

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⁾、Ghelfi Andrea¹⁾、Fawcett Jeffrey¹⁾、
白澤沙知子¹⁾、市原寿子²⁾、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

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

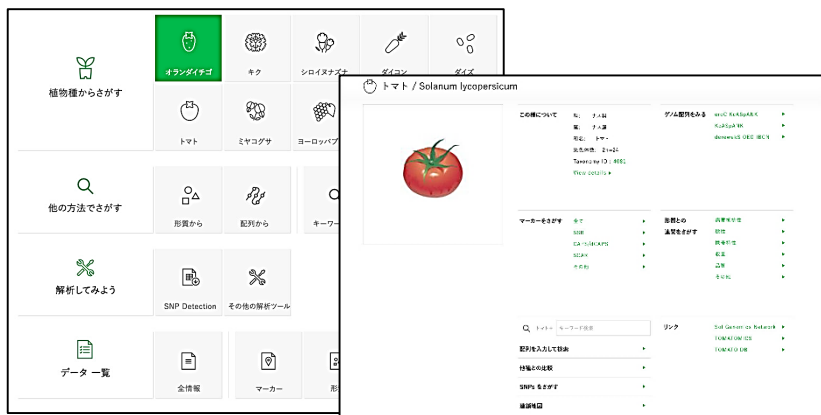


2018年～

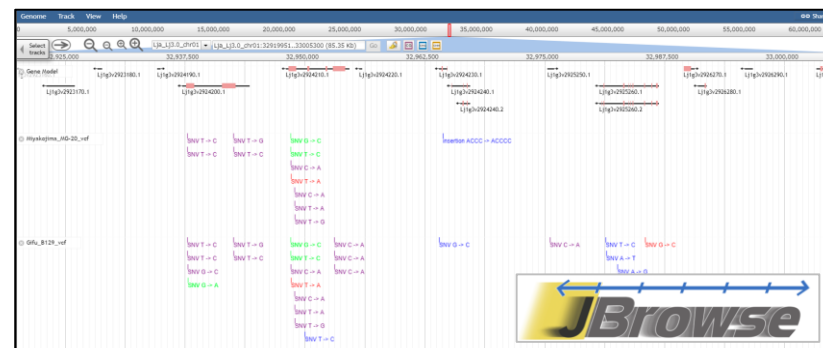
NGSに対応
ユーザ所有のデータを解析
プログラムの開発

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



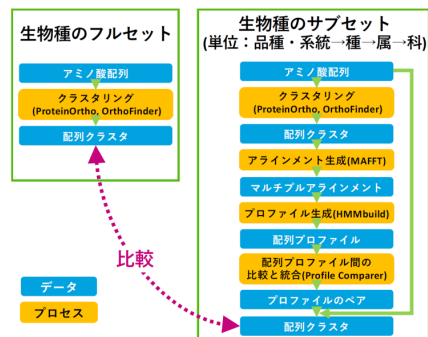
② ゲノム情報のブラウザへの集約



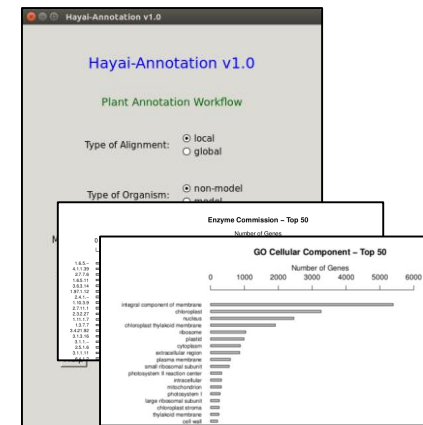
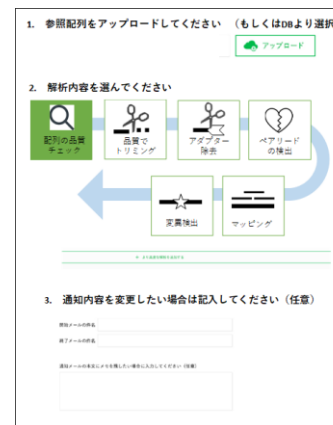
③ 種を超えた植物ゲノム情報統合のためのデータリンクの構築



