

## ウェブで使える作物家系図の作成ツール 「Pedigree Finder」の開発

### 一品種の祖先・子孫の情報を提供・利活用するデータベース

農研機構と情報・システム研究機構は、作物の家系図を利用者が自由に作成できるデータベース「Pedigree Finder」を開発しました。本データベースは、作物種・世代数などの情報を入力することで、品種・系統<sup>1)</sup>の類縁関係とそれらが示す作物特性を関連づけて表示できます。このため、育種関係者が目的とする形質を持った品種育成に最適な両親の組合せを効率的に選ぶことに利用できます。さらに、複数の作物種において同一の方法で簡便に利用できることから、生産者や消費者が品種の特性を理解するツールとするなど、家系情報<sup>2)</sup>活用ツールとしての幅広い利用が期待されます。

作物は、世界的な食料需要量の増加、食料安全保障対策の強化及び激しさを増す気候変動等に対応するため、常に新たな品種等の育成が求められています。この社会的な要求を満たすため、新品種育成の場面では、既存の品種・系統（以下、品種等）を交配<sup>3)</sup>することで多様な系統（品種の候補）を作出した上で、各系統の形質<sup>4)</sup>を適切に評価し、次世代の品種となる優秀な系統を選抜するという一連の作業を加速化・効率化する必要があります。

しかし、家系情報や形質情報、遺伝子情報等を含む育種に有用な諸情報は研究機関や作物ごとに散在しており、また、これらの育種情報のデータ形式も一律ではないことから、特に、最初に交配する既存品種等の選定段階において、最適な組合せを迅速に判断することが困難でした。

そこで、農研機構及び情報・システム研究機構は、誰もが作物の家系情報を簡単に活用することを可能にするデータベース「Pedigree Finder」を開発しました。「Pedigree Finder」は、イネ、イチゴ等の様々な作物の家系情報をウェブ上で一元管理し、かつ、家系情報と品種等の特性を視覚的に一覧できるデータベースとして提供することを実現しました。これにより、閲覧者ごとのデータの利用権限の範囲内で、作物横断的に整備された家系情報と品種等の特性に関する情報を、目的に従って自在に利用することができます。

「Pedigree Finder」は、育種者が近縁関係を考慮した交配親を選定する場面に限らず、研究者が分子育種学的な観点から親から子へと受け継がれる「遺伝的因子」を解明する場面、生産者が品種等の特性を考慮した栽培品種を選択する場面などにも利活用できます。なお、本データベースは試験運用を終え、本日より、5種の作物（イネ、イチゴ、小麦、サツマイモ、小豆）で国内向けに本格運用を開始します。

公開 URL: Pedigree Finder Top Page (<https://pedigree.db.naro.go.jp/>)

<関連情報>

予算:

農林水産省委託プロジェクト “種苗開発を支える「スマート育種システム」の開発”  
(BAC1001)

戦略的イノベーション創造プログラム (SIP) (スマートバイオ産業・農業基盤技術) “データ駆動型育種” 推進基盤技術の確立とその活用による作物開発” (DDB1002)

ROIS-DS-JOINT 2022 (012RP2020, 006RP2021, 022RP2022)

NAROchannel:

系譜情報データベース「Pedigree Finder」の使い方

<https://www.youtube.com/watch?v=DfauleZL4lQ>

## 問い合わせ先など

研究推進責任者: 農研機構 基盤技術研究本部 農業情報研究センター

センター長 中川路 哲男

同 作物研究部門 所長 石本 政男

同 農業機械研究部門 所長 天羽 弘一

情報・システム研究機構 データサイエンス共同利用基盤施設

ライフサイエンス統合データベースセンター センター長 小原 雄治

研究担当者: 農研機構 基盤技術研究本部 農業情報研究センター AI 研究推進室

上級研究員 鐘ヶ江 弘美

同 インキュベーションラボ ラボ長 米丸 淳一

同 作物研究部門 スマート育種基盤研究領域 上級研究員 松下 景

同 企画戦略本部 研究統括部国際課 課長 竹崎 あかね

(研究当時: 同 農業機械研究部門 知能化農機研究領域 上級研究員)

