

第43回日本分子生物学会年会  
生命科学のデータベース活用法2020

# 植物ゲノム情報統合ポータルサイト Plant GARDENの構築

かずさDNA研究所・東京大学・筑波大学

発表者：かずさDNA研究所・磯部祥子



Licensed under a Creative Commons表示4.0国際ライセンス  
(c) 2020 かずさDNA研究所

統合化推進プログラムにおいて  
2種類のDBを開発

<http://pgdbj.jp/>



H23~28 年度に開発

格納データ  
マーカー  
オルソログ (遺伝子配列の配列類似性情報)

リソース

オルソログDB

多種多様な植物ゲノム関連情報

様々な目的をもった幅広いユーザー

DNAマーカー他

育種向けDNAマーカー検索



<https://plantgarden.jp>

Plant GARDEN

H29年度から開発開始

格納データ  
ゲノム・遺伝子配列  
マーカー/QTL/連鎖地図  
SNPs

Plant GARDEN

Plant GARDEN (Genome And Resource Database Entry) は  
さまざまな種類のゲノムやマーカー情報を集めたポータルサイトです

検索

ゲノム配列  
マーカー  
QTL  
連鎖地図  
SNPs

GARDEN = Genome And Resource Database ENtry

覚えやすさを意識して命名

<https://plantgarden.jp>

B版公開：2019年3月

正規版公開：2020年7月

英語版

日本語版

The screenshot shows the Japanese version of the Plant GARDEN website. At the top, there is a logo with a green plant icon and the text "Plant GARDEN". Below the logo, it says "Plant GARDEN (Genome And Resource Database ENtry) はさまざまな植物のゲノムやマーカー情報を集めたポータルサイトです". There are language selection buttons for "JA" and "EN". A search bar with the placeholder "キーワード検索" and a "検索" button is present. Below the search bar is a grid of icons representing various plant species and functions. A "NEWS" section at the bottom lists updates from July 2020 to October 2019. At the very bottom, there are logos for PGDB and social media links for LinkedIn and Twitter.

The screenshot shows the English version of the Plant GARDEN website. At the top, there is a logo with a green plant icon and the text "Plant GARDEN". Below the logo, it says "Plant GARDEN (Genome And Resource Database ENtry) is a portal site that curates genome and marker information of various plant species." There are language selection buttons for "JA" and "EN". A search bar with the placeholder "Keyword search" and a "Search" button is present. Below the search bar is a grid of icons representing various plant species and functions. A "NEWS" section at the bottom lists updates from September 2020 to July 2020. At the very bottom, there are logos for PGDB and social media links for LinkedIn and Twitter.



- 「行うべきアクション」が分かりやすい項目表示  
4つのメニュー表示  
シンプルなデザイン・平易な言葉
- 想定するユーザー  
生物研究者/育種関係者 + 教育機関
- 高精度なゲノム配列が公開されている植物種は可能な限り網羅する方針

格納されている主なコンテンツ  
2020年12月現在

	植物種数	件数
ゲノム	119	133
遺伝子	119	5,890,957
DNAマーカー	34	287,703
形質と関連しているマーカー (QTL)	24	8,217
SNPs (SRAから検出した変異)	61	3,812

## 植物種からさがす

トップページの「植物種からさがす」または「すべてをみる」ボタンをクリック

Plant GARDEN ( Genome And Resource Database Entry ) は  
さまざまな植物のゲノムやマーカー情報を集めたポータルサイトです

② 使い方 ▶ ① このサイトについて ▶

キーワード検索 検索

植物種からさがす

キクタンゴク ブドウ欧州種 ミヤコサ シロイヌナズナ ダイコン イブゴ ラッカセイ ダイズ トマト すべてをみる ▶

他の方法でさがす

遺伝子から 配列から

BLAST

データ一覧

マーカー 形質

解析してみよう

SNP Detection

NEWS

- 23. Jul. 2020 PlantGARDEN 正規版を公開しました。▶
- 11. Dec. 2019 URIの更新を行いました。▶
- 15. Oct. 2019 ゲノム配列切り出し不具合について修正いたしました。▶

