

植物ゲノム情報統合ポータルサイトPlant GARDENの構築

平川英樹¹⁾、原田大士朗¹⁾、市原寿子²⁾、Ghelfi Andrea¹⁾、Fawcett Jeffrey¹⁾、白澤沙知子¹⁾、杉原英志³⁾、田村卓郎³⁾、中村保一¹⁾、中谷明弘²⁾、磯部祥子¹⁾、田畑哲之¹⁾

- 1) かずさDNA 研究所
- 2) 大阪大学大学院医学系研究科
- 3) 筑波大学プレシジョン・メディスン開発研究センター

植物ゲノム統合化データベースの構築



2011年～
<http://pgdbj.jp>

必要情報を閲覧して取得
オルソログ、DNAマーカー、リソース

次世代シーケンサー(NGS)の普及

新たな植物種で全ゲノム情報が次々と公開

品種間や変異株がもつゲノムワイドな多型情報を迅速・安価に収集

異種間でのゲノム情報の比較

増え続けるゲノム情報の統合

品種(系統)などの個体ゲノム解析による多型・ハプロタイプ検出



2018年～

NGSに対応
ユーザ所有のデータを解析
ツールの提供



Plant GARDEN

<https://plantgarden.jp>

ゲノム、トランスクリプトーム、プロテオーム、メタボローム、マーカー、変異、形質など。
国内外の様々な植物ゲノム関連データベースを統合し、横断検索できる。

 植物種からさがす	 オランダイチゴ	 キク	 シロイヌナズナ	 ダイコン
	 トマト	 ミヤコグサ	 ヨーロッパブドウ	 ラッカセイ ナンキンマメ
 他の方法でさがす	 形質から	 配列から	 キーワード検索	
	 解析してみよう	 SNP Detection	 その他の解析ツール	
 データ一覧	 全情報	 マーカー	 形質	

ゲノム解読された約400種を対象
(Pseudomolecule: 約90種)
100種を公開

リファレンス配列
文献からのキュレーション情報

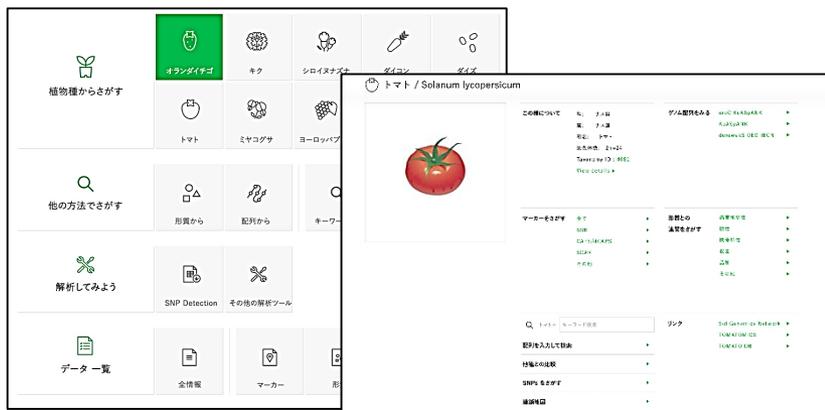
- ・連鎖地図
- ・DNAマーカー
- ・QTLなど

ツール

- ・SNP解析システム
- ・高速アノテーションシステム

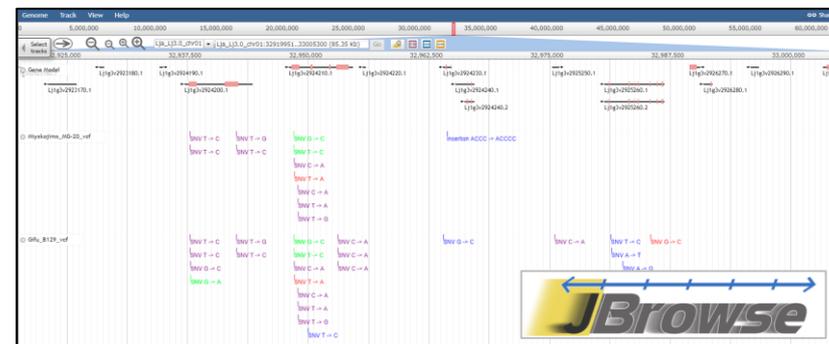
ユーザがより有益なゲノム情報を効率よく取得できるようにする

① Plant GARDENの構築



The interface includes a navigation menu with options like '植物種からさがす' (Search by species), '他の方法でさがす' (Search by other methods), '解析してみよう' (Analyze), and 'データ一覧' (Data list). The main view shows the tomato genome with various annotations and a search bar.