情報・システム研究機構 データサイエンス共同利用基盤施設

ライフサイエンス統合データベースセンター 特任准教授 川島 秀一

広報担当者: 農研機構 基盤技術研究本部 研究推進室 渉外チーム 東城 僚

TEL 029-838-6859 プレス用 e-mail [www\\_kiban@ml.affrc.go.jp](mailto:www_kiban@ml.affrc.go.jp)

情報・システム研究機構 データサイエンス共同利用基盤施設

ライフサイエンス統合データベースセンター 特任准教授 箕輪 真理

TEL 04-7135-5508 プレス用 e-mail [public\\_relations@dbcls.rois.ac.jp](mailto:public_relations@dbcls.rois.ac.jp)

本資料は農政クラブ、農林記者会、農業技術クラブ、筑波研究学園都市記者会、文部科学記者会、科学記者会に配付しています。

※農研機構 (のうけんきこう) は、国立研究開発法人 農業・食品産業技術総合研究機構のコミュニケーションネーム (通称) です。  
新聞、TV 等の報道でも当機構の名称としては「農研機構」のご使用をお願い申し上げます。

## 開発の社会的背景と研究の経緯

---

### データベース開発の経緯

作物は、全世界的な食料需要量の増加、食料安全保障を取り巻く国内外の状況の変化及び激しさを増す気候変動等に対応するため、常に新たな品種の育成が求められています。育種の加速化・効率化を進めるには、育種関連情報を整備し、かつ、複雑な家系情報の簡便で網羅的な検索・閲覧を可能とすることで、研究者や育種者の調査負担の軽減を図る必要がありました。また、品種等の育成過程を記録したものの多くが祖先方向に数世代しか系譜を遡ることができず、子孫方向の情報を得ることが難しいため、限定された範囲の家系情報で最適な交配組合せを判断せざるを得ない状況でした。そこで、誰もが作物の家系情報を簡単に活用することができるデータベースの開発が望まれていました。

### 家系情報に係るデータ整備の現状

作物の家系情報は、品種育成過程の重要な育種操作である交配親の選定において、最適な組合せを判断するために重要な情報ですが、データの利活用には以下の問題がありました。

- 育種関連情報は研究機関あるいは作物ごとに記録・管理されている。
- 紙媒体や電子媒体など、データの保管媒体が統一されていない。
- 家系情報が図示された家系図（絵）として保管されている場合も少なくないことから、新たに手作業で家系図を作成するには非常に労力を要し、データとして利活用しづらい問題がある。
- 育種関連情報として記録されている育種方法や交配親の表記に“ゆらぎ”があり、データとしての利活用を一層困難にしている。（例えば、ある記録中では「選抜」による育種が、「純系分離」や「抜穂」と記載されているなど語彙が統一されていない。）したがって、家系情報を整理・活用するためには、データ作成ルールとともに、育種の過程で用いられる語彙（オントロジー<sup>5)</sup>）などについても整備する必要がありました。

## 研究の内容・意義

---

### データベースの機能

#### ●効率よく検索・閲覧できるウェブシステム

系統名を入力後、作物を選択することで、品種等の家系情報を表示できます。また、家系情報が更新された場合には、適切なレイアウトで家系図が自動作成されます。

ウェブシステムの活用により、最新情報へのアクセスや共有が容易になりました。

#### ●作物の家系情報の可視化

グラフデータベース<sup>6)</sup>を活用し、作物の家系情報をネットワーク形式で分かりやすく可視化できます（図1、2）。

#### ●祖先・子孫の世代数を指定した表示

EvoTree Plusなどのツールと比較して、子孫方向への探索機能も充実させています（図3）。交配に使われた回数が多い系統や子孫の数が多い系統の調査など、様々な角度から品種等の重要度を検討し、品種等の起源や近縁関係、ゲノム情報を考慮して戦略的に育種素材を選ぶことが可能になります。