SEE ALL ▶

PGDB  
植物のオルソログ・マーカー情報 ▶

リンク集 ▶

Twitter ▶

## 45科に属する119植物種のリスト

よく検索される植物種

キクタンゴク ブドウ欧州種 ミヤコサ シロイヌナズナ ダイコン ラッカセイ ダイズ トマト


全植物種リスト


植物種名を入力してさがす 漢文字で表示 学名 和名

学名 ↓	和名 ↓	科名 ↓
Acer yangbiense	カエデ	ムクロシ科
Actinidia chinensis	キウイフルーツ (chinensis種), ゴールドキウイ	マタタビ科
Actinidia eriantha	キウイフルーツ (eriantha種)	マタタビ科
Aegilops tauschii	タルホコムギ	イネ科
Amaranthus hypochondriacus	アマランサス	ヒユ科
Ananas comosus	パイナップル	パイナップル科
Antirrhinum majus	キンギョソウ	オオバコ科
Apium graveolens	セロリ	セリ科
Arabidopsis thaliana	シロイヌナズナ	アブラナ科
Arabis alpina	ヤマハシザオ	アブラナ科
Arachis duranensis	ラッカセイ野生種 (duranensis種)	マメ科

# 植物種別ページ

https://plantgarden.jp/ja/list/species

 *Lotus japonicus*



この種について

科名: Fabaceae  
マメ科


属名: Lotus  
ミヤコグサ属

学名: Lotus japonicus  
和名: ミヤコグサ

Taxonomy ID: 34305

---


**ゲノム配列をみる**

 [すべて](#)

- ▶ [Lj2.5\(Miyakojima MG-20\)](#)
- ▶ [Lj3.0\(Miyakojima MG-20\)](#)

---


**遺伝子をさがす**


 [すべて](#)


- ▶ [Lj2.5\(Miyakojima MG-20\)](#)
- ▶ [Lj3.0\(Miyakojima MG-20\)](#)


---


**マーカーをさがす**

 [全て \(1155\)](#)

 [SSR \(1073\)](#)


 [CAPS \(52\)](#)


 [SCAR \(0\)](#)


 [その他 \(0\)](#)


---


**形質からさがす**


 [全て \(240\)](#)


 [ストロース別 \(28\)](#)

 [花性 \(4\)](#)

 [収量 \(22\)](#)

 [形態・成長 \(70\)](#)


 [島路・成分 \(86\)](#)


 [その他 \(36\)](#)


[さがす](#)

---

**その他の検索**

 [BLAST](#)  
配列からさがす





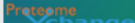
 [SNPs](#)  
変異をさがす

 [連鎖地図](#)

---

**リンク**


他の種類のデータ


    

メタボロミクス

タンパク質


種に関連するDB

 [miyakogusa.jp](#)

 [LIS - Legume Information System](#)  
Information about legume herbs for crop improvement

- ゲノム・遺伝子・SNPsは公開されている配列をキュレーション・解析してPlant GARDENに格納
- アセンブルされたゲノム配列はバージョン毎に表示
- DNAマーカー、形質と連関しているマーカー、連鎖地図はPGDBjでキュレーションしたコンテンツを移行中
- 各種に対応するメタボロミクス・タンパク質DBへのリンク
- 各種で整備されている代表的なDBのリンクを整備中

# アセンブルされたゲノム配列の情報

 *Brassica oleracea*

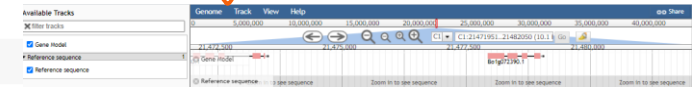
配列名 : Brassica\_oleracea.BOL

↓  
DOWNLOAD
↗  
JBrowse

**Index of /ja/download/t3712/t3712.G001**

- Parent Directory
- Brassica\_oleracea.BOL\_42.gff.gz
- Brassica\_oleracea.BOL\_cds.all.fa.gz
- Brassica\_oleracea.BOL\_cds.all.fa.gz
- Brassica\_oleracea.BOL\_dna.toplevel.fa
- Brassica\_oleracea.BOL\_dna.toplevel.fa.gz
- Brassica\_oleracea.BOL\_dna.rm.toplevel.fa.gz
- Brassica\_oleracea.BOL\_dna.rm.toplevel.fa.gz
- Brassica\_oleracea.BOL\_dna.rm.toplevel.fa.gz
- Brassica\_oleracea.BOL\_rnasec.all.fa.gz
- Brassica\_oleracea.BOL\_rnasec.all.fa.gz
- Brassica\_oleracea.11712.G001\_sam.v1.0.gz

