】**  
<https://dic.pixiv.net> › ... › 経済, 働く, 職業, 作家, 小説家, 文豪, 夏目漱石 ▾  
月が綺麗ですねがイラスト付きでわかる! "I love you"の日本の意訳。夏目漱石による訳と言われる(俗説で信憑性は低い)。通説 小説家・夏目漱石が英語教師をしていたとき、生徒が "I love you" の一文を「我君を愛す>愛する」と訳したのを聞き、「日本人は ...

こちらの検索結果です  
月が綺麗ですね (漫画)  
『月が綺麗ですね』は、伊藤ハチによる日本の漫画作品。『...

統合遺伝子検索 [Help](#) | [Advanced search](#) | [English](#) | [旧バージョン](#)

# GGRNA ver.2

Zoo (All organisms in RefSeq) ▾

遺伝子をGoogleのように検索できるサイトです。 [NCBI RefSeq](#) の transcript を全文検索します。

検索例：

- ・ 「[homeobox](#)」 「[claudin](#)」 ..... フリーワード検索
- ・ 「"RNA interference"」 ..... ダブルクオートで囲ってフレーズ検索
- ・ 「[Argonaute "PAZ domain"](#)」 ..... Argonaute かつ "PAZ domain" のAND検索
- ・ 「[NM\\_001518](#)」 「[10579](#)」 ..... RefSeq IDやGene IDなど各種IDから検索
- ・ 「[symbol:VIM](#)」 ..... 遺伝子名(symbolまたはsynonym)から検索
- ・ 「[ref:Naito](#)」 ..... 文献情報のなかからフリーワード検索
- ・ 「[1552311\\_a\\_at](#)」 ..... マイクロアレイのプローブIDから塩基配列を検索
- ・ 「[aa:KDEL](#)」 ..... アミノ酸配列を検索
- ・ 「[caagaagagattg](#)」 ..... 塩基配列を検索
- ・ 「[comp:caagaagagattg](#)」 ..... 相補鎖を検索
- ・ 「[iub:aggtcannrtgacct](#)」 ..... N, R, Y 等のあいまいな塩基を含む塩基配列を検索
- ・ [詳細な使い方](#)
- ・ ゲノム配列を検索したいときは [GGGenome](#) 《ゲゲゲノム》へ
  - ヒトゲノムの検索例：<http://GGGenome.dbcls.jp/TTCATTGACAACATT>

新着情報：

- 2018-05-30 データベースをRefSeq rel. 88 (May, 2018)に更新。
- 2015-01-13 HTTPSによる暗号化通信に対応 - <https://GGRNA.dbcls.jp/>
- 2013-07-24 ソースを公開 - [GitHub](#)
- 2013-07-08 GGRNA ver.2公開。全生物種のRefSeqを検索できます。
- 2012-05-29 下記論文の日本語による解説を「DBCLSからの成果発信」に掲載。
- 2012-05-29 GGRNAの論文がNucleic Acids Researchに掲載されました。
- [過去の新着情報](#)

# 遺伝子名を検索

統合遺伝子検索  
**GGRNA** ver.2



[Help](#) | [Advanced search](#) | [English](#) | [旧バージョン](#)

2018-06-22 16:28:51, GGRNA : RefSeq release 88 (May, 2018)

## Summary:

- [nanog \(50\)](#)
- [INTERSECTION \(50\)](#)

## Results:

検索語に色がつきます。重なると色が濃く表示されます。

## RefSeqの転写産物

- mRNA (NM\_, XM\_)
- ncRNA (NR\_, XR\_)

### [Homo sapiens Nanog homeobox retrogene P8 \(NANOGP8\), mRNA.](#) (1999 bp)

LOCUS NM\_001355281 1999 bp mRNA linear PRI 01-MAY-2018 DEFINITION Homo sapiens **Nanog** homeobox retrogene P8 (**NANOG**P8), mRNA. ACCESSION NM\_001355281 VERSION NM\_001355281.1 KEYWORDS RefSeq. SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo. REFERENCE COMMENT VALIDATED REFSEQ: This record has undergone validation or preliminary review. The reference sequence was derived from AC021231.8. Summary: This gene represents a transcribed retrogene of...

Synonym: **NANOG**P1; PN8

[NM\\_001355281.1](#) - Homo sapiens (human) - [NCBI](#) - [UCSC](#) - [RefEx\(expression\)](#)

### [Homo sapiens Nanog homeobox \(NANOG\), transcript variant 2, mRNA.](#) (2055 bp)

LOCUS NM\_001297698 2055 bp mRNA linear PRI 07-MAY-2018 DEFINITION Homo sapiens **Nanog** homeobox (**NANOG**), transcript variant 2, mRNA. ACCESSION NM\_001297698 XM\_005253484 VERSION NM\_001297698.1 KEYWORDS RefSeq. SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo. REFERENCE COMMENT REVIEWED REFSEQ: This record has been curated by NCBI staff. The reference sequence was derived from DC425336.1, AY578089.1, AC006517.46 and AI656990.1. On Jul 10, 2014...

[NM\\_001297698.1](#) - Homo sapiens (human) - [NCBI](#) - [UCSC](#) - [RefEx\(expression\)](#)

### [Homo sapiens Nanog homeobox \(NANOG\), transcript variant 1, mRNA.](#) (2103 bp)

LOCUS NM\_024865 2103 bp mRNA linear PRI 07-MAY-2018 DEFINITION Homo sapiens **Nanog** homeobox (**NANOG**), transcript variant 1, mRNA. ACCESSION NM\_024865 VERSION NM\_024865.3 KEYWORDS RefSeq. SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini;

# 遺伝子名を検索

*Gene & transcript search*

[Home](#) | [Help](#) | [Advanced search](#)

[Previous release \(v1\)](#)

**GGRNA** ver.2

2018-06-22 16:31:57, GGRNA.v2 : RefSeq release 88 (May, 2018)

**LOCUS** NM\_001297698 2055 bp mRNA linear PRI 07-MAY-2018  
**DEFINITION** Homo sapiens **Nanog** homeobox (**NANOG**), transcript variant 2, mRNA.  
**ACCESSION** NM\_001297698 XM\_005253484  
**VERSION** NM\_001297698.1  
**KEYWORDS** RefSeq.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
**REFERENCE** 1 (bases 1 to 2055)  
**AUTHORS** You L, Guo X and Huang Y.  
**TITLE** Correlation of Cancer Stem-Cell Markers OCT4, SOX2, and **NANOG** with  
Clinicopathological Features and Prognosis in Operative Patients  
with Rectal Cancer  
**JOURNAL** Yonsei Med. J. 59 (1), 35-42 (2018)  
**PUBMED** 29214774  
**REMARK** GeneRIF: Rectal tumor tissue OCT4 (p<0.001), SOX2 (p=0.003), and  
**NANOG** (p<0.001) expressions were higher than those in adjacent  
tissue.  
**REFERENCE** 2 (bases 1 to 2055)  
**AUTHORS** Seo S, Jeon HY and Kim H.  
**TITLE** Comparison of Cellular Transforming Activity of OCT4, **NANOG**, and  
SOX2 in Immortalized Astrocytes  
**JOURNAL** DNA Cell Biol. 36 (11), 1000-1009 (2017)  
**PUBMED** 28933914  
**REMARK** GeneRIF: High **NANOG** expression is associated with brain neoplasms.  
**REFERENCE** 3 (bases 1 to 2055)  
**AUTHORS** Song KH, Choi CH, Lee HJ, Oh SJ, Woo SR, Hong SO, Noh KH, Cho H,  
Chung EJ, Kim JH, Chung JY, Hewitt SM, Baek S, Lee KM, Yee C, Son  
M, Mao CP, Wu TC and Kim TW.  
**TITLE** HDAC1 Upregulation by **NANOG** Promotes Multidrug Resistance and a  
Stem-like Phenotype in Immune Edited Tumor Cells

# Accession番号を検索

統合遺伝子検索

**GGRNA**

ver.2



TOME PICTURES  
GALLERY

[Help](#) | [Advanced search](#) | [English](#)

[旧バージョン](#)

NM\_003380

検索

Homo sapiens (human)



2018-06-25 04:41:25, GGRNA : RefSeq release 88 (May, 2018)

## Summary:

- [refid:NM\\_003380 \(1\)](#)
- [INTERSECTION \(1\)](#)

## Results:

検索語に色がつきます。重なると色が濃く表示されます。

[Homo sapiens vimentin \(VIM\), mRNA.](#) (2195 bp)

LOCUS **NM\_003380** 2195 bp mRNA linear PRI 10-APR-2018 DEFINITION Homo sapiens vimentin (VIM), mRNA. ACCESSION **NM\_003380**

VERSION **NM\_003380.4** KEYWORDS RefSeq. SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata;

Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo. REFERENCE

COMMENT REVIEWED REFSEQ: This record has been curated by NCBI staff. The reference sequence was derived from DA980400.1, BQ050765.1, BC000163.2 and T50493.1. This sequence is a reference standard in the RefSeqGene project. On Aug...

**NM\_003380.4 - Homo sapiens (human) - NCBI - UCSC - RefEx(expression)**

◀◀ | ページ  / 1 | ▶▶ | C

1件中 1 - 1 を表示

## Data Export:

下記より最大10000件まで検索結果を取得できます。

- タブ区切りテキスト → [表示](#) | [ダウンロード](#)  
エクセル等の表計算ソフトに直接コピペできます。
- JSON形式 → [リンク](#) | [ダウンロード](#)

統合遺伝子検索  
**GGRNA** ver.2

TOE PICTURE GALLERY

[Help](#) | [Advanced search](#) | [English](#) | [旧バージョン](#)

caagaagagattgc

検索 Homo sapiens (human)

## **Summary:**

- seq:caagaagagattgc (10)
  - **INTERSECTION (10)**

## Results:

検索語に色がつきます。重なると色が濃く表示されます。

PREDICTED: Homo sapiens vimentin (VIM), transcript variant X1, mRNA. (1873 bp)

tcctgctggccgagctcgagcagctcaaggccaaggcaagtgcgcctggggacactacgaggagatgcggagctgcgcggcaggtgaccagctaaccacaaagccgcgtcagggt  
gagcgcgacaacctggccgaggacatcatgcgcctccggagaaaattgcaggaggagatgcctcagagagaggaagccaaaaacaccctgcaatcttcagacaggatgttgcataatgcgtctcggcacgtc  
ttgaccttgcacgcacaaagtggaatctttgc**caagaagagattgc**cttttgcagaaaactccacgaagaggaaatccaggagctgcaggctcagattcaggaacacgcattccaaatcgatgtggatgttccaa  
gcctgacctcacggctgcgcgtacgtcagcaatatgaaagtgtggctgccaagaacctgcaggaggcagaagaatggtacaaatccaagttgtgcacctctgaggctgccaaccggaaacaatg  
acgcctcgccaggcaagcaggactgacttccctcaccgtgaa...

position 832

[XM\\_006717500.2 - Homo sapiens \(human\)](#) - [NCBI](#) - [UCSC](#) - [RefEx\(expression\)](#)

## Homo sapiens vimentin (VIM), mRNA. (2195 bp)

tcctgctggccgagctcgagcagctcaaggccaaggcaagtgcgcctggggacactacgaggagatgcggagctgcgcggcagggtggaccagctaaccacaaagcccgctcgaggt  
gagcgcgacaacctggccgaggacatcatgcgcctccggagaaaattgcaggaggagatgcctcagagagaggaagccgaaaacaccctgcaatcttcagacaggatgttacaatgcgtctctggcacgtc  
ttgaccttgaacgc当地有三个不同的突变位点，分别用不同的颜色标注：  
1. 第一个突变位点在第100位，由G变为A，从“gc”变为“ca”。  
2. 第二个突变位点在第150位，由C变为T，从“gg”变为“tt”。  
3. 第三个突变位点在第200位，由G变为A，从“gg”变为“aa”。  
这些突变导致了氨基酸序列的变化，从而影响了蛋白质的功能。

position 1144

NM\_003380.4 - Homo sapiens (human) - NCBI - UCSC - RefEx(expression)

PREDICTED: Homo sapiens dedicator of cytokinesis 5 (DOCK5), transcript variant X8, mRNA. (9068 bp)

統合遺伝子検索  
**GGRNA** ver.2

TOE PICTURE GALLERY

[Help](#) | [Advanced search](#) | [English](#) | [旧バージョン](#)

caagaagagat

検索 Homo sapiens (human)