## ●形質・遺伝子型情報による色分け

家系図を形質・遺伝子型情報によって異なる色で塗り分けることにより、家系に特徴的な特性の検索、及び遺伝様式についての手がかりを視覚的にとらえることに役立ちます（図3）。

## ●作物の家系情報の分析基盤の整備

近縁係数計算プログラムの整備により、作物の交配を計画する際に、遺伝的な近縁性を考慮することが可能です。

## ●ユーザーごとのデータアクセス管理

データの公開・非公開についてはデータ提供機関の意見をふまえて設定しており、アクセス権の制御により特定のユーザー間におけるデータ共有が可能となっています。

## あらゆる作物に利用可能なデータ形式の整備

- ・ 育種関連情報を整備するためにオントロジー「Pedigree Finder Ontology (PFO)」 (<https://github.com/dbcls/pfo>) を構築し、家系情報の標準フォーマットの決定及び情報の電子化を行いました。
- ・ 作物横断的にデータを活用するために、RDF (Resource Description Framework) <sup>7)</sup> でデータを整備し、育種関連情報をユーザー（研究者や育種者など）が効率よく検索できるようになりました。
- ・ RDF でデータを整備することで、データの意味を理解したプログラムの作成が容易になり、AI 分野での更なる活用も期待されます。
- ・ 系統を ID で管理し、品種等名の表記の“ゆらぎ”に対応しました。
- ・ 作物の家系情報を取り扱うデータフォーマットはこれまでも幾つか開発・公開されていますが、突然変異育種や純系分離などの交配育種以外で得られた品種等を扱えるものはなく、系統の識別性という観点からも本データベースには新規性があります。

## まとめ

近年、形質データに加えて、ゲノム情報データや画像データなど、様々な育種に有用な情報が作物品種等に関連付けられて利用可能になっています。その一方で、家系情報にこれらのデータをひも付ける仕組みがなかったため、各種データを連携した有効利用ができませんでした。本研究で開発した「Pedigree Finder」によって、家系情報をハブとして育種に必要な各種データの統合利用が可能になると期待されます。データベース開発の過程では、様々な作物の育種者の要望やデータ解析担当者のニーズが数多く反映されており、情報が最適に集約されているため、ユーザーが必要な情報を簡単に探し出すことができます。

## 今後の予定・期待

今後は、作物の家系情報を簡単に更新できる仕組みを構築するとともに、作物の種類も増やしていく予定です。また、品種の特性やゲノム情報も合わせて利用できるように、機能の追加を計画しています。

「Pedigree Finder」により育種データの蓄積・共有が促進され、多くのデータをお互いに参照することで品種育成が加速化することが期待されます。例えば、たくさんの情報の中から、機能性成分の含有量の高さに注目する場合、その情報を品種の家系情報と合わせて表示させることにより、含有量の高い品種や家系を視覚的に把握でき、交配親を効率的に選定することができます。

「Pedigree Finder」で整備されたデータは、育種のみならず、様々な食農情報に対してAIなどによる新たな技術開発の基盤となり、豊かな食及び農を実現するために利用することができます。

## 用語の解説

---

### 1. 品種・系統（又は品種等）

例えば、「コシヒカリ」の種子は、全て「コシヒカリ」のゲノムを持ち、どれを植えても「コシヒカリ」の形質を示します。このように共通のゲノムを持ち、他と区別できる形質を示す集団のうち、種苗法による登録が行われたものを「登録品種」と呼び、一般に栽培されている在来種、品種登録されたことがない品種、品種登録期間が終了した品種を「一般品種」と呼びます。一方、「系統」とは、品種登録前の育成段階のものや研究素材等を指します。

### 2. 家系情報（系譜情報）

親子関係の連鎖を指します。作物の場合は交配育種だけでなく、突然変異育種や純系分離もあり、その場合には元となる品種の情報を記載します。

### 3. 交配

品種改良や育種などのために、人為的に生物の個体間の受粉や受精を行うことを意味します。色、形や性質等の形質が異なり、それぞれに長所を持つ2つの品種等を親として選び、交配して新しい品種を作り出します。