データベース化



④ 解析ツール群の開発・提供



① Plant GARDENの構築

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。

The screenshot displays the Plant GARDEN web interface with the following components:

- 植物種からさがす** (Search by plant species): A grid of icons for various plants including オランダイチゴ (Strawberry), キク (Chrysanthemum), シロイヌナズナ (Arabidopsis thaliana), ダイコン (Carrot), ダイズ (Soybean), トマト (Tomato), ミヤコグサ (Miyakogusa), ヨーロッパブドウ (European grape), and ラッカセイナンキンマメ (Castor bean). A "すべてみる ▶" (View all) link is also present.
- 他の方法でさがす** (Search by other methods): A search bar with options for 形質から (By trait), 配列から (By sequence), and イーノード検索 (By node).
- 解析してみよう** (Try analysis): Tools for SNP Detection and other analysis tools.
- データ一覧** (Data overview): A list of data types including 全情報 (All information), マーカー (Marker), and 形質 (Trait).

ゲノム解読された360種以上を対象
(Pseudomolecule: 約90種)

現在構築中
今年度末にβ版を公開予定

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
(SL2.5、SL3.0)
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, エボシグサ

Lj2.5

Lj3.0

キーワード検索

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

その他の検索

検索(配列、変異、連鎖地図など)

すべて	SSR	CAPS	SCAR	その他	BLAST 配列からさがす	SNPs 変異をさがす	連鎖地図	他種との比較
-----	-----	------	------	-----	------------------	----------------	------	--------

形質との関連をさがす

形質との関連についての検索

すべて	病害抵抗性	澱性	携帯特性	収量	品質	その他
-----	-------	----	------	----	----	-----

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

さがす	解析してみよう	データ一覧
植物種から	形質から	配列から
SNP Detection	その他の解析ツール	全情報
マーカー	形質	

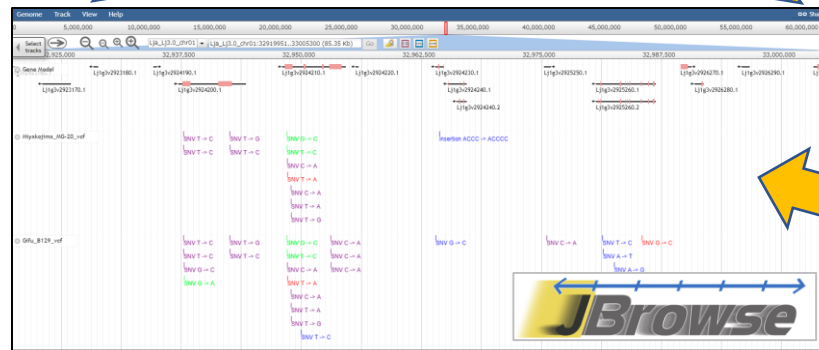
② ゲノム情報のゲノムブラウザへの集約

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

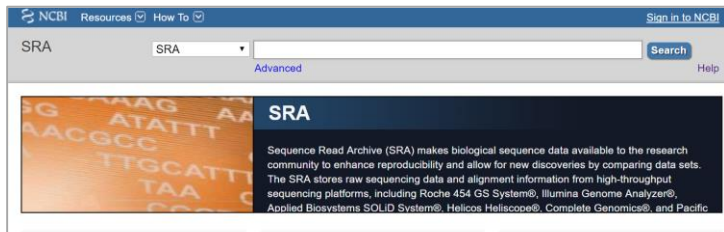
DNAマーカー情報

- 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 (リシーケンス)
(トランスクリプトーム (RNA-Seq、BS-Seqなど))