② ゲノム情報のブラウザへの集約



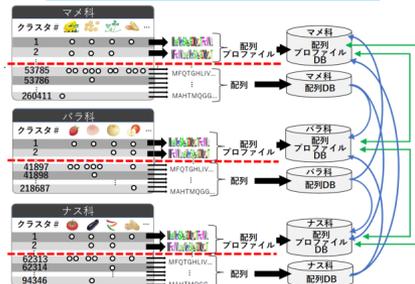
The JBrowse interface displays a genomic track with various annotations, including gene models and sequence alignments. The JBrowse logo is visible in the bottom right corner.

③ 種を超えた植物ゲノム情報統合のためのデータリンクの構築

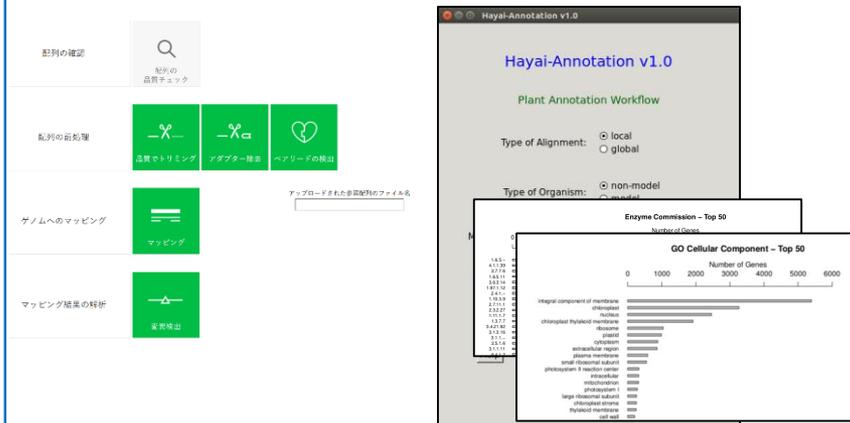
全植物種でクラスタリング



科でクラスタリング



④ 解析ツール群の開発・提供



The Hayai-Annotation v1.0 interface shows the plant annotation workflow, including options for alignment type (local/global) and organism type (non-model). The resulting GO Cellular Component analysis is displayed as a bar chart, showing the number of genes for various cellular components.

① Plant GARDENの構築 (植物のゲノム関連情報ページ)

ミヤコグサ / *Lotus japonicus*

ゲノム配列をみる

一覧

ゲノム配列の閲覧
(Lj2.5、Lj3.0)
バージョン管理



この種について

基本情報

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

俗名: ミヤコグサ, エポングサ

Lj2.5

Lj3.0

キーワード検索

Q ミヤコグサ+

キーワード検索

リンク

他の種類のデータ



メタプロテオミクス

タンパク質

Taxonomy ID: 34305

種に関連するDB



他DBへのリンク

マーカーをさがす

DNAマーカー検索

検索 (配列、変異、連鎖地図など)

Navigation buttons for DNA marker search: すべて (All), SSR, CAPS, SCAR, その他 (Others).

Navigation buttons for search: BLAST (配列からさがす), SNPs (変異をさがす), 連鎖地図 (Linkage map).

形質との関連についての検索

Trait-related search filters: 全て (199), ストレス耐性 (0), 稔性 (0), 収量 (0), 形態・成長 (73), 品質・成分 (92), その他 (34).

旧サイトへのリンク



植物のオルソログ・マーカー情報



リンク集

リンク集

検索

解析ツール

データ一覧

100種を公開

公開ゲノム情報

配列名 : Lj3.0

ゲノム配列の詳細

学名	<i>Lotus japonicus</i>
配列名	Lj3.0
系統名	MG-20
アセンブリレベル	Chromosome
配列数	9
染色体	
配列	
N50	
シー	
シート	
取得した配列量	35x
アセンブリ方法	Paracel Genome Assembler
推定ゲノムサイズ (Mb)	465
登録者	Kazusa DNA Research Institute