### 4. 形質

エンドウマメの花の色や豆の形などように、生物の特性を形質と呼びます。例えば、イネの多収性や病害抵抗性などは農業的に重要な形質です。

### 5. オントロジー

対象分野における事物の概念、及び概念間の関係を、コンピュータで理解可能な形式で記述するための方法です。特に機械学習や人工知能（AI）の分野で活用が進んでいます。

### 6. グラフデータベース

グラフ形式のデータを格納・検索することのできるデータベース管理システムのことです。データの構造がネットワーク状になっている場合に、格納・検索の面で高い性能を発揮します。

### 7. RDF (Resource Description Framework)

主語、述語、目的語の3つ組を単位として、情報をグラフとして記述するデータモデルです。情報をコンピュータに理解しやすく、共有・再利用しやすい形で記述することができます。

## 発表論文

鐘ヶ江弘美・松下景・林武司・川島秀一・後藤明俊・竹崎あかね・矢野昌裕・菊井玄一郎・米丸淳一（2022）系譜情報グラフデータベース「Pedigree Finder」．育種学研究，24:115-123 <https://doi.org/10.1270/jsbbr.22J02>

## 参考図

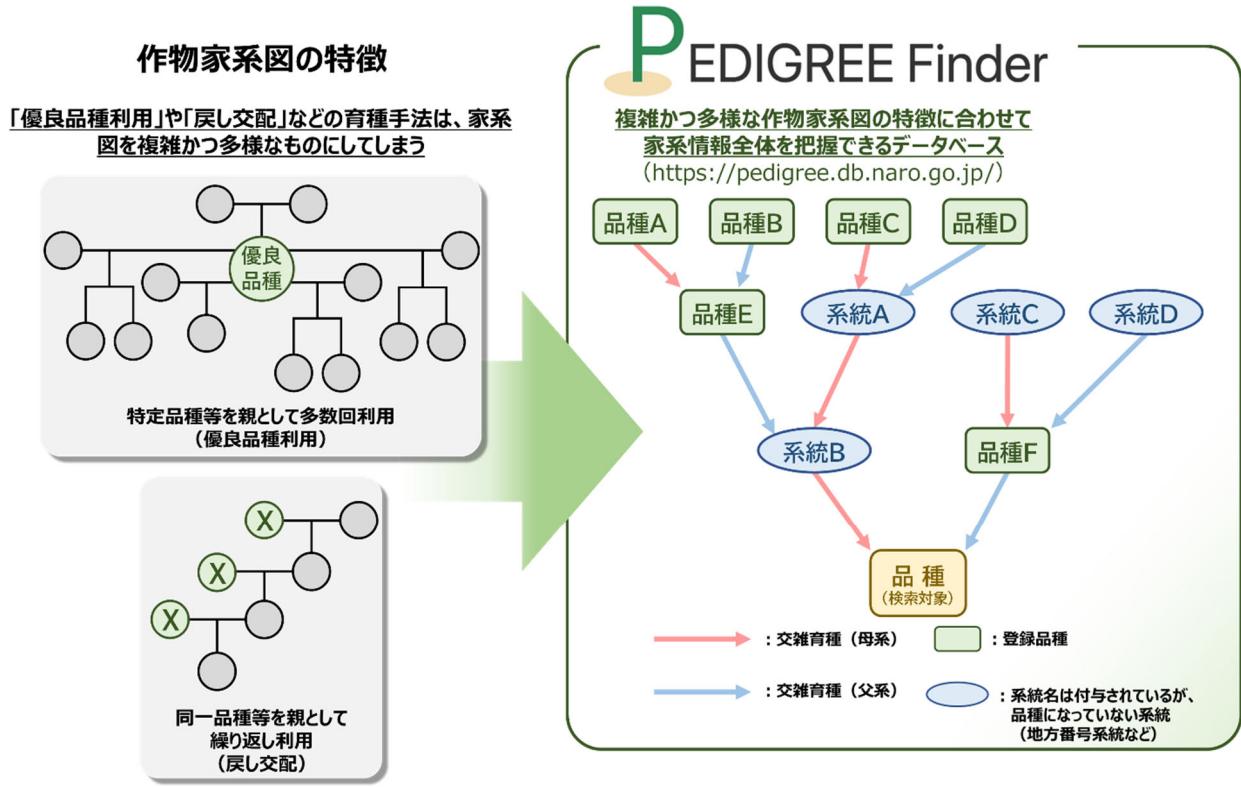


図1 作物家系図の特徴

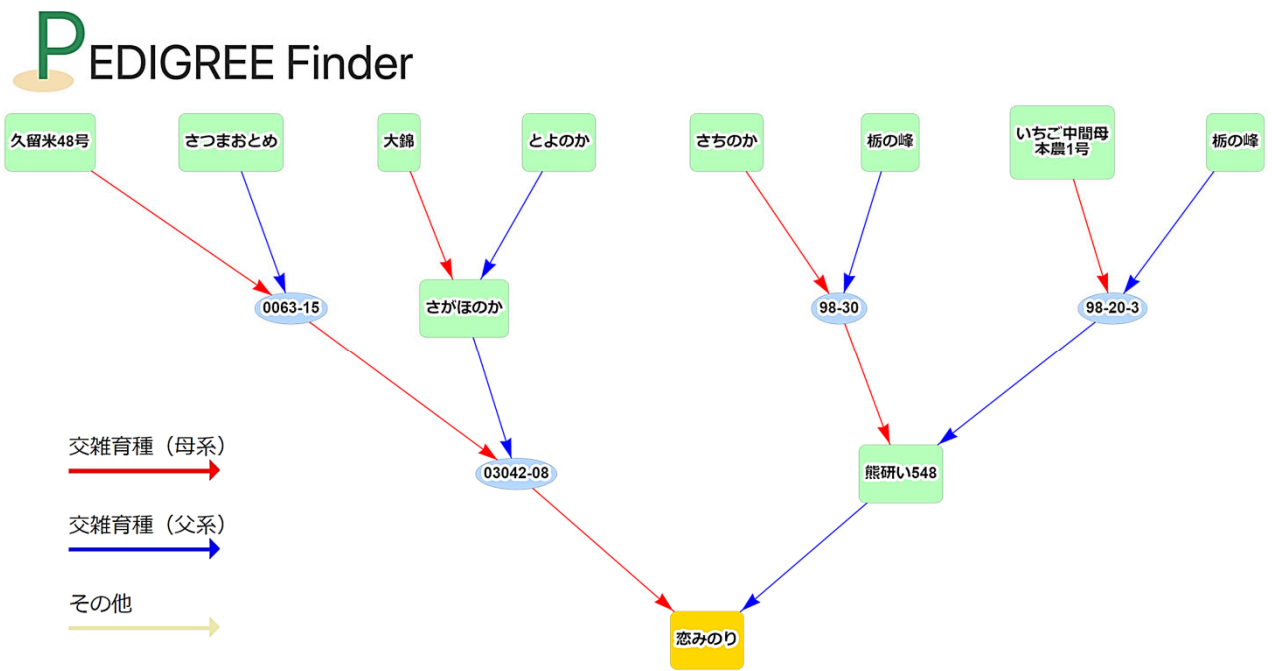


図2 Pedigree Finder での家系図の表示例

### イチゴ「恋みのり」家系情報

「恋みのり」は多収性の早生系統「03042-08」を母親に、食味に優れる「熊研い548」（商標名「ひのしずく」）を父親としています。

## ユーザーインターフェースの特徴

### 検索可能作物種

イチゴ、サツマイモ、イネ、小豆、小麦  
 ゲストユーザー(アカウント未登録)での利用可能

### 祖先方向の表示世代数

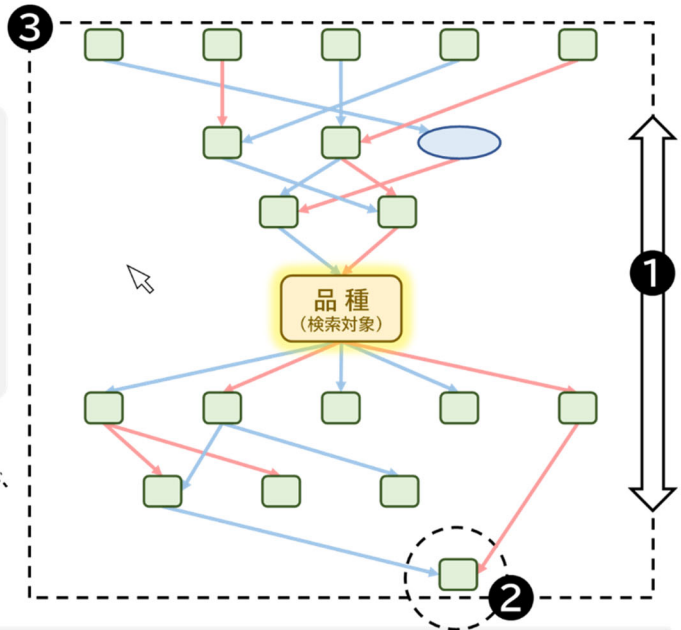
初期表示時: 3 (最大: 15)

### 子孫方向の表示世代数

初期表示時: 1 (最大: 5)

- (赤) : 交雑育種 (母系)    □ (緑) : 登録品種