- 品種 (系統) 間のゲノムワイド多型の閲覧

遺伝子と多型との関係

- SNPアノテーション (SnEff)
遺伝子機能への影響

対象とする植物種（優先度が高いもの）

かずさDNA研究所で解読（15種）

Pseudomolecule



ミヤコグサ



サブクローバ



シバ



オウトウ



イチジク



キヌア

スキャフォールド



ダイコン



イチゴ



バラ野生種



ナス



サツマイモ
野生種



カーネーション



ソバ



ユーカリ



ジャトロファ

外部研究機関（優先度1）

Pseudomolecule



シロイヌナズナ



ハクサイ



ミヤコグサ



ラッカセイ



キク



トマト



キャベツ



ダイズ



ブドウ

（優先度2）



トウガラシ



メロン



イネ



スイカ



キュウリ



コムギ



オレンジ



サツマイモ



トウモロコシ

ユーザの希望により植物種の追加、データ更新を行う

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

各植物種ゲノムデータベースのエントリーを繋ぐ基盤の構築

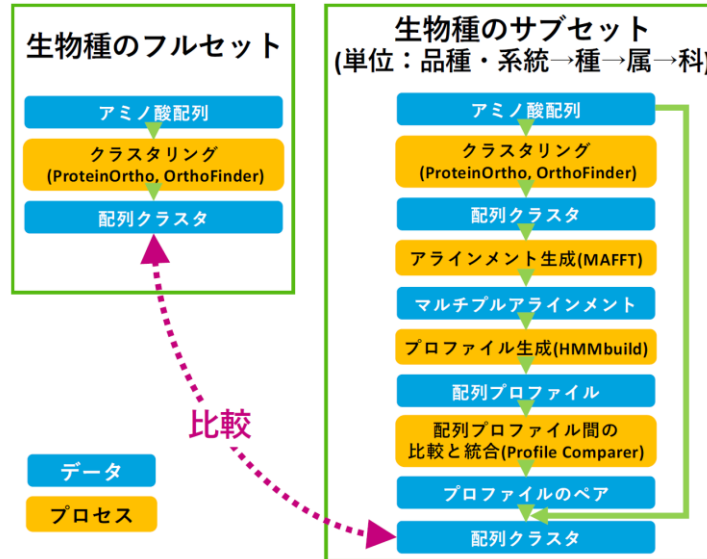
- ・遺伝子のアミノ酸配列の類似性
- ・遺伝子のゲノム配列上の位置情報



配列類似情報に基づくデータリンク



新規植物データの追加への対応



パンゲノム(着目した系統群の遺伝子セットの和集合)の考慮(計算量の軽減)



コア遺伝子群: 着目する階層における系統(群)の全てのゲノムに存在する遺伝子セット

ユニーク遺伝子群: 着目する階層における個々の系統(群)のゲノムにのみ存在する遺伝子セット

対象: 27植物種(かずさDNA研で解読)