ゲノム配列の詳細

- ・学名、配列バージョン、染色体数
- ・シーケンシング、アセンブリ方法
- ・推定ゲノムサイズ、論文など

遺伝子をさがす

遺伝子名を入力してください

例を入力

検索

リセット

- ・遺伝子をさがす
- ・遺伝子名の入力

ゲノム配列を切り出す

染色体 / Scaffolds

Lj3.0_chr0

開始位置

1

終了位置

100

相補鎖を出力する

配列を得る

リセット

ゲノム配列を切り出す

- ・染色体名
- ・開始位置、終始位置
- ・相補鎖 (on/off)

その他の配列

配列の種類	ファイル名	データダウンロード	コメント
遺伝子 (cds / gene)	Lj3.0_cds.ffn.gz		
ゲノム	Lj3.0_pseudomol.fna.gz		
アノテーション (gff / gff3)	Lj3.0_gene_models2.gff3.gz		
Hayai Annotation ZEN	hayai_annotation_zen_PlantGARDEN_Kegg.tsv.gz		
タンパク (pep / aa)	Lj3.0_pep.fna.gz		
トランスクリプト (cDNA / mRNA)	Lj3.0_cDNA.ffn.gz		

その他の配列

- ・ゲノム
- ・遺伝子 (cds)
- ・タンパク質 (pep)
- ・トランスクリプト (cDNA)
- ・アノテーション (gff)
- ・Hayai-Annotation ZEN

ゲノム (100種)

- ・Hayai-Annotation (14種)
- ・DNAマーカー (26種)
- ・形質連関 (1種)
- ・連鎖地図 (1種)
- ・SRA (8種)
- ・リンク (8種)
- ・ホモロジー (10種)

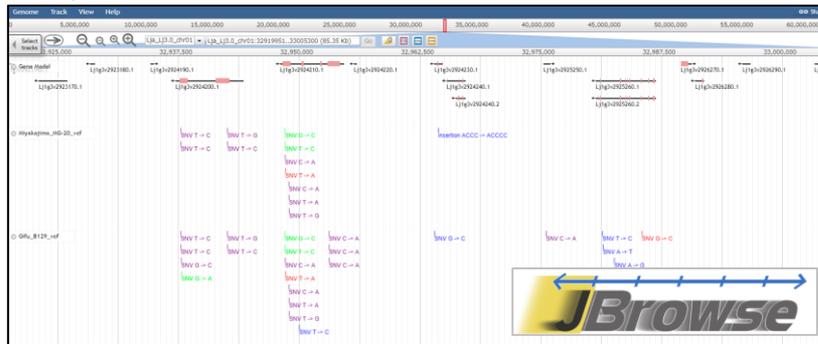
② ゲノム情報のゲノムブラウザへの集約

ゲノム情報

- ゲノム配列
Pseudomolecule (染色体数に収束したもの)
- 遺伝子配列、アミノ酸配列

DNAマーカー情報

- EST-SSR、SNP、CAPSなど
- ### QTL情報
- QTL領域 (物理地図上)



文献からのキュレーション

- 有用遺伝子
- DNAマーカー
- QTL

ゲノムワイド多型情報 (NCBI SRA)

- ゲノム配列 (リシークエンス)

- 品種 (系統) 間のゲノムワイド多型
- SNPアノテーション (SnpEff)

- 遺伝子 (49種)
- SRA (SNP; 8種)
- DNAマーカー (25種)

③ 種を超えた植物ゲノム情報統合のためのデータリンクの構築

Plant GARDENのコンテンツを植物種間で結びつけるための基盤情報を生成

アミノ酸配列の類似度に基づく遺伝子クラスタリングの実施(52植物種、2019年10月)

今後の植物ゲノムの追加、更新データへの置換に対応した方法を開発

* 400植物ゲノムの報告(2019年10月)

系統情報(植物の科)に基づいた配列プロファイルの生成

配列プロファイルを用いた植物科間の遺伝子クラスタの対応付け

全ての植物種でクラスタリング

科でクラスタリング

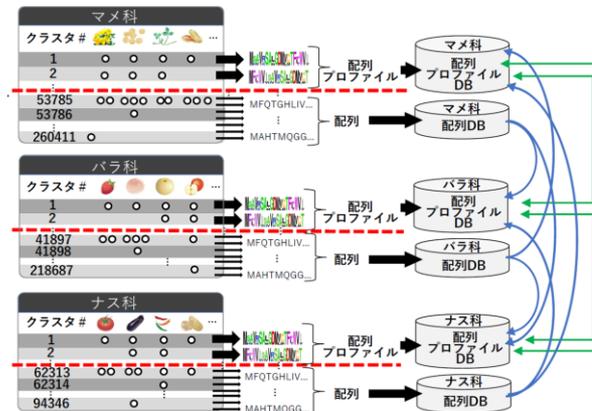
配列
VS
配列



比較



生成される
クラスタの
メンバー



配列プロファイル
VS
配列プロファイル

配列
VS
配列プロファイル

対象: 52植物ゲノム(2019年10月)