### **Summary:**

- seq:caagaagagat (465)
  - **INTERSECTION** (465)

## Results:

検索語に色がつきます。重なると色が濃く表示されます。

Homo sapiens family with sequence similarity 90 member A27, pseudogene (FAM90A27P), non-coding RNA. (496 bp)

TITLE The DNA sequence and biology of human chromosome 19 JOURNAL Nature 428 (6982), 529-535 (2004) PUBMED 15057824

ttacgacgaatgacagaaaaaggagcgaagtccaaagtccccagcagcagcagcagcgaagctccgacgcacatcccagaactccccaaagaaaaatgcaggaagccttggaaaggagccagcagaagatttgttgcctgaggcatctaccatgccactgcctgtccacaccac**caagaagagat**ctgtcctggccctgttccacaggtccaccgcctgtcaacaaacccgagatgagattactctgccttcgggtcacaacgattcacctaactgagcacctgtggaccaccaaaggacatggcagggacgttactgcctccctgtccctgttctgaagagctccaccagacccactctcagtgccaggctgccagccaacaggcctgacatgtcctccatggtgctctccagcctgccatgcaggcgcttgcctggccttaatcccaggcagaatcaa...

position 179

NR\_046365.1 - Homo sapiens (human) - NCBI - UCSC - RefEx(expression)

Homo sapiens SUZ RNA binding domain containing 1 pseudogene (LOC100130075), non-coding RNA. (1058 bp)

caccagcaacgggtggtcagcagccccaaagtccgctagcagggccgcctccagtcagtcactggcacagtggaaagccgatcccgaggccaggaaagcggatctgggcagcgccaaccggaga  
gaagcaggagaaaccatcctcgataggcttcctgtatcttcccttcaggccaaaccaggatctctaaccggaaagacagcagacagccaaataatgtatcagacagcctctgggtctgtatgggtcacac  
ggctcaaacacgcgcagataaatgcagg**caagaagagat**ggcgcactgccgtcaacgcgtctgggtcgtccgcaagggttgcactaccgtggcagacagctggacttgagcagcgggaacttgact  
tacttgccctggatccccgttgcctcccccactgtgaccttgaatcccatgactgtgaccccttcctccactgtgattggcacttgacaaggactgtcccaagtcaatggaaaggaaaaagg  
tgagggttaggagaaggtggggaaaccaccaattactcagactcagatcagacagggc...

position 499

NR\_073494.1 - Homo sapiens (human) - NCBI - UCSC - RefEx(expression)

[Homo sapiens long intergenic non-protein coding RNA 2136 \(LINC02136\), long non-coding RNA.](#) (529 bp)

統合遺伝子検索 GGRNA ver.2

[TOP PICTURE GALLERY](#)

[Help](#) | [Advanced search](#) | [English](#) | [旧バージョン](#)

caagaaga

検索

Homo sapiens (human)

## **Summary:**

- seq:caagaaga (32288)
  - INTERSECTION (32288)

## Results:

検索語に色がつきます。重なると色が濃く表示されます。

## [Homo sapiens microRNA 1248 \(MIR1248\)](#), microRNA. (106 bp)

Jones S, Grocock RJ, van Dongen S, Bateman A and Enright AJ. TITLE miRBase: microRNA sequences, targets and gene nomenclature JOURNAL Nucleic Acids Res. 34 (DATABASE ISSUE), D140-D144 (2006) PUBMED 16381832

tttaccttcttgtataaggactgtctaaaattgcagacacttaggaccatgtctggcaataatgctagcagagtacaca**caagaaga**aaagtaaacagca

position 87

Synonym: hsa-mir-1248; mir-1248; MIRN1248

NR\_031650.1 - Homo sapiens (human) - NCBI - UCSC - RefEx(expression)

## Homo sapiens uncharacterized LOC101927523 (LOC101927523), long non-coding RNA. (1014 bp)

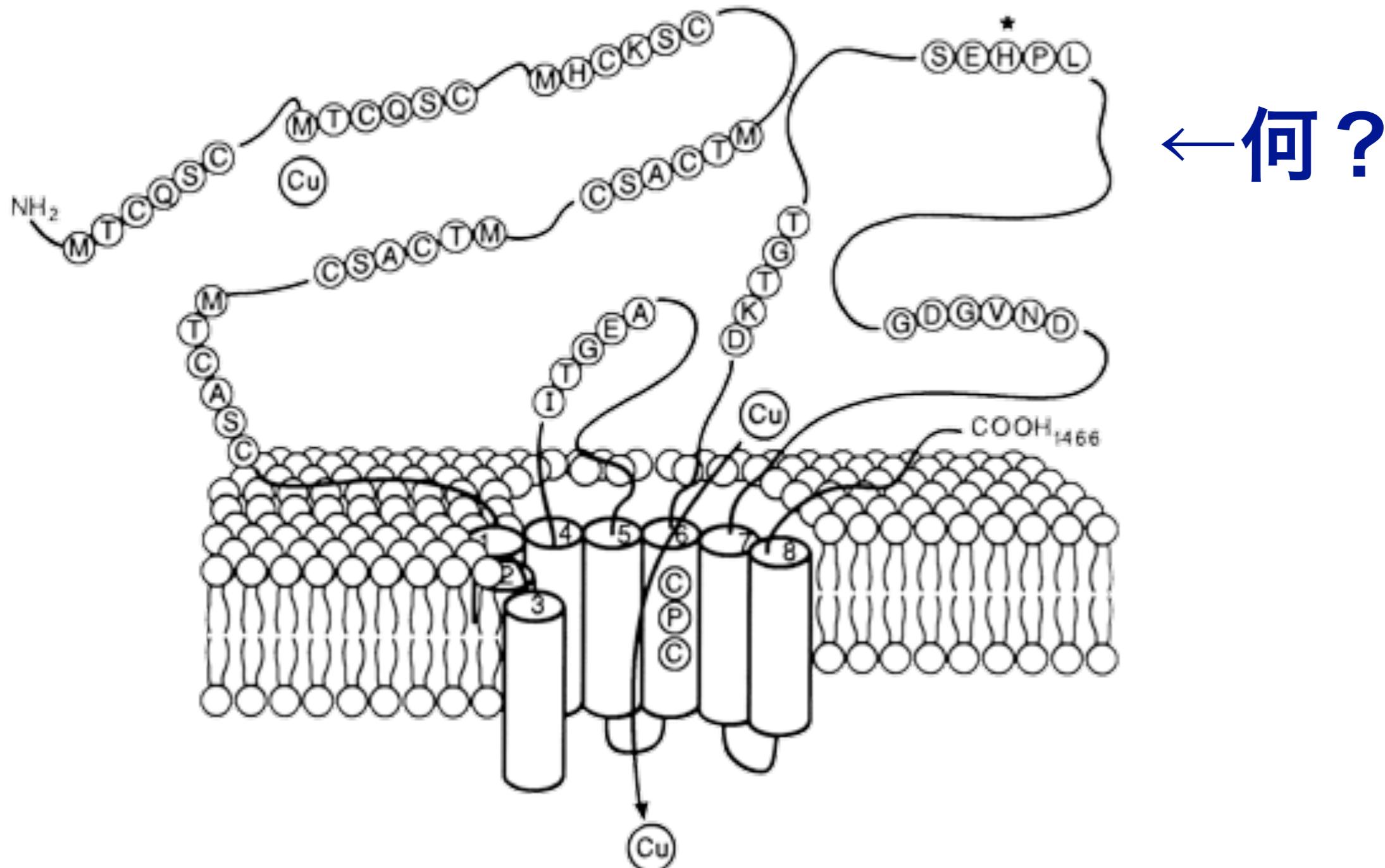
g" /replace="t" /db\_xref="dbSNP:1420081411" variation complement(1011) /gene="LOC101927523" /replace="c" /replace="t" /db\_xref="dbSNP:1047866110" variation complement(1013) /gene="LOC101927523" /replace="a" /replace="c" /db\_xref="dbSNP:1243001703" ORIGIN //

ggttgccaaatggtgtttggctcttcttccaagaagaactagcccattgtcaccaatgcagagctggaaaggctttaaggcagaatgaaaggagaaaagcacgtcctgagcaccaagaag  
aagataactcgaagttaaaaagacgtgatggactttggagaagaagcaacacctgactgagcaacgttcttggatattgataaatctccaggtcttcaacaaggggagtcga  
aataggaaactcattcaaatcatctcagggtgacaagaaggattaccatatttgaggcagcctgtgggtgc...

position 35

NR\_120624.1 - Homo sapiens (human) - NCBI - UCSC - RefEx(expression)

Homo sapiens HLA-DQB1 antisense RNA 1 (HLA-DQB1-AS1), long non-coding RNA. (552 bp)



Schaefer et al. (1999) IV. Wilson's disease and Menkes disease.  
Am. J. Physiol. Gastrointest. Liver Physiol. 276, G311-G314



Licensed under CC-BY 4.0 ©2020 内藤雄樹 (DBCLS)

# アミノ酸配列を検索

統合遺伝子検索

**GGRNA** ver.2

[Help](#) | [Advanced search](#) | [English](#)

[旧バージョン](#)

SEHPL MTCQSC

検索

Zoo (All organisms in RefSeq)

▼

2018-06-25 04:49:29, GGRNA : RefSeq release 88 (May, 2018)

## Summary:

- [SEHPL \(8838\)](#)
- [MTCQSC \(993\)](#)
- [INTERSECTION \(555\)](#)

## Results:

検索語に色がつきます。重なると色が濃く表示されます。

[PREDICTED: Macaca nemestrina ATPase copper transporting beta \(ATP7B\), transcript variant X8, mRNA.](#) (6516 bp)

copper-transporting ATPase 2 isoform X4" /protein\_id="XP\_011754963.1" /db\_xref="GeneID:105490756"

/translation="MKKSFAFDNVGYEGGLDGLGPSSQVDTSTIRILG**MTCQSC**VKSIEDRISSLKGIVSMKVSLEQGSATVKYVPSVVSQQVCHQIGDMGFEASIAEGK  
AASWPSRSLPAQEAVVKLRVEG**MTCQSC**VGSIEGKVRKLQGVVRVKVSLSNQEAVITYQPYLIQPEDLRDHVNDMGFEAAIKNKVAPLSLG...RFSGYFVPLIIIMSTLT  
LVWWIVIGFIDFGVVQKYFPNPNKHISQTTEVIIRFAFQTTSITVLCIACPCSLGLATPTAVMVGTVAAQNGILIKGGKPLEMAHKIKTVMDKTGTITHGVPRVMRVLLGD  
VATLPLRKVLAVVGTAEAS**SEHPL**GVAVTKYCKEELGTETLGYCTDFQAVPGCGIGCKVSNVEGILAHSERPLSAPASHLNEAGNLPAEKDAAPQTFSVLIGNREWLRR  
NGLTISSDVSDAMTDHEMKGQTAILVAIDGMLCGMIAIADAVKQEAALAVHTLQSMGVDVVL..."

AA\_position 35 1035

[XM\\_011756661.2 - Macaca nemestrina \(pig-tailed macaque\) - NCBI](#)

[PREDICTED: Macaca nemestrina ATPase copper transporting beta \(ATP7B\), transcript variant X10, mRNA.](#) (6827 bp)

copper-transporting ATPase 2 isoform X4" /protein\_id="XP\_011754965.1" /db\_xref="GeneID:105490756"

/translation="MKKSFAFDNVGYEGGLDGLGPSSQVDTSTIRILG**MTCQSC**VKSIEDRISSLKGIVSMKVSLEQGSATVKYVPSVVSQQVCHQIGDMGFEASIAEGK  
AASWPSRSLPAQEAVVKLRVEG**MTCQSC**VGSIEGKVRKLQGVVRVKVSLSNQEAVITYQPYLIQPEDLRDHVNDMGFEAAIKNKVAPLSLG...RFSGYFVPLIIIMSTLT  
LVWWIVIGFIDFGVVQKYFPNPNKHISQTTEVIIRFAFQTTSITVLCIACPCSLGLATPTAVMVGTVAAQNGILIKGGKPLEMAHKIKTVMDKTGTITHGVPRVMRVLLGD  
VATLPLRKVLAVVGTAEAS**SEHPL**GVAVTKYCKEELGTETLGYCTDFQAVPGCGIGCKVSNVEGILAHSERPLSAPASHLNEAGNLPAEKDAAPQTFSVLIGNREWLRR  
NGLTISSDVSDAMTDHEMKGQTAILVAIDGMLCGMIAIADAVKQEAALAVHTLQSMGVDVVL..."