- ・双子葉
イチゴ、野生イチゴ、オウトウ、リンゴ、サツマイモ、ナス、トマトなど
- ・単子葉
イネ、シバ

④ 解析ツール群の開発・提供

様々な解析ツールを公開。ユーザが所有する配列データの解析やデータベースに登録されている情報との比較も可能。

画面の説明に沿ってデータをアップロードするだけで簡単に解析可能。
高度なカスタム解析も可能。

解析方法は2通りあります。該当する解析ツールの解析方法を選択してください。



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

SNP Detection

SNPやIn/Delを検出する解析パイプライン

SNP検出



SNP Viewer

SNPデータやグラフ遺伝子型を表示するツール

SNPビューワ



Hayai-Annotation

遺伝子配列のアノテーション

高速アノテーション



Genomic Selection

Genomic Selectionの解析パイプライン

Genomic Selection



TAS Genotyping

Target Amplicon Sequenceの遺伝子型解析

TAS解析



HSE

遺伝子のハプロタイプ特異的発現解析

ハプロタイプ特異的発現



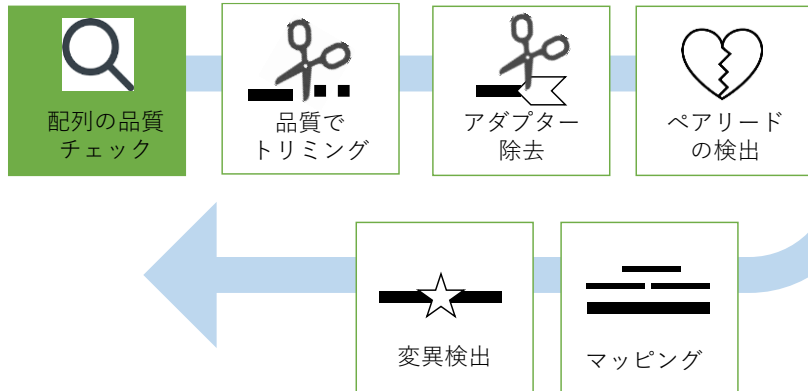
カスタム型多型・ハプロタイプ検出システムの構築

多型検出

1. 参照配列をアップロードしてください (もしくはDBより選択)

アップロード

2. 解析内容を選んでください



+ より高度な解析を追加する

3. 通知内容を

開始メールの件名

終了メールの件名

通知メールの本文にメモを

- ・ユーザの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)
- ・メールで通知

ハプロタイプ検出

多型データをアップロード

リファレンス配列を指定

LD、ハプロタイプブロック検出



ハプロタイプ同定

ブラウザで結果の表示

結果のダウンロード

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

an ultra-fast and comprehensive gene annotation system in plants

UniProtKB (緑色植物) に対するUSEARCH、GO解析など

The screenshot shows the Hayai-Annotation Plants web interface. The main title is "Hayai-Annotation Plants". On the left, there are options for "Type of Alignment" (Local and Global) and "Type of Algorithm" (Protein Existence Level and Alignment Score). There is a "Submit" button and a "Browse..." button for uploading a FASTA file. The main area contains several input fields: "Max hits per query" (set to 1), "Evaluate 1e-" (set to 6), "Minimum Sequence Identity (%)" (set to 90), and "Minimum Query Coverage (%)" (set to 80). Below these are three horizontal bar charts: "EC番号" (EC numbers), "BP (Biological Process)", and "CC (Cellular Component)". To the right of the charts is a "MF (Molecular Function)" chart. At the bottom, there is a "Hayai-Annotation Output Table" with columns for uniprot, gene_code, gene_name, Protein_Evidence, EC, evidence_type, GO_BP, GO_BP_name, GO_MF, and GO_MF_name. The table shows four entries. A red box at the bottom contains the URL: <https://github.com/kdri-genomics/Hayai-Annotation-Plants>.

アライメント手法 (局所、全体)

ヒット数

相同性 (%)

E-value

ヒット領域 (%)

配列のアップロード

EC番号

BP (Biological Process)

CC (Cellular Component)

MF (Molecular Function)

集計表

uniprot	gene_code	gene_name	Protein_Evidence	EC	evidence_type	GO_BP	GO_BP_name	GO_MF	GO_MF_name
14338_ARATH	Pav_sc0000257.1_g890.1.mk	14-3-3-like protein GF14 kappa	PE 1: Evidence at protein level	IMP	IMP	GO:0050826	response to freezing	GO:0019904	protein domain specific binding
14339_ARATH	Pav_sc0002597.1_g010.1.mk	14-3-3-like protein GF14 kappa	PE 1: Evidence at protein level	IMP	IMP	GO:0019222	regulation of metabolic process	GO:0019904	protein domain specific binding
1A13_SOLLIC	Pav_sc0003746.1_g060.1.mk	1-aminocyclopropane-1-carboxylate synthase							1-aminocyclopropane-1-carboxylate synthase
AA087WNH2_FICBE	Pav_sc0009909.1_g230.1.mk	Peroxisomal membrane protein 1							Peroxisomal membrane protein 1

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

Plant GARDENでSNP解析の デモユーザー募集！

SNP解析システムのデモユーザーを募集しております。
アカウント情報をお知らせしますので、是非ご連絡下さい。

pgdbj@kazusa.or.jp

お手持ちのデータからリードの精査やマッピング、
SNPやInDel解析が行えます。

新たに加えたい植物種や機能がありましたらご連絡ください。
皆様のご要望を反映し充実したデータベースを構築します。

今年度末にPlant GARDEN(β版)公開予定

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 「植物ゲノム情報統合データベースPGDBj / Plant GARDEN」