ゲノム配列の詳細			
配列名	Brassica_oleracea.BOL	系統名	TO1000DH3, kale-like type
配列数	32,928	染色体数	2n=18
配列長 (bp)	488,622,507	N50長 (bp)	48,366,697
シーケンシングの方法	Illumina, 454	取得した配列量	94x
アセンブリ方法	SOAPdenovo v. 1.05	推定ゲノムサイズ (Mb)	648
シーケンシングの方法のコメント		コメント	Kale-like type. A doubled haploid line TO1000DH3 derived from <i>B. oleracea</i> cultivar TO1434



論文 (DOIコード) [10.1186/gb-2014-15-6-r77](https://doi.org/10.1186/gb-2014-15-6-r77)

データソース名 EnsemblePlants

遺伝子をさがす

キーワードを入力して検索

Agri-Food  
orsite Evry  
(National  
tml

ゲノム配列を切り出す

染色体 / Scaffolds

開始位置

終了位置


相補銀を出力する

詳細情報

配列の切り出し



# 遺伝子配列の検索

 *Brassica oleracea*



この種について

科名: Brassicaceae  
アブラナ科

属名: Brassica  
アブラナ属

学名: Brassica oleracea

和名: ヤセイカンラン, キャベツ, テリメンキャベツ, ブロッコリー, ケール

Taxonomy 1

ゲノム配列をみる



すべて

- ▶ Brassica\_oleracea.BOL(TO1000DH3, kale-like type)
- ▶ B.oleracea\_v1.0.chromosomes(var. capitata, 02-12)

遺伝子をさがす



すべて

- ▶ Brassica\_oleracea.BOL(TO1000DH3, kale-like type)
- ▶ B.oleracea\_v1.0.chromosomes(var. capitata, 02-12)

アセンブルされたゲノム上に推定された遺伝子情報を格納→ゲノムのバージョン別に整理

マーカーをさがす



全て  
(1987)



SSR  
(1040)



CAPS  
(282)



SCAR  
(26)



その他  
(639)

形質からさがす



全て  
(267)



ストレス耐性  
(65)



種子  
(0)



収穫  
(1)



形態・成長  
(51)



品質・成分  
(29)



その他  
(118)

さがす

その他の検索



BLAST  
配列からさがす

SNPs

変異をさがす

## 遺伝子キーワード検索

遺伝子をさがす



- ▶ Lj2.5(Miyakojima MG-20)
- ▶ Lj3.0(Miyakojima MG-20)

例えばゲノムのバージョンを選ぶと該当するゲノム内で絞り込んだキーワード検索ができる

キーワード検索

▶ 事前絞り込み条件  植物種名 Lotus japonicus (ID:t34305)  配列名 Lj2.5 (ID:t34305.G001)

Lj2.5

リセット 検索

Total (36468 Hit(s))

Excel CSV JSON

URL	Title
<a href="/ja/list/t34305/genome/t34305.G001/chr1.CM0133.450.r2.m">/ja/list/t34305/genome/t34305.G001/chr1.CM0133.450.r2.m</a>	chr1.CM0133.450.r2.m   Lotus japonicus
<a href="/ja/list/t34305/genome/t34305.G001/chr5.CM0158.410.r2.d">/ja/list/t34305/genome/t34305.G001/chr5.CM0158.410.r2.d</a>	chr5.CM0158.410.r2.d   Lotus japonicus
<a href="/ja/list/t34305/genome/t34305.G001/LjSGA_081454.0.1">/ja/list/t34305/genome/t34305.G001/LjSGA_081454.0.1</a>	LjSGA_081454.0.1   Lotus japonicus
<a href="/ja/list/t34305/genome/t34305.G001/chr1.CM0166.150.r2.m">/ja/list/t34305/genome/t34305.G001/chr1.CM0166.150.r2.m</a>	chr1.CM0166.150.r2.m   Lotus japonicus

# 遺伝子のページ

*Brassica oleracea*

配列名: B.oleracea\_v1.0.chromosomes



遺伝子配列の詳細			
物理上の位置	遺伝子名	CDS	Bol022761
		PEP	Bol022761
	ゲノム配列上の位置	Chromosome	C07
		Start Position	7226852
		End Position	7230434
Strand	+		
アノテーション情報 (Hayai Annotation ZEN)	KusakiDB	KusakiDB ID	KUNI00411397
		Transcript or Protein existence	YES
	OrthoDB	Protein name	Serine/threonine-protein kinase
		OrthoDB unique gene ID	3711_0:0046d6
		OG unique ID	95186at33090
	UniProt	Protein name	Protein kinase domain-containing protein
		Accession	A0A0D3CM56
		Comment	NA
		EC	NA
		KO ID	NA
KO Description	NA		