AA\_position 35 1035

[XM\\_011756663.2 - Macaca nemestrina \(pig-tailed macaque\) - NCBI](#)

ある遺伝子に対して  
RT-PCRをかけようとしたら  
なぜかバンドが2本・・・

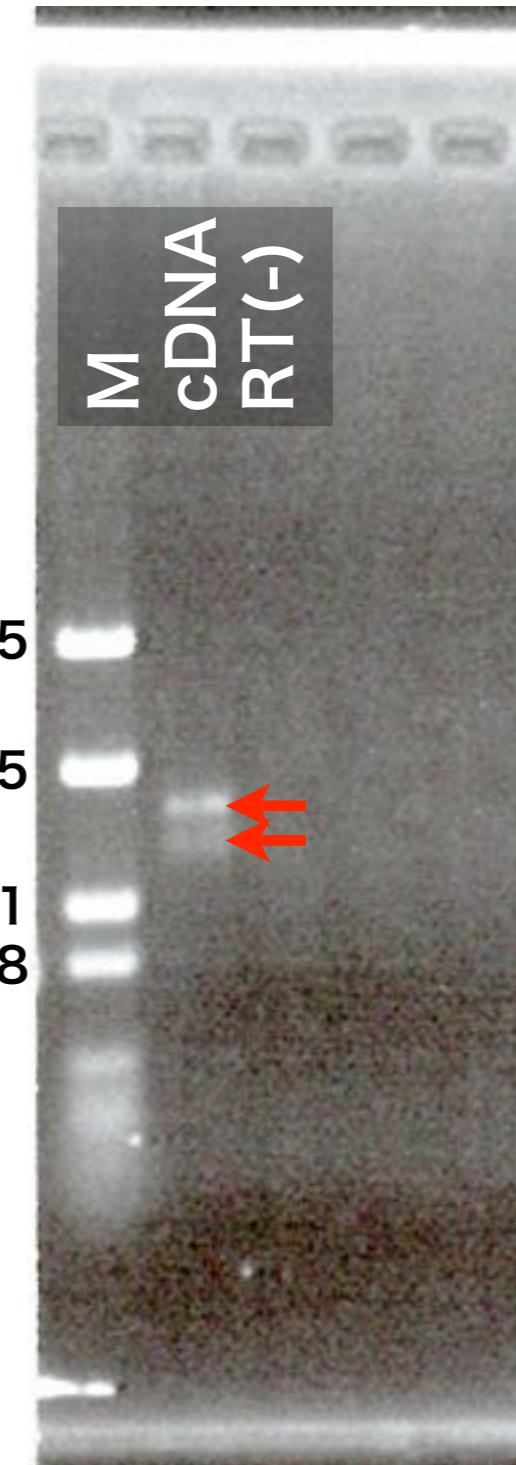
primer-F:

agctcattactttatcagtgcA

primer-R:

tgacgtattcaactcttctggtt

増幅遺伝子は何か？  
予想されるバンドのサイズは？



# PCRのプライマー

- [seq:agctcattactttatcaqtqca \(2\)](#)
- [comp:tqacqtattcactcttctqqtt \(2\)](#)
- **INTERSECTION (2)**

## Results:

トップ50件を表示。検索語に色がつきます。重なると色が濃く表示されます。

[Homo sapiens DGCR8 microprocessor complex subunit \(DGCR8\), transcript variant 2, mRNA. \(4437 bp\)](#)  
 tgaaaaaaattcaggacttgggctgagcggcggaattcaatcggaaatgaagcggaaagcaggcgagtcgcggccatctgccagccaatcaga **agctcattactttatcagtgtca** agatgcacccacaaag  
 aaagagtttattaaaccccaacgggaaatccgaggctgcattgcacgatcatgcgcgtgtcccaaggccgcctgtctataatttttgaatgtgcccggactacactggaaatcctcatccctgactttgtta  
 aacagacctctgaagagaagccaaagacagtgaagaactcgagtatttaaccacatcagcatcgaggactcgccggctacgagctgaccagcaaggctggcttgtctccatatcagatcctccacgagtgccta  
 ...  
 position 1892 2272 **2272 - 1892 = 380**  
 Synonym: C22orf12; DGCR8; Gyr; pasha  
[NM\\_001190326.1 - Homo sapiens \(human\) - NCBI - UCSC - RefEx\(expression\)](#)

[Homo sapiens DGCR8 microprocessor complex subunit \(DGCR8\), transcript variant 1, mRNA. \(4536 bp\)](#)  
 gaagcggaaagcaggcgagtcgcggccatctgccagccaatcaga **agctcattactttatcagtgtca** agatgcacccacaaagaaagatgttattaaaccccaacggaaatccgaggctgcattgcac  
 cgagtagatgcgcgtgtcccaaggccgcctgtctataatttttgaatgtgagaacccaaagtgcgcgtgtccgcgtgaccattgtggatctggacttgcacggatctggaaactgcacggaaaacttgcga  
 agaataaaagctgcggagctacactggaaatccatccctgactttgttacccatcagatcccacgactgcctaaagaaaccatggatggatggatggacttgcacgttatcaagttgaatgtggtctggaaa **aaccagaagagtgtacgtca** tggcgtgt  
 ggcaaggcacacagtgcgcgtgtgttataaaaacaagatgt...  
 position 1892 2371 **2371 - 1892 = 479**  
 Synonym: C22orf12; DGCR8; Gyr; pasha  
[NM\\_022720.6 - Homo sapiens \(human\) - NCBI - UCSC - RefEx\(expression\)](#)

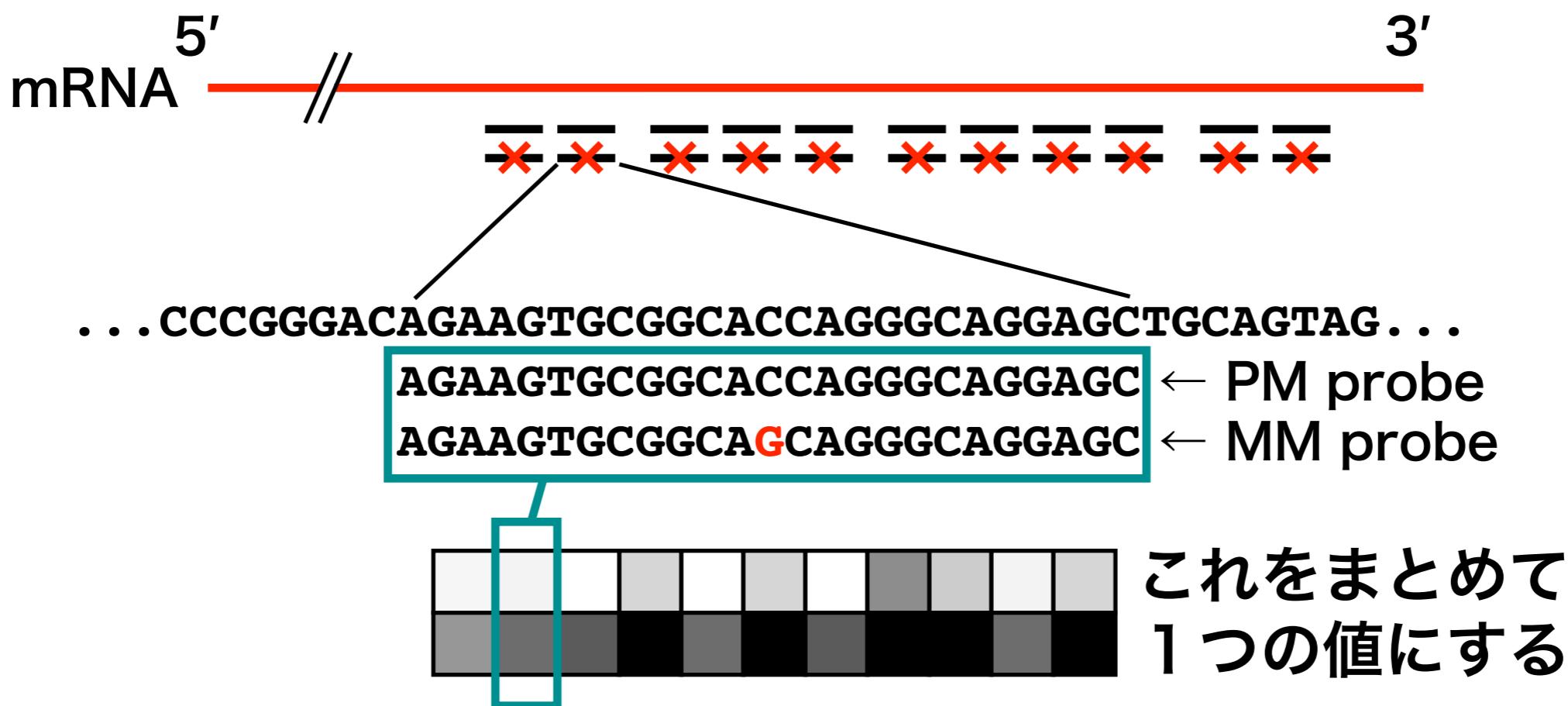
## Data Export:

下記より最大10000件まで検索結果を取得できます。

- タブ区切りテキスト → [表示 | ダウンロード](#)  
 エクセル等の表計算ソフトに直接コピペできます。
- JSON形式 → [リンク | ダウンロード](#)

## Debug Info:

Affymetrix社GeneChipマイクロアレイ  
1遺伝子につき25塩基 × 11箇所のプローブ  
→ 「プローブセット」 例) 1552311\_a\_at



統合遺伝子検索  
**GGRNA** ver.2

*Toddler PICTURE GALLERY*

[Help](#) | [Advanced search](#) | [English](#) | [旧バージョン](#)

1552311\_a\_at

検索 Homo sapiens (human)

## **Summary:**

- seq: TCTCCACAAACGTTTTAAAATGTG (2)
  - seq: GGGACACGGCAGTAAGCACAAGAAA (2)
  - seq: GGCAGGAGCTGCAGTAGCTACCCCTC (2)
  - seq: GCATGGGATGGGACACTCTGGGCCA (2)
  - seq: ATGTGCCGGGTGTACTGGTGCACAC (2)
  - seq: AGGTCACCCCATCTCTAGGCCGAC (2)
  - seq: AGATCACTCCCAGATCACCAAGGTCA (2)
  - seq: AGAAGTGCAGCACCAGGGCAGGAGC (2)
  - seq: ACGGCAGTAAGCACAAGAAAAGATT (2)
  - seq: AATGTCACCGCACACCAGGCAGTGG (2)
  - seq: AAAATGTGCCGGGTGTACTGGTGCA (2)

• **INTERSECTION (2)**

## Results:

検索語に色がつきます。重なると色が濃く表示されます。

*Homo sapiens retina and anterior neural fold homeobox 2 (RAX2), transcript variant 2, mRNA.* (2190 bp)

cagccgcatacgcccaagagctctgagcaaggaggctgtcgccccggcagaacccgtggactggcaagcacggctggcccagtgcagcaggagggggccctgaggcatggatggacagtctgg  
gccagccgcaccctcccgacagaagtgcggcaccaggcaggagctgcagtagctaccctcccgctccagcctggctccccagatcactcccagatcaccaggtcacccatctctaggg  
ggcaccctcacacaccagtcctgtggtccaacgcggccatcaccaatgtcaccgcacaccaggcagttggacacggcagtaaggcacaagaaagattttttaaagctaaaccaggccagg  
gcgggtggctcatgcctgtaatcccagtgtttggaggctgaggtggaggattgcttgagaccagcctgggtgacacagcaagacccatctccacaaacgttttaaaatgtgccgggtgtactgg  
acacctgtcatcccaqctaccqaqaqctqaggcaqaqaqqatcattqaqccqaqaqgtcqaggctcaqggagctgtqatcacactq...

position 1592 1634 1650 1698 1717 1783 1807 1812 1955 1972 1975

Synonym: ARMD6; CORD11; QRX; RAXL1

NM\_032753.3 - Homo sapiens (human) - NCBI - UCSC - RefEx(expression)

Homo sapiens retina and anterior neural fold homeobox 2 (RAX2), transcript variant 1 mRNA (2428 bp)