- (青) : 交雑育種 (父系)    ○ (青) : 系統名は付与されているが、品種になっていない系統 (地方番号系統など)

- ① 祖先方向・子孫方向の家系情報を表示
- ② ダブルクリックで注目系統の家系図を新たに表示可能

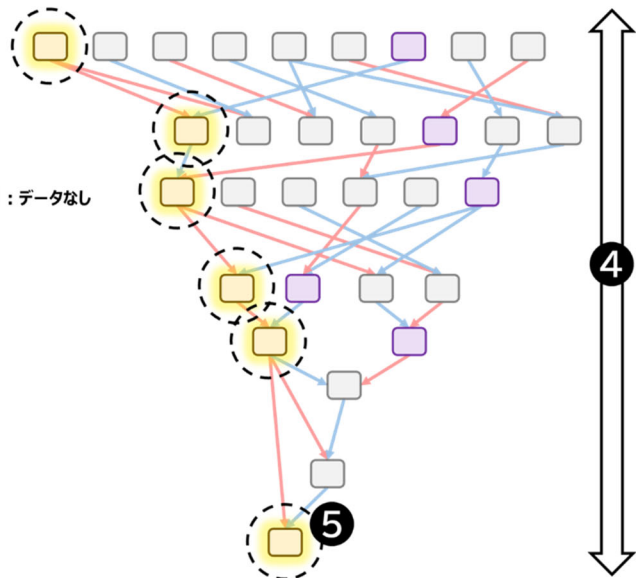


### データセットに対応した表示

品種登録データや近縁係数などのデータセットに基づいて、  
 形質・遺伝子情報により色付けして表示することも可能

- (赤) : 交雑育種 (母系)    □ (黄) : 病害抵抗性あり    □ (白) : データなし
- (青) : 交雑育種 (父系)    □ (紫) : 病害抵抗性なし

- ③ 右クリックで家系図の画像を保存可能
- ④ “家系”と“品種等が持つ特性”の関係を一覧的に表示
- ⑤ 病害抵抗性の遺伝子の伝達を視覚的に表現





「Pedigree Finder」の  
使用方法紹介動画



「Pedigree Finder」公開HP

さらに詳しく知りたい方はこちらをクリック

図 3 Pedigree Finder ユーザーインターフェースの特徴