遺伝子アノテーションはDB内の各遺伝子のアノテーション項目とアノテーション基準を統一させるために、オリジナルのアノテーションではなく、独自に開発したアノテーションツールにより再アノテーションしています。

Hayai-Annotation  
Ghelfi et al. (2019) Bioinformatics 35:4427-4429

PFAM	KD Description	duar-specificity kinase
	PFAM ID	PF00069
InterPro	InterPro Description	<a href="#">Details</a>
	InterPro ID	IPR008271,IPR000719,IPR011009
Protein Database (NCBI)	InterPro Description	<a href="#">Details</a>
	Protein name	Uncharacterized protein
	Accession	NA
Gene Ontology	Database source	UniProt
	GD Description	<a href="#">Details</a>
	Database source	UniProt
Best hit taxonomy	Species	<i>Brassica oleracea</i> var. <i>oleracea</i>
	Family	Brassicaceae
	Phylum	Streptophyta
	Superkingdom	Eukaryota
Comments		

## Annotation Details



ID	Group/Domain	Name	LINK
UniProt Annotation (2 items)			
A0A0D3CM56		Protein kinase domain-containing protein	UniProt
IPR00069		duar-specificity kinase	IPR00069
OrthoDB (1 item)			
3711_0:0046d6	61202at33090	Serine/threonine-protein kinase, active site	OrthoDB (3)
InterPro (1 item)			
IPR00069		Kinase	IPR00069
InterPro (2 items)			
IPR008271		Ser/Thr_kinase_A0	InterPro (3)
IPR000719		Prog_kinase_dom	InterPro (3)
IPR011009		Kinase-like_dom_sf	InterPro (3)

## 遺伝子機能からの詮索

遺伝子をさがす



- ▶ Brassica\_oleracea.BOL(TO1000DH3, kale-like type)
- ▶ B.oleracea\_v1.0.chromosomes(var. capitata, 02-12)

遺伝子機能 (GO, KEGG) からの検索も可能  
(Hayai-annotationによるアノテーション結果に基づきます)

詳細検索

Gene Ontology

KO(KEGG Ontology)

ゲノム配列名

abaxial cell fate specification:GO:0010158(2) ✕ abscisic acid binding:GO:0010427(3) ✕ ここに選択したアイテムが出ます

abaxial cell fate specification:GO:0010158(2)

abscisic acid binding:GO:0010427(3)

abscisic acid biosynthetic process:GO:0009688(4)

abscisic acid glucosyltransferase activity:GO:0010294(1)

abscisic acid homeostasis:GO:1902265(2)

abscisic acid metabolic process:GO:0009687(1)

()内の数字はヒットした遺伝子数

# SNPの検索

*Brassica oleracea*



この種について

科目: Brassicaceae  
アブラナ科

属名: Brassica  
アブラナ属

学名: Brassica oleracea

和名: ヤセイカンラン, キャベツ, チリメンキャベツ, ブロッコリー, ケール

Taxonomy 1

ゲノム配列をみる



すべて

- ▶ Brassica\_oleracea.BOL(TO1000DH3, kale-like type)
- ▶ B.oleracea\_v1.0.chromosomes(var. capitata, 02-12)

遺伝子をさがす



すべて

- ▶ Brassica\_oleracea.BOL(TO1000DH3, kale-like type)
- ▶ B.oleracea\_v1.0.chromosomes(var. capitata, 02-12)

マーカーをさがす



全て  
(1987)



SSR  
(1940)



CAPS  
(282)



SCAR  
(26)



その他  
(639)

形質からさがす



全て  
(267)



ストレス耐性  
(65)



硬性  
(0)



収穫  
(1)



病害・疫病  
(54)



品質・成分  
(29)



その他  
(118)

さがす

その他の検索



配列からさがす



変異をさがす

*Brassica oleracea*

ID	Population Name	Line List	Variant Type	Reference Sequence	Sequence Type	Variant Counts	Data Source	Download	JBrowse
13712.SNP000001	Brassica_oleracea_SRA_100_lines_20191119		SNPs, In/dels	B.oleracea_v1.0.chromosomes	WGS	26,965,708	Plant GARDE		