<http://GGRNA.dbcls.jp/>

- 遺伝子をGoogleのように検索
- 検索対象はRefSeqの転写産物
  - RefSeq ID : NM\_\*, XM\_\* = mRNA
  - RefSeq ID : NR\_\*, XR\_\* = ncRNA
- さまざまなキーワードに対応
- 塩基配列、アミノ酸配列も素早く検索
  - とくに短い配列の検索が得意
- 無償で自由に利用可能
  - 商用利用であっても無償。
  - 利用記録は当センターにおいて機密扱い。

<http://GGGenome.dbcls.jp/>

### 検索例：

- [ [TTCATTGACAAACATT](#) ] ..... 塩基配列を検索
- [詳細な使い方](#)
- 遺伝子や転写産物のキーワード検索は [GGRNA](#) 《ぐぐるな》へ
  - 例) ヒトの「nanog」を検索：<http://GGRNA.dbcls.jp/hs/nanog>

### 検索結果へのリンク：

- [http\[s\]://GGGenome.dbcls.jp/db/k/\[strand\]/sequence\[.format\]\[.download\]](http://GGGenome.dbcls.jp/db/k/[strand]/sequence[.format][.download])
  - **db** → hg19, mm10, dm3, ce10, TAIR10, pombe, refseq, etc. 省略時は hg19  
[検索可能なデータベース一覧](#)
  - **k** → 許容するミスマッチ/ギャップの数。あまり大きいとしぼうする。省略時は 0
  - **strand** → '+' ('plus') または '-' ('minus') で特定の方向のみ検索。省略時は両方向を検索
  - **sequence** → 塩基配列。大文字・小文字は区別しない
  - **format** → html, txt, csv, bed, gff, json。省略時は html
  - **download** → URLの最後に付加すると検索結果をファイルとしてダウンロードできる
- 例1：<http://GGGenome.dbcls.jp/TTCATTGACAAACATT>
  - ヒトゲノム **hg19** (省略可) で
  - ミスマッチ/ギャップを許容せず (省略可)
  - **TTCATTGACAAACATT** を検索し
  - **html** 形式 (省略可) で結果を返す
- 例2：<http://GGGenome.dbcls.jp/mm10/2/+//TTCATTGACAAACATTGCGT.txt>
  - マウスゲノム **mm10** で
  - 2 ミスマッチ/ギャップまで許容して

超絶高速ゲノム配列検索  
**GGGenome**

[Help](#) | [English](#)

TTCACTGACAACATTGAGTA  Human genome, GRCh37/hg19 (Feb, 2009) 

許容するミスマッチ/ギャップの数:  (検索する塩基配列の長さの25%まで)

双方向を検索  +方向のみ検索  -方向のみ検索

2017-06-28 06:20:32, GGGenome : Human genome, GRCh37/hg19 (Feb, 2009)

**Summary:**

- [TTCACTGACAACATTGAGTA \(1\)](#)
- [TACTCAATGTTGTCAGTGAA \(0\)](#)
- **TOTAL (1)**

**Results:**

検索語に色がつきます (ミスマッチ・挿入欠失)。

[chr2:30898313-30898332](#) ▼30898313  
GCCCGCAGCCACTTATATGTATTTTTAAATTCACTGACAACATTGAGTAGAAAAGATAATTTTTTTGAGACAG

« « | ページ  / 1 | > > | C

1件中

**Data Export:**

下記のフォーマットで最大100000件まで検索結果を取得できます。

- タブ区切りテキスト → [表示 | ダウンロード](#)  
エクセル等の表計算ソフトに直接コピペできます。
- CSV形式 → [表示 | ダウンロード](#)  
エクセル等の表計算ソフトにて開くことができます

超絶高速ゲノム配列検索  
**GGGenome**

[Help](#) | [English](#)

TTCACTGACAACATT  Human genome, GRCh37/hg19 (Feb, 2009) 

許容するミスマッチ/ギャップの数:  (検索する塩基配列の長さの25%まで)

双方向を検索  +方向のみ検索  -方向のみ検索

2017-06-28 06:20:47, GGGenome : Human genome, GRCh37/hg19 (Feb, 2009)

#### Summary:

- [TTCACTGACAACATT \(10\)](#)
- [AATGTTCTAGTGAA \(7\)](#)
- **TOTAL (17)**

#### Results:

検索語に色がつきます (ミスマッチ・挿入欠失)。

[chr1:173400829-173400843](#) ▼173400829  
 TTTTCACATGCACTACCTGCCTGCTGTCTTCACTGACAACATTTCAGGGTCTCGCAAGGCATGGAATACAATG

[chr1:197463616-197463630](#) ▼197463616  
 AATGATCTTACTGTTAACACTGCACTTATTCACTGACAACATTACTTAAAGAAAAGAGATTATGATTGGATAAA

[chr2:30898313-30898327](#) ▼30898313  
 GCCCCCAGCCACTTATATGTATTTTTAATTCACTGACAACATTGAGTAGAAAAGATAATTTTTTTTTGA

[chr2:163535086-163535100](#) ▼163535086  
 TGAGAAAGAGGACACCCTAACGAAATTATGGCTTCACTGACAACATTAGAGAACATATCAAACGAAATATT

[chr3:117135824-117135838](#) ▼117135824  
 TCCCCTAGCAAATCAATTCAACTTCTGTTATTCACTGACAACATTGTTTTTCTACTGGCTGCAGCTTTCACAA

chr1:128011007-128011021

超絶高速ゲノム配列検索  
**GGGenome**

[Help](#) | [English](#)

TTCACTGACA  Human genome, GRCh37/hg19 (Feb, 2009) 

許容するミスマッチ/ギャップの数: 0 (検索する塩基配列の長さの25%まで)

双方向を検索  +方向のみ検索  -方向のみ検索

2017-06-28 06:21:00, GGGenome : Human genome, GRCh37/hg19 (Feb, 2009)

#### Summary:

- [TTCACTGACA \(4957\)](#)
- [TGTCAGTGAA \(5068\)](#)
- **TOTAL (10025)**

#### Results:

検索語に色がつきます (ミスマッチ・挿入欠失)。

[chr1:620618-620627](#) ▼620618  
 ATACTAGAATTAAAAGCAAATGTCTTTTATTCACTGACAACATAATCATCTATAAAGAAAATCCTACAT

[chr1:728363-728372](#) ▼728363  
 CCCTCAGCCACACAGCTGACGCTGTAGAGATTCACTGACATCCAGTCTGATGGACACTGTTCTCCAGCC

[chr1:1336263-1336272](#) ▼1336263  
 CCGCAGGGCTCTGCCACATCCCTGTCTTCTTCACTGACATGAAACGCAGAAAGGCAGCTTGCACAA

[chr1:3274566-3274575](#) ▼3274566  
 AGAGGCCCACTCGTGAGCCTTGCTGTGCCCTTCAGATTCACTGACATGCTCGCCCATGAGCTGAGTCTGCCTTGAC

[chr1:3783925-3783934](#) ▼3783925  
 CTCCGCTGCGTGTAGATACCACACTGTCTATTCACTGACACTTGGGTTGCTTCCACCTTTGGCCTGGCC

chr1:6720605-6720604

超絶高速ゲノム配列検索  
**GGGenome**

[Help](#) | [English](#)

TTCACTGACAACATTGAGTA  Human genome, GRCh37/hg19 (Feb, 2009) 

許容するミスマッチ/ギャップの数 :  (検索する塩基配列の長さの25%まで)

双方向を検索  +方向のみ検索  -方向のみ検索

2017-06-28 06:21:13, GGGenome : Human genome, GRCh37/hg19 (Feb, 2009)

#### Summary:

- TTCACTGACAACATTGAGTA (36)
- TACTCAATGTTGTCACTGAA (33)
- **TOTAL (69)**

#### Results:

検索語に色がつきます (ミスマッチ・挿入欠失)。

chr1:77930953-77930971 ▼77930953  
 AGCAGTACTCAAGATCACCTTTTTAAAGTTCACTG-CAACATTCAGTAGATGAAAAATGTGTAGCTGTAATTCATTGA

chr1:106947513-106947532 ▼106947513  
 TAGAAAACATAACCACATAAAGAACATCTTTCACTGCACAA-ATTGAGTAAAAATTAAACCCGTTGAAAGGGTCATTA

chr2:30898313-30898332 ▼30898313  
 GCCCCCAGCCACTTATATGTATTTTTATTCACTGACAACATTGAGTAGAAAAGATAATTTTTTTTTGAGACAG

chr2:141517099-141517116 ▼141517099  
 TTTATTTCTTAACACAGCCTTGATAAATTCACTGACATTGAGT-TAGCCATAGTGTGAATGGGAGACTGCCA

chr2:163535086-163535103 ▼163535086  
 TGAGAAAGAGGACACCCTAAGAATTATGGTTCACTGACAACATT-AG-AGAACATATCAAACCAAACGAATATTCTGG

chr2:186261205-186261224

# GGGenome による塩基配列検索

- アドレス = 検索クエリ になっている
- 結果をさまざまな形式で出力できる

## 検索例：

http[s]://GGGenome.dbcls.jp/hg38/2/ATAGGACCATA[.txt]