アブラナ科(7)、ウリ科(5)、バラ科(7)、マメ科(7)、イネ科(10)、ナス科(6)、キク科(2)、
トウダイグサ科(1)、ヒルガオ科(1)、ブドウ科(1)、ミカン科(1)

④解析ツールの提供 (SNP解析ツール)



<https://pgdbjsnp.kazusa.or.jp>

配列解析システム 1.03	
アカウント	<input type="text"/>
パスワード	<input type="password"/>
ログイン	
●アカウントの新規登録はここをクリック	

無料で利用可能

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)



アカウント、パスワード設定

配列の確認		配列の品質チェック		
配列の前処理				品質でトリミング アダプター除去 ベアリードの検出
ゲノムへのマッピング		マッピング	アップロードされた参照配列のファイル名 <input type="text"/>	
マッピング結果の解析		変異検出		

データのアップロード
・イルミナリード
・リファレンス配列

クオリティチェック (FastQC)
トリミング (PRINSEQ)
マッピング (Bowtie2、TopHat)
SNP検出 (SAMtools)



ユーザに結果の送信
・VCFファイル
(SNP、InDel)



IGVなどで閲覧可能

高速アノテーションツール (Hayai-Annotation Plants)

UniProtKB (緑色植物) に対する USEARCH、GO 解析など

Andrea Ghelfi 作

アライメント手法
(局所、全体)

Type of Alignment

- Local
- Global

Type of Algorithm

- Protein Existence Level
- Alignment Score

Upload FASTA File

Browse... No file selected

Submit

集計表

Show 10 entries

query uniprot seqID

Showing 0 to 0 of 0 entries

uniprot	gene_code	gene_name	Protein_Evidence	EC	evidence_type	GO_BP	GO_BP_name	GO_MF	GO_MF_name
1	14338_ARATH	Pav_sc0000257.1_g090.1.mk GF14 kappa	PE 1: Evidence at protein level	IMP	IMP	GO:005026	response to freezing	GO:0019904	protein domain specific binding
2	14339_ARATH	Pav_sc0002597.1_g010.1.mk GF14 kappa	PE 1: Evidence at protein level	IMP	IMP	GO:0019222	regulation of metabolic process	GO:0019904	protein domain specific binding
3	1A13_SOLLC	Pav_sc0003746.1_g060.1.mk 1-aminocyclopropane-1-carboxylate synthase							1-aminocyclopropane-1-carboxylate synthase
4	AA087WNH2_FICBE	Pav_sc0009909.1_g230.1.mk Peroxidase							Peroxidase

ヒット数

Max hits per query

1

E-value

Evaluate 1e-

6

相同性 (%)

Minimum Sequence Identity (%)

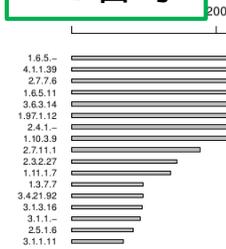
90

ヒット領域 (%)

Minimum Query Coverage (%)

80

EC番号



BP (Biological Process)

CC (Cellular Component)

MF (Molecular Function)

<https://github.com/kdri-genomics/Hayai-Annotation-Plants>

今後の予定

- ・植物種の拡張
- ・キュレーションの半自動化
DNAマーカー、QTL、有用遺伝子など
- ・他データベースとの連携

Microbe DB^{JP} (ホロゲノム解析; 植物と微生物間の相互作用)

- ・キーワード検索機能の付与
筑波大学との共同開発
- ・解析ツールのコンテナ化
SNP解析(国立遺伝学研究所との共同開発)

- ・育種プラットフォームGOBiiとの連携



特別企画「使ってみようバイオデータベース-つながるデータ、広がる世界(BioDB)」
BioDB 8 「Plant GARDEN and PGDBj (Plant Genome DataBase Japan)」