*Brassica oleracea*

Bio Sample	Sequence Name	Accession	JBrowse	gVCF	Comment
Badger Inbred	SRA476457	SRX2184910			Cabbage
BOL909	SRA476457	SRX2184913			Cauliflower
HRIGRU009617	SRA476457	SRX2184912			Kohlrabi
QNF11	SRA923688	SRX6461273			var. botrytis
QNF11	SRA923688	SRX6461271			var. botrytis
QNF11	SRA923688	SRX6461270			var. botrytis
QNF11	SRA923688	SRX6461272			var. botrytis
TO1000	SRA150686	SRXS07314			var. oleracea
TO1000	SRA150686	SRXS07224			var. oleracea
TO1000	SRA150686	SRXS07227			var. oleracea
TO1000	SRA150686	SRXS07225			var. oleracea

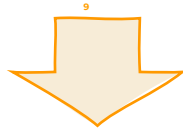
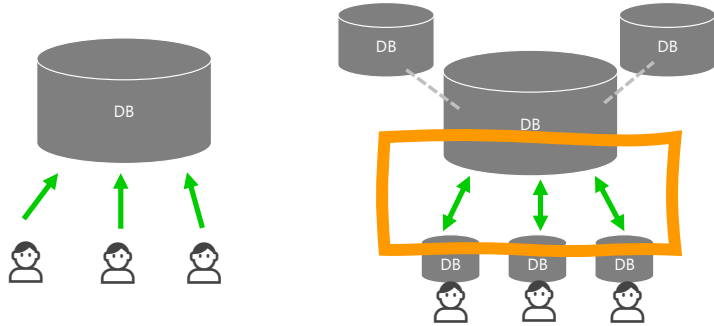
公開されたSRAからReferenceに対するVariantを検出→  
個別のgvcf + まとめたvcfファイルを提供



# Plant GARDEN

DBをとりまく環境

ユーザー自身が大規模なデータを持つ時代  
百科事典→DB間ネットワークのハブ



## カスタム型データ解析システム



Plant GARDEN内のデータを参照データとし、  
ユーザーのデータとできるだけシームレスに比較  
できるツールの開発

[? 使い方](#) [i このサイトに](#)

 植物種からさがす	 キクタニキク	 ブドウ欧州種	 ミヤコグサ	 シロイヌナズナ	 ダイコン
 他の方法でさがす	 遺伝子から	 BLAST 配列から			
 解析してみよう	 SNP Detection				

# 現在提供中のツール









## 解析ツールを選ぶ

-  お持ちのデータをアップロードしてブラウザ上で解析する
-  解析ツールをダウンロードしてローカル環境で解析する

アカウント登録をして  
ログインします

配列解析システム 1.03	
アカウント	sisobe
パスワード	.....
ログイン	
<a href="#">●アカウントの新規登録はこちらをクリック</a>	

パスワードを忘れてしまった場合はplantgarden-at-kazusa.or.jpまでお問い合わせ下さい。(データはリセットされます)

SNP Detection	SNPやIn/Delを検出する解析パイプライン		
Hayai-Annotation	遺伝子配列のアノテーション		
Hayai-Annotation Plants v2.0	植物種に特化した機能アノテーションシステム		
KusakiDB v1.0	タンパク質同族体群の検証と完全性のための新しいアプローチ		

Launching Hayai\_annotation\_v2.0

ユーザーごとにDocker containerを自動で作成

### Hayai-Annotation Plants v2.0 - Functional Annotation System Specialized in Plant Species

Database: KusakiDB v1.0  
Reference: Hayai-Annotation Plants

Hayai Annotation Search Proteins

Perform Full Annotation

Query Sequence Type  
 Protein  
 DNA

Minimum Sequence Identity (%)  
25 50 75 100

Minimum Query Coverage (%)  
25 50 75 100

Maximum E-value  
1 1e-05 1e-10 1e-15 1e-20

Minimum Target Coverage (%)  
25 50 75 100

ユーザーの皆様へ

<https://plantgarden.jp>

## Plant GARDEN

- Plant GARDENは、発展途上のDBです。不具合、もしくは機能不足が生じてご迷惑をおかけすることもあると思いますが、お使いになってみて問題点があれば、ぜひお知らせください。
- 格納するデータについては、できるだけユーザーの意見を取り入れたいと思います。格納をご希望するデータがございましたらお知らせください。優先的に格納してまいります。

ご意見・ご要望をお待ちしています  
[plantgarden@Kazusa.or.jp](mailto:plantgarden@Kazusa.or.jp)