暗号化通信  
に対応

①ヒトゲノム  
hg38に対し

②最大ミスマッチ・  
挿入欠失 = 2 で

③この配列を検索

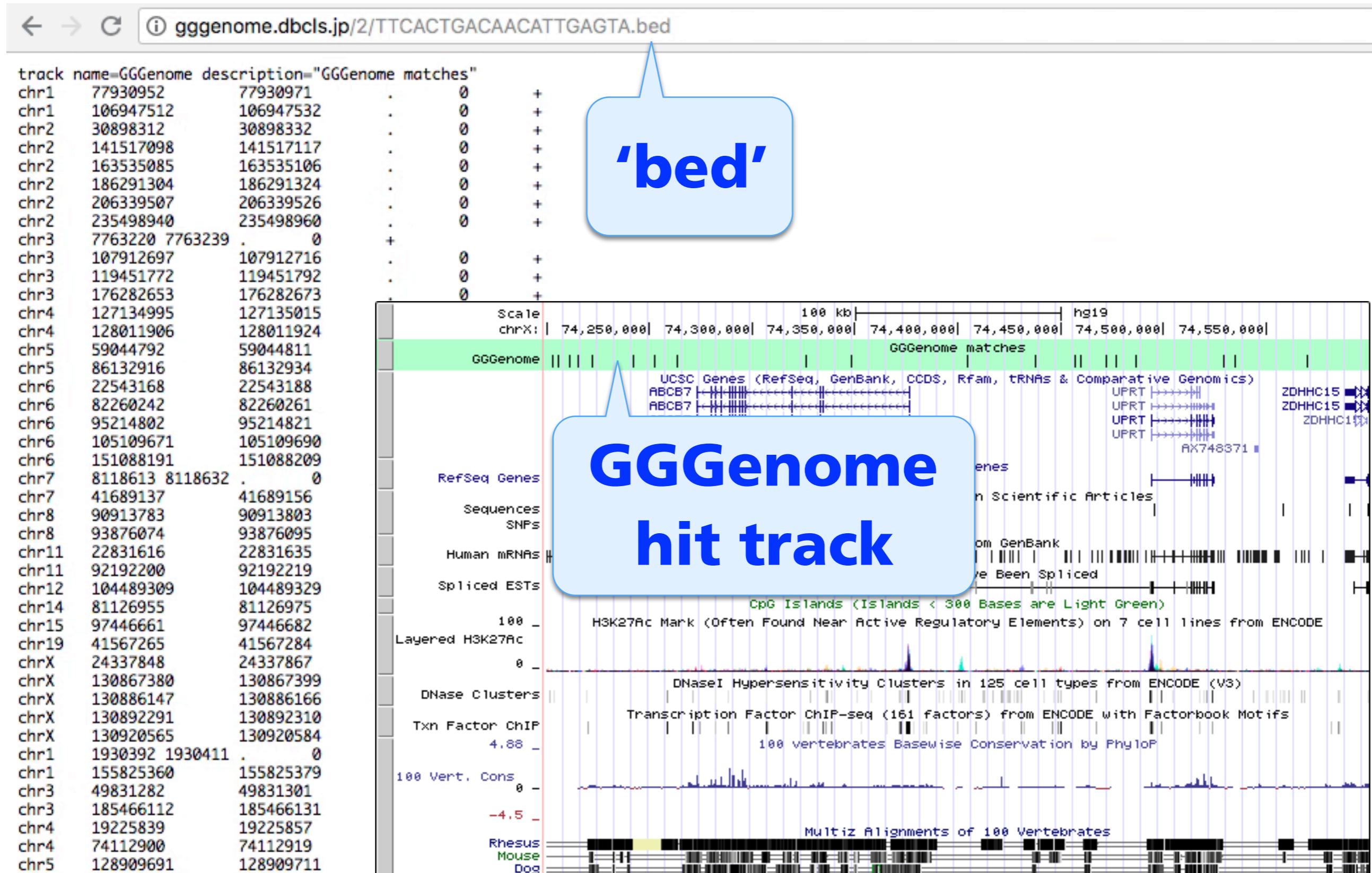
④結果はテキスト  
(txt) 形式で。

# GGGenome を他のツールから呼び出す

38

# GGGenome を他のツールから呼び出す

39



# GGGenome を他のツールから呼び出す

40



The screenshot shows a web browser window with the URL `gggenome.dbcls.jp/2/TTCACTGACAACATTGAGTA.json`. The page displays a JSON object representing a sequence alignment. A blue callout bubble with the text 'json' points to the file extension in the URL.

```
{...

- database: "Human genome, GRCh37/hg19 (Feb, 2009)",
- error: "none",
- results:
  - [...
    - {
      - align: "||||||| ||||| ||||",
      - del: 0,
      - edit: "-----I-----X---",
      - ins: 1,
      - match: 18,
      - mis: 1,
      - name: "chr1",
      - position: 77930953,
      - position_end: 77930971,
      - query: "TTCACTGACAACATTGAGTA",
      - sbjct: "TTCACTG-CAACATTGAGTA",
      - snippet: "TCATGAAATGGAAGTGCTGCCACTTCAGGGTCTGTAAAGAGATAATGCTGATGCCAGGCCAGGGAAAAAGCAGTACTCAAGATCACCTTTTAAAGTTCACTGCAAC",
      - snippet_end: 77931071,
      - snippet_pos: 77930853,
      - strand: "+"
,
      - {
        - align: "||||||| ||||| |||||",
        - del: 1,
        - edit: "-----D-----I-----",
        - ins: 1,
        - match: 19,
        - mis: 0,
        - name: "chr1",
        - position: 106947513,
        - position_end: 106947532,
        - query: "TTCACTG-ACAACATTGAGTA",
        - sbjct: "TTCACTGCACAA-ATTGAGTA",

```

```
19:46:08 meso@mesobook:~ curl -s 'https://GGGenome.dbcls.jp/hg38/GTCGTATGCTAG.bed'  
track name=GGGenome description="GGGenome matches"  
chr1 194830236 194830248 . 0 +  
chr2 78670750  
chr2 218571531  
chr12 55980766  
chr12 116320454  
chr15 91363640  
chrX 151249981  
chrX 151250002  
chr1 158411570 158411582 . 0 -  
chr2 14459609 14459621 . 0 -  
chr4 27155130 27155142 . 0 -  
chr7 145345362 145345374 . 0 -  
chr8 4213930 4213942 . 0 -  
chr8 91253328 91253340 . 0 -  
chr8 128369312 128369324 . 0 -  
chr10 10395010 10395022 . 0 -  
chr10 127341893 127341905 . 0 -  
chr15 52581403 52581415 . 0 -
```

curl コマンド：  
webから情報を取得

wc コマンド：  
行数をカウント

```
19:46:25 meso@mesobook:~ curl -s 'https://GGGenome.dbcls.jp/hg38/GTCGTATGCTAG.bed' | wc -l  
19  
19:46:31 meso@mesobook:~ curl -s 'https://GGGenome.dbcls.jp/hg38/GTCGTATGCTAG.bed' | grep chr8  
chr8 4213930 4213942 . 0 -  
chr8 91253328 91253340 . 0 -  
chr8 128369312 128369324 . 0 -  
19:46:36 meso@mesobook:~
```

grep コマンド：  
キーワード抽出

# 表計算ソフトから GGGenome 検索

42

GGGenome primer search ☆ meso cacase

ファイル 編集 表示挿入 表示形式 データ ツール ヘルプ 変更内容をすべてドライブに保存しました コメント 共有

fx =ImportData(D2)

A	B	C	D	E	F	G	H	I
1	Primer name	(1) Sequence	(2) GGGenome URL	(3) ImportDATA function	name	strand	start	end
2	YN001-F	caatcaccctcaccctttatatgc	<a href="http://GGGenome.dbcls.jp/rice/caatcaccctcaccctttatatgc.txt">http://GGGenome.dbcls.jp/rice/caatcaccctcaccctttatatgc.txt</a>	=ImportData(D2)	chr01	+	27707	27
3	YN001-R	ccgctgtgaac		chr01 - 28071 28092				
4	YN002-F	tgatccaata		ATATTGTTGGTACCGAGTA				
5	YN002-R	catgcaaggta		27971 28192	chr01	-	28071	28
6	YN003-F	acgtactgtgg						
7	YN003-R	acccaaccta						
8	YN004-F	atcagattccggccggccg						
9	YN004-R	ggagagatctgggtgggag						
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								

= ImportData(D2)

PCRプライマー配列

# 表計算ソフトから GGGenome 検索

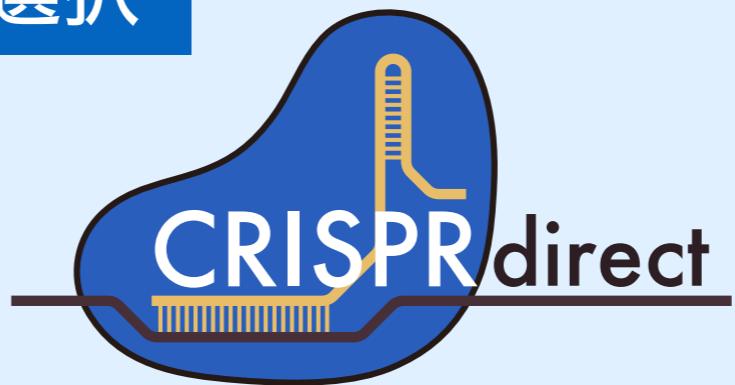
43

<http://GGGenome.dbcls.jp/>

- ゲノムを高速に検索可能  
350種の生物種（うち植物は80種）に対応
- 短い塩基配列の検索に強い  
ミスマッチや挿入欠失が多くても見落としのない検索
- 他のツールから呼び出すことができる（REST API）  
Googleスプレッドシートの ImportData() 関数  
Excel 2013以降の WEBSERVICE() 関数  
R、Ruby、Galaxy、ゲノムブラウザ・・・
- 無償で自由に利用可能  
商用利用であっても無償。  
利用記録は当センターにおいて機密扱い。

PAMに隣接する配列の選択

ガイドRNA  
設計支援ツール



オフターゲット配列検索

GGGenome

高速で正確な塩基配列検索ツール

GAAGGCAGTCCAGTGAAATG  
| | | | | | | | | | | | | |  
GAAGCAGTACAGAGAAATG



wordが完全一致せず  
BLASTで検索できない



<http://crispr.dbcls.jp/>

Enter an accession number (e.g. NM\_006299) or genome location (e.g. hg19:chr7:1)

 retrieve sequence

Accession番号

or Paste a nucleotide sequence: ?

```
agtgcaccccccgttcgcctcgctcgagctcgcttactttacaccggccggccgagctctccagactccggaggaggaaagcg  
atcggttacacgtacgcctcgtaaggcagaaggcagaacatctgaagaatggctgacgaggatattcaacctatttgtgtgcgacaatg  
gcactggaatggtaaggcagggtttgtgtatgtgcaccaaggccgtcttcccttagcatttagggagaccacgtcacacc  
ggtgtcatggttgggatgggccaaggatgcctatgtgggtatgaagctcaggcaaaaagggtatcctgactctaaagtaccc  
aattgaacatggaaattgtcaataactgggatgacatggagaaaaatatggcaccacccatcataatgagcttcgtgtgcacctg  
aagatcacccctgtattactaactgaagccccctctcaatcccaaagccaacagagagaagatgacacagatdatgtttgagacctt  
aattgcccaatgtatgtcgcaatccaggctgttatcctgt
```

または  
塩基配列 を入力

or upload sequence file: ? ファイルを選択 選択されていません

PAM sequence requirement: NGG (e.g. NGG, NRG) ?

Specificity check: Human (Homo sapiens) genome, GRCh37/hg19 (Feb, 2009)



?

design

What's new:

- 2017-06-19 Added 8 species - List

# CRISPRdirect によるガイド RNA 設計

47

- Target sequences with TTTTs are also shown in gray. Avoid TTTTs in gRNA vectors with pol III promoter.

show **highly specific** target only

Show 20 entries Search:

position		target sequence	sequence information				number of target sites		
start - end	+/-	20mer+PAM (total 23mer)	GC% of 20mer	Tm of 20mer	TTTT in 20mer	restriction sites	20mer +PAM	12mer +PAM	8mer +PAM
7 - 29	-	cctcccggttcgcctcgctcgcg [gRNA]	75.00 %	83.48 °C	-	NruI	1 [detail]	5 [detail]	370 [detail]
<b>10 - 32</b>	-	cccgtttcgccctcgctcgcgagct [gRNA]	70.00 %	83.41 °C	-		1 [detail]	1 [detail]	276 [detail]
11 - 33	-	ccgttccgcctcgctcgcgagctc [gRNA]	70.00 %	83.41 °C	-		4 [detail]	4 [detail]	911 [detail]
18 - 40	-	cctcgctcgcgagctcgcttac [gRNA]	70.00 %	83.41 °C	-		4 [detail]	4 [detail]	1072 [detail]
47 - 69	-	ccggccggccgagctctccaga [gRNA]	75.00 %	80.41 °C	-		305 [detail]	305 [detail]	79033 [detail]
50 - 72	-	ccggccggagctctccagactc [gRNA]	70.00 %	80.41 °C	-	SacI	1 [detail]	37 [detail]	3555 [detail]
53 - 75	+	ccggccggagctctccagactc [gRNA]	70.00 %	80.41 °C	-	SacI	1 [detail]	3 [detail]	175 [detail]
53 - 75	-	ccggccggagctctccagactccgg [gRNA]	70.00 %	80.27 °C	-	SacI	1 [detail]	15 [detail]	2679 [detail]
56 - 78	+	ccggccggagctctccagactccgg [gRNA]	70.00 %	80.27 °C	-	BspEI SacI	1 [detail]	4 [detail]	793 [detail]
<b>56 - 78</b>	-	ccggccggagctctccagactccggagg [gRNA]	65.00 %	79.68 °C	-	BspEI SacI	1 [detail]	1 [detail]	1295 [detail]
59 - 81	+	agctctccagactccggaggagg [gRNA]	65.00 %	79.68 °C	-	BspEI	1 [detail]	2 [detail]	608 [detail]
65 - 87	-	ccaaactccggaggaggaggacg [gRNA]	65.00 %	78.93 °C	-	BspEI	1 [detail]	6 [detail]	375 [detail]
72 - 94	-	ccggaggaggaggacgatcgttac [gRNA]	55.00 %	73.00 °C	-	BsiEI PvuI	1 [detail]	14 [detail]	3180 [detail]
102 - 124	-	cctcggtcaaggcagaaggcagaaca [gRNA]	50.00 %	71.33 °C	-		0 [detail]	3 [detail]	588 [detail]
114 - 136	+	gaaggcagaacatctgaagaaatgg [gRNA]	40.00 %	65.99 °C	-		0 [detail]	7 [detail]	1117 [detail]
123 - 145	+	catctgaagaatggctgacgtgg [gRNA]	50.00 %	69.80 °C	-		1 [detail]	2 [detail]	1388 [detail]
151 - 173	+	caaccttattgtgtgcgacaaatgg [gRNA]	45.00 %	69.15 °C	-		1 [detail]	3 [detail]	303 [detail]
<b>154 - 176</b>	-	cctattgtgtgcgacaaatggcac [gRNA]	50.00 %	72.27 °C	-		1 [detail]	1 [detail]	667 [detail]
157 - 179	+	attgtgtgcgacaaatggcacatgg [gRNA]	50.00 %	72.27 °C	-		1 [detail]	2 [detail]	499 [detail]
162 - 184	+	gtgcgacaaatggcactggaaatgg [gRNA]	55.00 %	74.69 °C	-		1 [detail]	6 [detail]	677 [detail]

Showing 1 to 20 of 71 entries

First Previous 1 2 3 4 Next Last

標的サイトの  
位置と塩基配列

Graphical View:

# CRISPRdirect によるガイド RNA 設計

48

- Target sequences with TTTTs are also shown in gray. Avoid TTTTs in gRNA vectors with pol III promoter.

show **highly specific** target only

Show 20 entries		sequence information				number of target sites <small>?</small>		
position start - end	target sequence	GC% of 20mer	Tm of 20mer	TTTT in 20mer	restriction sites	20mer +PAM	12mer +PAM	8mer +PAM
7 - 29	cctcccggttcgcctcgctcgcg [gRNA]	75.00 %	83.48 °C	-	NruI	1 [detail]	5 [detail]	370 [detail]
<b>10 - 32</b>	cccgtttcgccctcgctcgcgagct [gRNA]	70.00 %	81.73 °C	-	NruI	1 [detail]	1 [detail]	276 [detail]
11 - 33	ccgttccgcctcgctcgcgagctc [gRNA]	70.00 %	81.09 °C	-	NruI SacI	1 [detail]	4 [detail]	911 [detail]
18 - 40	cctcgctcgcgagctcgcttac [gRNA]	70.00 %	80.24 °C	-	NruI SacI	1 [det]		
47 - 69	ccggccggccgagctctccaga [gRNA]	75.00 %	84.25 °C	-	SacI	1 [det]		
50 - 72	ccggccggagctctccagactc [gRNA]	70.00 %	80.41 °C	-	SacI	1 [det]		
53 - 75	ccggccgagctctccagactcgg [gRNA]	70.00 %	80.41 °C	-	SacI	1 [det]		
53 - 75	ccggccgagctctccagactccgg [gRNA]	70.00 %	80.27 °C	-	SacI	1 [det]		
56 - 78	ccggagctctccagactccggagg [gRNA]	70.00 %	80.27 °C	-	BspEI SacI	1 [det]	4 [detail]	793 [detail]
<b>56 - 78</b>	ccggagctctccagactccggagg [gRNA]	65.00 %	79.68 °C	-	BspEI SacI	1 [det]	1 [det]	1295 [detail]
59 - 81	agctctccagactccggaggagg [gRNA]	65.00 %	79.68 °C	-	BspEI	1 [det]	2 [det]	608 [detail]
65 - 87	ccaaactccggaggaggaaagcga [gRNA]	65.00 %	78.93 °C	-	BspEI	1 [det]	6 [det]	375 [detail]
72 - 94	ccggaggaggaaagcgatcgttac [gRNA]	55.00 %	73.00 °C	-	BsiEI PvuI	1 [det]	14 [det]	3180 [detail]
102 - 124	cctcggtcaaggcagaaggcagaaca [gRNA]	50.00 %	71.33 °C	-		0 [det]	3 [det]	588 [detail]
114 - 136	gaaggcagaacatctgaagaaatgg [gRNA]	40.00 %	65.99 °C	-		0 [det]	7 [det]	1117 [detail]
123 - 145	catctgaagaatggctgacgtgg [gRNA]	50.00 %	69.80 °C	-		1 [det]	2 [det]	1388 [detail]
151 - 173	caaccttattgtgtcgacaaatggc [gRNA]	45.00 %	69.15 °C	-		1 [det]	3 [det]	303 [detail]
<b>154 - 176</b>	cctattgtgtcgacaaatggcac [gRNA]	50.00 %	72.27 °C	-		1 [det]	1 [det]	667 [detail]
157 - 179	attgtgtcgacaaatggcacatgg [gRNA]	50.00 %	72.27 °C	-		1 [det]	2 [det]	499 [detail]
162 - 184	gtgcgacaaatggcactggaaatgg [gRNA]	55.00 %	74.69 °C	-		1 [det]	6 [det]	677 [detail]

Showing 1 to 20 of 71 entries

First Previous 1 2 3 4 Next Last

配列に関する情報  
制限酵素サイト も

# CRISPRdirect によるガイド RNA 設計

49

- Target sequences with TTTTs are also shown in gray. Avoid TTTTs in gRNA vectors with pol III promoter.

show **highly specific** target only

Show 20 entries

position		target sequence	sequence information				number of target sites <small>?</small>		
start - end	-	20mer+PAM (total 23mer)	GC% of 20mer	Tm of 20mer	TTTT in 20mer	restriction sites	20mer +PAM	12mer +PAM	8mer +PAM
7 - 29	-	cctcccggttcgcctcgctcgca [gRNA]	75.00 %	83.48 °C	-	NruI	1 [detail]	5 [detail]	370 [detail]
10 - 32						NruI	1 [detail]	1 [detail]	276 [detail]
11 - 32						NruI SacI	1 [detail]	4 [detail]	911 [detail]
18 - 40						NruI SacI	1 [detail]	4 [detail]	1072 [detail]
47 - 69						SacI	1 [detail]	305 [detail]	79033 [detail]
50 - 72						SacI	1 [detail]	37 [detail]	3555 [detail]
53 - 75						SacI	1 [detail]	3 [detail]	175 [detail]
53 - 75						SacI	1 [detail]	15 [detail]	2679 [detail]
56 - 78						BspEI SacI	1 [detail]	4 [detail]	793 [detail]
56 - 78						BspEI SacI	1 [detail]	1 [detail]	1295 [detail]
59 - 81						BspEI	1 [detail]	2 [detail]	608 [detail]
65 - 87						BspEI	1 [detail]	6 [detail]	375 [detail]
72 - 94	-	ccggaggagaagcgatcgttac [gRNA]	55.00 %	73.00 °C	-	BsiEI PvuI	1 [detail]	14 [detail]	3180 [detail]
102 - 124	-	cctcgtaaggcagaaggcagaaca [gRNA]	50.00 %	71.33 °C	-		0 [detail]	3 [detail]	588 [detail]
114 - 136	+	gaagcagaacatctgaagaa[tgg] [gRNA]	40.00 %	65.99 °C	-		0 [detail]	7 [detail]	1117 [detail]
123 - 145	+	catctgaagaatggctgacg[agg] [gRNA]	50.00 %	69.80 °C	-		1 [detail]	2 [detail]	1388 [detail]
151 - 173	+	caaccttattgtgtgcacaa[tgg] [gRNA]	45.00 %	69.15 °C	-		1 [detail]	3 [detail]	303 [detail]
154 - 176	-	cctatttgtgtgcacaaatggcac [gRNA]	50.00 %	72.27 °C	-		1 [detail]	1 [detail]	667 [detail]
157 - 179	+	atttgtgtgcacaaatggcac[tgg] [gRNA]	50.00 %	72.27 °C	-		1 [detail]	2 [detail]	499 [detail]
162 - 184	+	gtgcgacaaatggcactggaa[tgg] [gRNA]	55.00 %	74.69 °C	-		1 [detail]	6 [detail]	677 [detail]

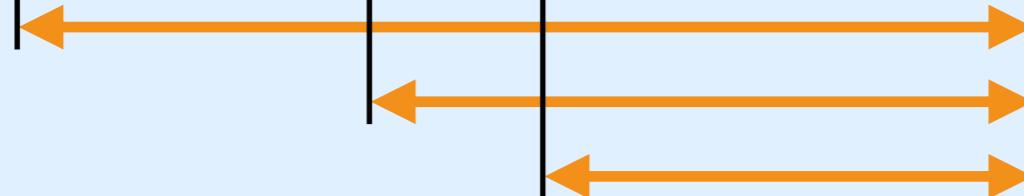
Showing 1 to 20 of 71 entries

First Previous 1 2 3 4 Next Last

ゲノムに何箇所あるか？

20mer PAM

NNNNNNNNNNNNNNNNNN NGG



# CRISPRdirect によるガイド RNA 設計

50

- Target sequences with TTTTs are also shown in gray. Avoid TTTTs in gRNA

show highly specific target only

Sequence		sequence information			off-target sites			
		GC% of 20mer	Tm of 20mer	TTTT in 20mer	SITES	+PAM	+PAM	8mer +PAM
-	-	cctcccggttcgcgcgtcgcgca [gRNA]	75.00 %	83.48 °C	-	NruI	1 [detail]	5 [detail]
<b>10 - 32</b>	-	cccgttgcgcgtcgcgagct [gRNA]	70.00 %	81.73 °C	-	NruI	1 [detail]	1 [detail]
11 - 33	-	ccgttgcgcgtcgcgagctc [gRNA]	70.00 %	81.09 °C	-	NruI SacI	1 [detail]	4 [detail]
18 - 40	-	cctcgctcgcgagctcggttac [gRNA]	70.00 %	80.24 °C	-	NruI SacI	1 [detail]	4 [detail]
47 - 69	-	ccggccggccgagctctccaga [gRNA]	75.00 %	84.25 °C	-	SacI	1 [detail]	305 [detail]
50 - 72	-	ccggccggagctctccagactc [gRNA]	70.00 %	80.41 °C	-	SacI	1 [detail]	37 [detail]
53 - 75	+	ccggccggagctctccagactc [gRNA]	70.00 %	80.41 °C	-	SacI	1 [detail]	3 [detail]
53 - 75	-	ccggagctctccagactccgg [gRNA]	70.00 %	80.27 °C	-	SacI	1 [detail]	1 [detail]
56 - 78	+	ccggagctctccagactccgg [gRNA]	70.00 %	80.27 °C	-	BspEI SacI	1 [detail]	1 [detail]
<b>56 - 78</b>	-	ccggagctctccagactccggagg [gRNA]	65.00 %	79.68 °C	-	BspEI SacI	1 [detail]	1 [detail]
59 - 81	+	agctctccagactccggagg [gRNA]	65.00 %	79.68 °C	-	BspEI	1 [detail]	1 [detail]
65 - 87	-	ccaaactccggaggaggaaagcga [gRNA]	65.00 %	78.93 °C	-	BspEI	1 [detail]	1 [detail]
72 - 94	-	ccggaggaggaaagcgatcgttac [gRNA]	55.00 %	73.00 °C	-	BsiEI PvuI	1 [detail]	14 [detail]
102 - 124	-	cctcggtcaaggcagaaggcagaaca [gRNA]	50.00 %	71.33 °C	-		0 [detail]	3 [detail]
114 - 136	+	gaaggcagaacatctgaagaa [gRNA]	40.00 %	65.99 °C	-		0 [detail]	7 [detail]
123 - 145	+	catctgaagaatggctgacg [gRNA]	50.00 %	69.80 °C	-		1 [detail]	2 [detail]
151 - 173	+	caaccttattgtgtcgacaa [gRNA]	45.00 %	69.15 °C	-		1 [detail]	3 [detail]
<b>154 - 176</b>	-	cctattgtgtcgacaaatggcac [gRNA]	50.00 %	72.27 °C	-		1 [detail]	1 [detail]
157 - 179	+	attgtgtgtcgacaaatggcac [gRNA]	50.00 %	72.27 °C	-		1 [detail]	2 [detail]
162 - 184	+	gtgcgacaaatggcactggaa [gRNA]	55.00 %	74.69 °C	-		1 [detail]	6 [detail]

Showing 1 to 20 of 71 entries

First Previous 1 2 3 4 Next Last

緑：おすすめ

オフターゲット  
サイトが少ない

オフターゲット  
サイトが多い

# CRISPRdirect によるガイド RNA 設計

51

- Target sequences with TTTTs are also shown in gray. Avoid TTTTs in gRNA vectors with pol III promoter.

show **highly specific** target only

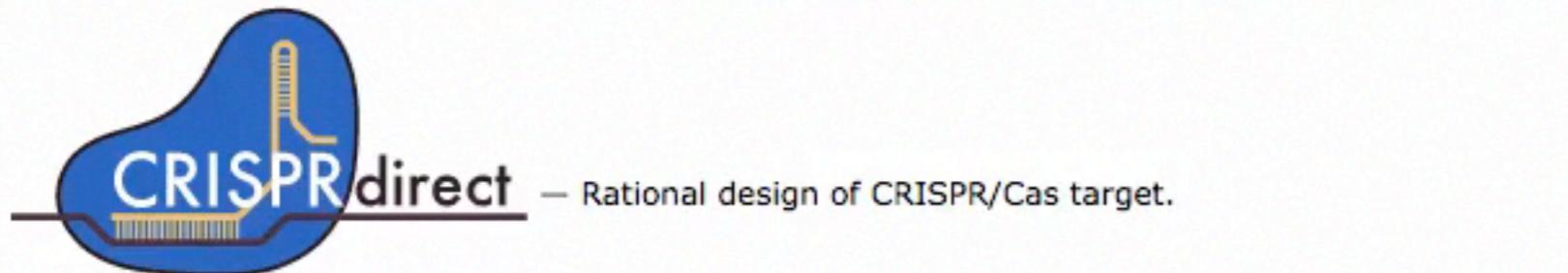
Show 20 entries

position		target sequence	sequence information				number of target sites ?		
start - end	-	20mer+PAM (total 23mer)	GC% of 20mer	Tm of 20mer	TTTT in 20mer	restriction sites	20mer +PAM	2mer +PAM	8mer +PAM
7 - 29	-	cctcccggttcgcgcctcgctcgcgaa [gRNA]	75.00 %	83.48 °C	-	NruI	1 [detail]	[detail]	370 [detail]
<b>10 - 32</b>	-	cccgtttcgccctcgctcgcgagct [gRNA]	70.00 %	81.73 °C	-	NruI	1 [detail]	1 [detail]	276 [detail]
11 - 33	-	ccgttccgcctcgctcgcgagctc [gRNA]	70.00 %	81.09 °C	-	NruI SacI	1 [detail]	4 [detail]	911 [detail]
18 - 40	-	cctcgctcgcgagctcgcttac [gRNA]	70.00 %	80.24 °C	-	NruI SacI	1 [detail]	4 [detail]	1072 [detail]
47 - 69	-	ccggccggccggagctctccaga [gRNA]	75.00 %	84.25 °C	-	SacI	1 [detail]	305 [detail]	79033 [detail]
50 - 72	-	ccggccggagctctccagactc [gRNA]	70.00 %	80.41 °C	-	SacI	1 [detail]	37 [detail]	3555 [detail]
53 - 75	+	ccggccggagctctccagactcgg [gRNA]	70.00 %	80.41 °C	-	SacI	1 [detail]	3 [detail]	175 [detail]
53 - 75	-	ccggccggagctctccagactccgg [gRNA]	70.00 %	80.27 °C	-	SacI	1 [detail]	15 [detail]	2679 [detail]
56 - 78	+	ccggccggagctctccagactccggagg [gRNA]	70.00 %	80.27 °C	-	BspEI SacI	1 [detail]	4 [detail]	793 [detail]
<b>56 - 78</b>	-	ccggccggagctctccagactccggagg [gRNA]	65.00 %	79.68 °C	-	BspEI SacI	1 [detail]	1 [detail]	1295 [detail]
59 - 81	+	agctctccagactccggaggagg [gRNA]	65.00 %	79.68 °C	-	BspEI	1 [detail]	2 [detail]	608 [detail]
65 - 87	-	ccggccggagggaggaggacgca [gRNA]	65.00 %	78.93 °C	-	BspEI	1 [detail]	6 [detail]	375 [detail]
72 - 94	-	ccggccggagggaggacgatcgttac [gRNA]	55.00 %	73.00 °C	-	BsiEI PvuI	1 [detail]	14 [detail]	3180 [detail]
102 - 124	-	cctcggtcaaggcagaaggcagaaca [gRNA]	50.00 %	71.33 °C	-		0 [detail]	3 [detail]	588 [detail]
114 - 136	+	gaaggcagaacatctgaagaaatgg [gRNA]	40.00 %	65.99 °C	-		0 [detail]	7 [detail]	1117 [detail]
123 - 145	+	catctgaagaatggctgacggagg [gRNA]	50.00 %	69.80 °C	-		1 [detail]	2 [detail]	1388 [detail]
151 - 173	+	caacccatttgtgtgcgacaaatggc [gRNA]	45.00 %	69.15 °C	-		1 [detail]	3 [detail]	303 [detail]
<b>154 - 176</b>	-	cctatttgtgtgcgacaaatggcac [gRNA]	50.00 %	72.27 °C	-		1 [detail]	1 [detail]	667 [detail]
157 - 179	+	atttgtgtgcgacaaatggcacatgg [gRNA]	50.00 %	72.27 °C	-		1 [detail]	2 [detail]	499 [detail]
162 - 184	+	gtgcgacaaatggcactggaaatgg [gRNA]	55.00 %	74.69 °C	-		1 [detail]	6 [detail]	677 [detail]

Showing 1 to 20 of 71 entries

First Previous 1 2 3 4 Next Last

Graphical View:



0 mismatch/gap | ≤1 mismatch/gap | ≤2 mismatches/gaps | more: 0 show

Search for:  both strand  plus strand  minus strand

2017-06-28 07:16:48, GGGenome : Rice genome, Os-Nipponbare-Reference-IRGSP-1.0 (Oct, 2011)

## Summary:

- CCNcccattcgccctcgactcgaa (0)
- tcagcqaaqcqaggacqaacgggNGG (1)
- **TOTAL (1)**

## Results:

Showing first 200 results for each strand of the query sequence.

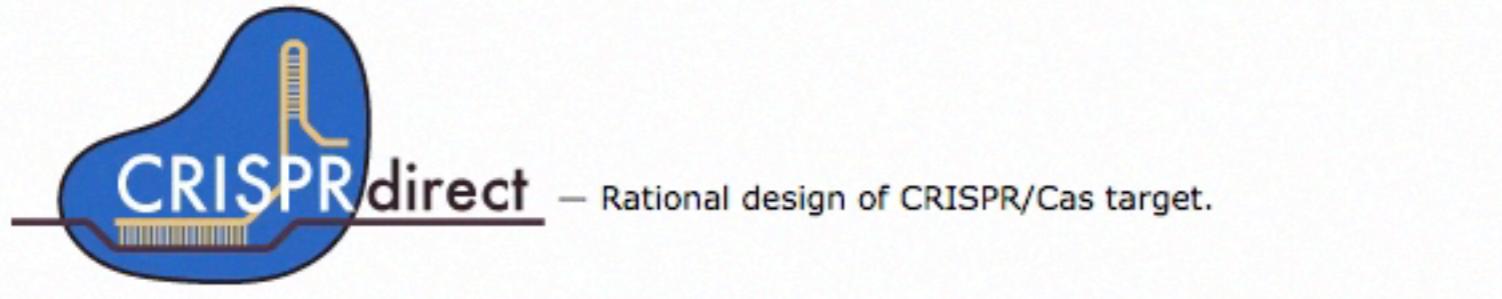
Matches are highlighted with blue background. **Mismatches** and **indels** are marked in red.

[chr05:21504153-21504175](#) ▼21504153  
TCGGCGGCCGGCGGTGTAAAGTAACGCGAGCTCGCGAGCGAGGGGAACGGGAGGTGCACTGCTGGTCTGGACGGACGAAGA

## Data Export:

Maximum 100000 results can be retrieved in various formats shown below:

- Tab-delimited text: [Open in new window](#) | [Download](#)  
You can copy-paste the result into spreadsheet softwares (e.g. Excel) or text editors.
- CSV format: [Open in new window](#) | [Download](#)  
You can open CSV files using spreadsheet softwares (e.g. Excel) or text editors.
- BED format: [Link](#) | [Download](#)
- GFF format: [Link](#) | [Download](#)



0 mismatch/gap | ≤1 mismatch/gap | ≤2 mismatches/gaps | more: 2 show

Search for:  both strand  plus strand  minus strand

2017-06-28 07:17:02, GGGenome : Rice genome, Os-Nipponbare-Reference-IRGSP-1.0 (Oct, 2011)

## Summary:

- CCNcccattcgcctcgactcgaa (1)
- tgcgcqaaqaaaggacgaaacgggNGG (4)
- **TOTAL (5)**

## Results:

Showing first 200 results for each strand of the query sequence.

Matches are highlighted with blue background. **Mismatches** and **indels** are marked in red.

chr12:23545769-23545793 ▼23545769  
CCAACGAGGCCAGGTGAGCTCTCCCCCTCCCCCGATTGCCCTCGCTTCGCGACCGAGTTCACTTGAGCTGGGTTGCACTGAC

chr05:21504153-21504175 ▼21504153  
TCGGCGGCCGGCGGTGTAAAGTAACGCGAGCTCGCGAGCGAGGCCAACGGGAGGTGCACTGCTGGTGGACGGACGAAGA

chr05:21504153-21504175 ▼21504153  
TCGGCGGCCGGCGGTGTAAAGTAACGCGAGCTCGCGAGCGAGGCCAACGGGAGGTGCACTGCTGGTGGACGGACGAAGA

chr05:21504153-21504175 ▼21504153  
TCGGCGGCCGGCGGTGTAAAGTAACGCGAGCTCGCGAGCGAGGCCAACGGGAGGTGCACTGCTGGTGGACGGACGAAGA

chr05:21504153-21504175 ▼21504153  
TCGGCGGCCGGCGGTGTAAAGTAACGCGAGCTCGCGAGCGAGGCCAACGGGAGGAGTGCACTGCTGGTGGACGGACGAAGA

# CRISPRdirect によるガイド RNA 設計

54

- Target sequences with TTTTs are also shown in gray. Avoid TTTTs in gRNA vectors with pol III promoter.

show **highly specific** target only

Show 20 entries Search:

position		target sequence	sequence information				number of target sites <small>?</small>		
start - end	-	20mer+PAM (total 23mer)	GC% of 20mer	Tm of 20mer	TTTT in 20mer	restriction sites	20mer +PAM	12mer +PAM	8mer +PAM
7 - 29	-	cctcccggttcgcctcgctcgcg [gRNA]	75.00 %	83.48 °C	-	NruI	1 [detail]	5 [detail]	370 [detail]
10 - 32	-	cccgtttcgccctcgctcgcgagct [gRNA]	70.00 %	81.73 °C	-	NruI	1 [detail]	1 [detail]	276 [detail]
11 - 33	-	ccgttccgcctcgctcgcgagctc [gRNA]	70.00 %	81.09 °C	-	NruI SacI	1 [detail]	4 [detail]	911 [detail]
18 - 40	-	cctcgctcgcgagctcgcttac [gRNA]	70.00 %	80.24 °C	-	NruI SacI	1 [detail]	4 [detail]	1072 [detail]
47 - 69	-	ccggccggccgagctctccaga [gRNA]	75.00 %	84.25 °C	-	SacI	1 [detail]	305 [detail]	79033 [detail]
50 - 72	-	ccggccggagctctccagactc [gRNA]	70.00 %	80.41 °C	-	SacI	1 [detail]	37 [detail]	3555 [detail]
53 - 75	+	ccggccgagctctccagactcgg [gRNA]	70.00 %	80.41 °C	-	SacI	1 [detail]	3 [detail]	175 [detail]
53 - 75	-	ccggccgagctctccagactccgg [gRNA]	70.00 %	80.27 °C	-	SacI	1 [detail]	15 [detail]	2679 [detail]
56 - 78	+	ccggccgagctctccagactccggagg [gRNA]	70.00 %	80.27 °C	-	BspEI SacI	1 [detail]	4 [detail]	793 [detail]
56 - 78	-	ccggccgagctctccagactccggagg [gRNA]	65.00 %	79.68 °C	-	BspEI SacI	1 [detail]	1 [detail]	1295 [detail]
59 - 81	+	agctctccagactccggaggagg [gRNA]	65.00 %	79.68 °C	-	BspEI	1 [detail]	2 [detail]	608 [detail]
65 - 87	-	ccaaactccggaggaggaaagcga [gRNA]	65.00 %	78.93 °C	-	BspEI	1 [detail]	6 [detail]	375 [detail]
72 - 94	-	ccggccgagctctccagactccggagg [gRNA]	55.00 %	73.00 °C	-	BsiEI PvuI	1 [detail]	14 [detail]	3180 [detail]
102 - 124	-	cctcggtcaaggcagaaggcagaaca [gRNA]	50.00 %	71.33 °C	-		0 [detail]	3 [detail]	588 [detail]
114 - 136	+	gaaggcagaacatctgaagaaatgg [gRNA]	40.00 %	65.99 °C	-		0 [detail]	7 [detail]	1117 [detail]
123 - 145	+	catctgaagaatggctgacgtgg [gRNA]	50.00 %	69.80 °C	-		1 [detail]	2 [detail]	1388 [detail]
151 - 173	+	caacccatttgtgtgcgacaaatgg [gRNA]	45.00 %	69.15 °C	-		1 [detail]	3 [detail]	303 [detail]
154 - 176	-	cctatttgtgtgcgacaaatggcac [gRNA]	50.00 %	72.27 °C	-		1 [detail]	1 [detail]	667 [detail]
157 - 179	+	atttgtgtgcgacaaatggcacatgg [gRNA]	50.00 %	72.27 °C	-		1 [detail]	2 [detail]	499 [detail]
162 - 184	+	gtgcgacaaatggcactggaaatgg [gRNA]	55.00 %	74.69 °C	-		1 [detail]	6 [detail]	677 [detail]

Showing 1 to 20 of 71 entries

First Previous 1 2 3 4 Next Last

Graphical View:

# CRISPRdirect によるガイド RNA 設計

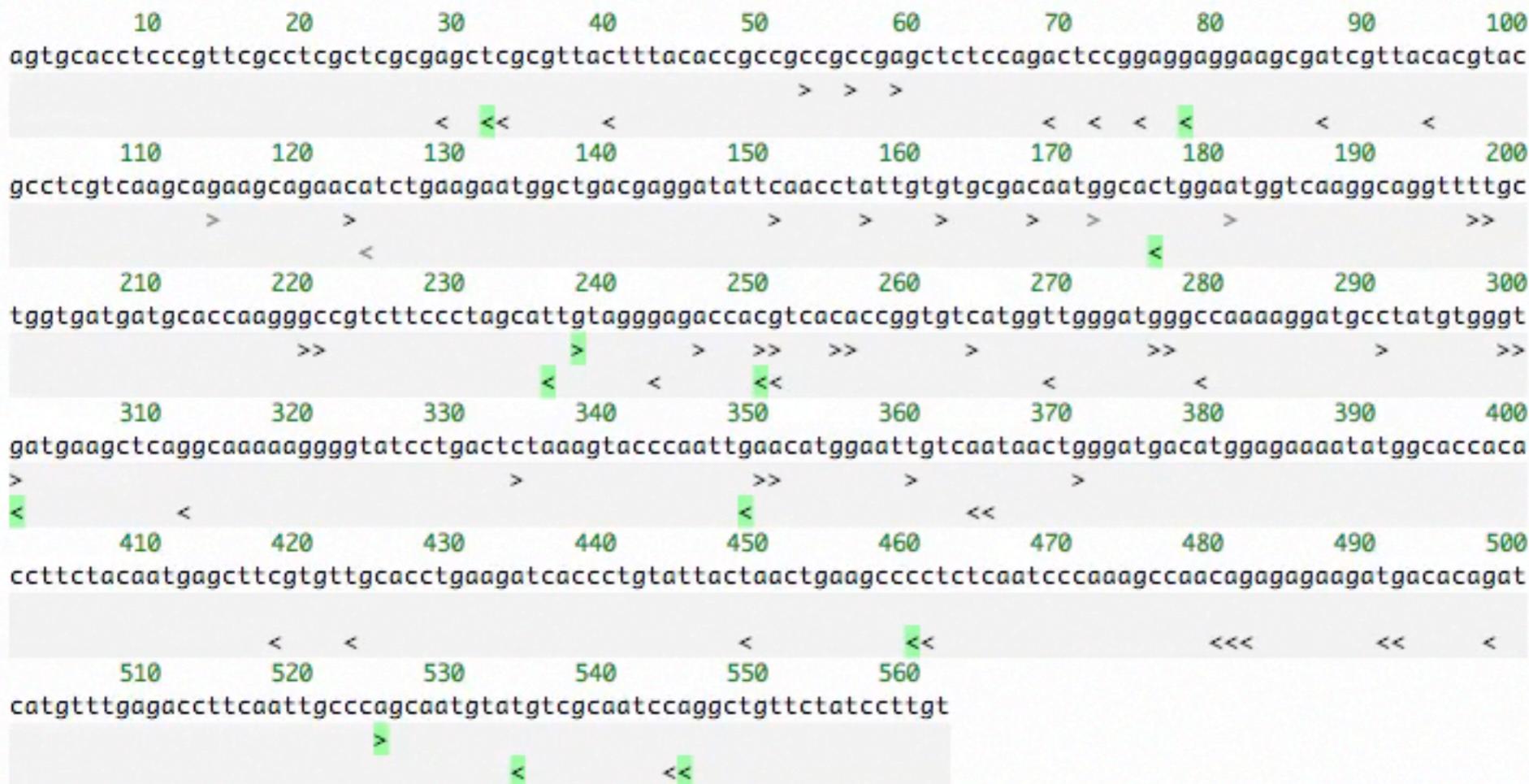
55

151 - 173	+	caacctattgtgtcgacaa[tgg] [gRNA]	45.00 %	69.15 °C	-		1 [detail]	3 [detail]	303 [detail]
154 - 176	-	cctattgtgtcgacaaatggcac [gRNA]	50.00 %	72.27 °C	-		1 [detail]	1 [detail]	667 [detail]
157 - 179	+	attgtgtcgacaaatggcac[tgg] [gRNA]	50.00 %	72.27 °C	-		1 [detail]	2 [detail]	499 [detail]
162 - 184	+	gtgcgacaaatggcactggaa[tgg] [gRNA]	55.00 %	74.69 °C	-		1 [detail]	6 [detail]	677 [detail]

Showing 1 to 20 of 71 entries

First Previous 1 2 3 4 Next Last

## Graphical View:



## Data Export:

- Tab-delimited text: [Open in new window](#) | [Download](#)
- JSON format: [Open in new window](#) | [Download](#)

Tab-delimited text can be copy-pasted into spreadsheet softwares (e.g. Excel) or text editors.

# CRISPRdirect によるガイド RNA 設計

56

510 < < 520 530 540 550 560 << <<< << <  
catgtttgagacaccttcaattgcccccaggcaatgttatgtcgcaatccaggctgttctatccttgt  
> < <<

### **Data Export:**

- Tab-delimited text: [Open in new window](#) | [Download](#)
  - JSON format: [Open in new window](#) | [Download](#)

Tab-delimited text can be copy-pasted into spreadsheet softwares (e.g. Excel) or text editors.

# CRISPRdirect によるガイド RNA 設計

57

# CRISPRdirect によるガイド RNA 設計

◆	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	# [ CRISPRdirect ]	2017-06-28 12:37:35 ]													
2	# sequence_name:														
3	# pam_sequence	NGG													
4	# specificity_db	Rice (Oryza sativa ssp. japonica) genome, Os-Nipponbare-Reference-IRGSP-1.0 (Oct, 2011)													
5	# start	end	strand	sequence	GC	Tm	TTTT	RE_sites	hit_20mer	hit_12mer	hit_8mer				
6	#														
7	7	29	-	cctcccggttcgcctcgctcgca	75	83.48	0	Nrul	1	5	370				
8	10	32	-	cccggttcgcctcgctcgagct	70	81.73	0	Nrul	1	1	276				
9	11	33	-	ccggttcgcctcgctcgagctc	70	81.09	0	Nrul,Sacl	1	4	911				
10	18	40	-	cctcgctcgcgagctcgcttac	70	80.24	0	Nrul,Sacl	1	4	1072				
11	47	69	-	ccggccggcccgagctctccaga	75	84.25	0	Sacl	1	305	79033				
12	50	72	-	ccggccggcgagctctccagactc	70	80.41	0	Sacl	1	37	3555				
13	53	75	+	ccggccgagctctccagactccgg	70	80.41	0	Sacl	1	3	175				
14	53	75	-	ccggccgagctctccagactccgg	70	80.27	0	Sacl	1	15	2679				
15	56	78	+	ccgagctctccagactccggagg	70	80.27	0	BspEI,Sacl	1	4	793				
16	56	78	-	ccgagctctccagactccggagg	65	79.68	0	BspEI,Sacl	1	1	1295				
17	59	81	+	agctctccagactccggaggagg	65	79.68	0	BspEI	1	2	608				
18	65	87	-	ccagactccggaggaggaaagcga	65	78.93	0	BspEI	1	6	375				
19	72	94	-	ccggaggaggaagcgatcgttac	55	73	0	BsiEI,PvuI	1	14	3180				
20	102	124	-	cctcgtaaggcagaaggcagaaca	50	71.33	0		0	3	588				
21	114	136	+	gaagcagaacatctgaagaatgg	40	65.99	0		0	7	1117				
22	123	145	+	catctgaagaatggctgacgagg	50	69.8	0		1	2	1388				
23	151	173	+	caacctattgtgtcgacaatgg	45	69.15	0		1	3	303				
24	154	176	-	cctattgtgtcgacaatggcac	50	72.27	0		1	1	667				
25	157	179	+	attgtgtcgacaatggcactgg	50	72.27	0		1	2	499				
26	162	184	+	gtgcgacaatggcactggaatgg	55	74.69	0		1	6	677				
27	168	190	+	caatggcactggaatggtaagg	50	72.31	0		5	8	545				
28	172	194	+	ggcactggaatggtaaggcagg	60	77.26	0		0	2	447				
29	181	203	+	atggtcaaggcagggtttgtgg	50	73.64	1		0	1	1511				
30	197	219	+	ttgctggtgatgtatgcaccaagg	50	73.43	0		1	4	553				
31	198	220	+	tgctggtgatgtatgcaccaagg	50	73.43	0		1	9	616				
32	214	236	-	ccaaggcccgtttcccttagcat	60	80.62	0		1	1	121				
33	220	242	+	gccgtttccctagcattgtagg	55	76.01	0		1	2	796				
34	221	243	+	ccgtttccctagcattgtagg	50	72.87	0		1	3	538				
35	221	243	-	ccgtttccctagcattgtagg	50	73.65	0		1	6	1777				

<http://crispr.dbcls.jp/>

- PAMに隣接し、特異性の高い20塩基を選択

ゲノムのほかの領域とは塩基配列が一致せず、標的とする部位とだけ完全一致する配列を選択。特にPAM近傍のseed(12塩基または8塩基) の特異性を重視。

- 大量設計にも対応

ヘルプページにてスクリプトを公開

- 多くの生物種に対応

350種の生物種に対応 (GGGenomeと同じ)

- 無償で自由に利用可能

商用利用であっても無償。

利用記録は当センターにおいて機密扱い。



ゲノムが公開されていれば対応します。  
ゲノム配列の所在、生物種に関する情報  
をご連絡ください（項目はヘルプページ  
の一覧を参照）

## 検索可能なデータベース一覧(ゲノム)



Database	Source	Group	Species	生物種	Latin name	Assembly information
ASM34733v1	EnsemblPlants	Plant	Tausch's goatgrass	タルホコムギ	Aegilops tauschii	<a href="#">ASM34733v1 (Dec, 2013)</a>
AMTR1.0	EnsemblPlants	Plant	A. trichopoda	アムボレラ・トリコポダ	Amborella trichopoda	<a href="#">AMTR1.0 (Jan, 2014)</a>
Araly v.1.0	EnsemblPlants	Plant	Lyre-leaved rock-cress	シロイヌナズナ属	Arabidopsis lyrata	<a href="#">v.1.0 (Dec, 2008)</a>
TAIR10_en	EnsemblPlants	Plant	Thale cress	シロイヌナズナ	Arabidopsis thaliana	<a href="#">TAIR10 (Sep, 2010)</a>
Bradi v1.0	EnsemblPlants	Plant	Purple false brome	セイヨウヤマカモジ, ミナトカモジグサ	Brachypodium distachyon	<a href="#">v1.0 (Jan, 2009)</a>
Braol v2.1	EnsemblPlants	Plant	Wild cabbage	ヤセイカンラン, ワイルドキャベツ	Brassica oleracea	<a href="#">v2.1</a>
IVFCAASv1	EnsemblPlants	Plant	Chinese cabbage	ハクサイ, サントウサイ	Brassica rapa ssp. pekinensis	<a href="#">IVFCAASv1 (Aug, 2009)</a>
Chlre v3.1	EnsemblPlants	Plant	Green algae	クラミドモナス	Chlamydomonas reinhardtii	<a href="#">v3.1 (Nov, 2007)</a>
ASM9120v1	EnsemblPlants	Plant	Red alga	シアニティオシゾン	Cyanidioschyzon merolae	<a href="#">ASM9120v1 (Nov, 2008)</a>
Soybn V1.0	EnsemblPlants	Plant	Soybean	ダイズ	Glycine max	<a href="#">V1.0 (Jan, 2010)</a>



生命科学研究を支える  
便利なデータベースやツール  
約100個を簡潔に紹介

内藤雄樹／編 実験医学増刊号（羊土社）  
今日から使える！ データベース・ウェブ  
ツール 達人になるための実践ガイド100

本日はありがとうございました