

2024 年度 研究開発実施報告

概要

研究開発課題名	非モデル植物のための遺伝子ネットワーク情報活用基盤
開発対象データベースの名称(URL)	ATTED-II (https://atted.jp/)
研究代表者氏名	大林 武(50397048)
所属・役職	東北大学 情報科学研究科 教授 (2025 年 3 月時点)



©2024 大林 武(東北大学) licensed under CC表示4.0国際

□目次

概要	1
§1. 研究実施体制	4
§2. 研究開発対象とするデータベース・ツール等	5
(1) データベース一覧	5
(2) ツール等一覧	5
§3. 実施内容	6
(1) 本年度に計画されていた研究開発項目・タスク	6
(2) 進捗状況	6
1) 共発現提供生物種の情報拡充のためのシステム開発	6
2) 種間比較機能の高度化	7
3) 遺伝子共発現情報の構築と更新	7
4) 公開環境の整備と負荷増大対応	8
§4. 成果発表等	9
(1) 原著論文発表	9
1) 論文数概要	9
2) 論文詳細情報	9
(2) その他の著作物(総説、書籍など)	9
(3) 国際学会および国内学会発表	9
1) 概要	9
2) 招待講演	9
3) 口頭講演	9
4) ポスター発表	10
(4) 知的財産権の出願（国内の出願件数のみ公開）	10
1) 出願件数	10
2) 一覧	10
(5) 受賞・報道等	10
1) 受賞	10
2) メディア報道	10
その他の成果発表	10
§5. 主要なデータベースの利活用状況	11
(1) アクセス数	11
1) 実績	11
2) 分析	11
(2) データベースの利用状況を示すアクセス数以外の指標	12
(3) データベースの利活用により得られた研究成果(生命科学研究への波及効果)	12
(4) データベースの利活用によりもたらされた産業への波及効果や科学技術のイノベーション(産業や科学技術への波及効果)	14
§6. 研究開発期間中に主催した活動(ワークショップ等)	15

(1) 進捗ミーティング	15
(2) 主催したワークショップ、シンポジウム、アウトリーチ活動等	15

§1. 研究実施体制

グループ名	研究代表者・ 研究分担者 氏名	所属機関・役職名	研究題目
大林グループ	大林 武	東北大学・教授	遺伝子ネットワーク情報活用基盤 ATTED-IIの開発

§2. 研究開発対象とするデータベース・ツール等

(1) データベース一覧

【主なデータベース】

No.	名称	別称・略称	URL
1	植物の遺伝子共発現データベース ATTED-II	ATTED-II	https://atted.jp/

【その他のデータベース】

No.	名称	別称・略称	URL
1			

(2) ツール等一覧

No.	名称	別称・略称	URL
1			

§3. 実施内容

(1) 本年度に計画されていた研究開発項目・タスク

- 1) 共発現提供生物種の情報拡充のためのシステム開発
 - ・ RNAseq 共発現の高精度化
 - ・ 共発現モジュールが機能するサンプル条件の提示する機能の開発
 - ・ データベース間の連携機能の高度化
- 2) 種間比較機能の高度化
 - ・ マクロスケールの種間比較システムの高度化
 - ・ メゾスケールの種間比較システムの開発
- 3) 遺伝子共発現情報の構築と更新
 - ・ 遺伝子共発現データの計算
 - ・ 遺伝子共発現の解釈のためのデータの計算と公開

(2) 進捗状況

1) 共発現提供生物種の情報拡充のためのシステム開発

1)-1. RNAseq 共発現の高精度化

昨年度の成果として得られた標準種と非標準種の RNAseq 共発現データ(特に *Arabidopsis* の Col-0 以外を含むデータ)を基に、2024 年度は共発現情報の精度向上を目的としたサンプル再構成とその評価を実施した。RNAseq 定量法として、アライメントベースの bowtie2 を採用し、アライメントフリーの手法と比較して非標準種データにおける KEGG スコアの改善を確認した。これにより、RNAseq データの多様性を考慮した定量手法の選択が、共発現の精度向上に有効であることを確認した。また、次項の「PC View」機能を通じて、共発現計算にも用いている主成分の特徴の見える化が達成され、今後、種間比較を通じた性能評価や更なる共発現の開発に繋げていく。

1)-2. 共発現モジュールが機能するサンプル条件を提示する機能の開発

昨年度に試作した主成分空間におけるサンプル条件提示ビューアを「PC View」として発展させ、ATTED-II v12.0 に実装・公開した。具体的には、遺伝子発現量テーブルの主成分分析により得られた各主成分について、次の点を吟味できるようにした。

- どの遺伝子とその主成分に応答しているか(主成分スコア):主成分スコアの上位と下位の 100 遺伝子を表示
- どのサンプルがその主成分に寄与しているか(因子負荷量):SRA メタデータから自動的に比較軸を抽出して表示

なお、主成分分析により得られる上記 2 種の情報については、単一のサンプルの環境ではなく、サンプル間の環境の差異が主成分になっている。各々のサンプルは多数の要因によって特徴づけられるため、このような要因を分解するアプローチが本質的に重要である。

このシステムにより、ユーザーは任意の遺伝子ペアに対して、どの主成分が共発現に寄与しているかを可視化と関連情報へのアクセスが可能となった。今後、主成分スコアが変動している遺伝子リストの要約

(KEGG Pathway の濃縮検定など)、ならびに複数の研究プロジェクトの比較軸の要約を提示する機構を開発し、ユーザーが着目する遺伝子ペアが機能するための外部環境と細胞内環境のより直感的な理解を目指す。

1)-3. データベース間の連携機能の高度化

前年度に整備した API を基盤として、ATTED-II の Arabidopsis 遺伝子ページから Plant GARDEN へのリンク機能を実装・公開した。これにより、ゲノム多型情報などとの相互参照が可能となり、連携機能の高度化が一部達成された。さらに、共発現データに関連する各種情報を Web 上から容易にダウンロード可能とするため、以下の 4 点について整備を行なった。

- PCA 関連データ(発現行列、主成分、ローディング情報)
- 各遺伝子の共発現リストに対する KEGG Pathway 濃縮解析結果
- CoexMap 関連データ
- 共発現データや発現データなどの大規模データの Zenodo リポジトリへの登録と ATTED-II ダウンロードページからの直リンクの自動生成

これにより、ユーザーによる再解析やローカルでの利活用が容易になり、データベースの利便性・透明性・拡張性が向上した。

2) 種間比較機能の高度化

昨年度に開発した外れ値補正アルゴリズムを活用し、UMAP ベースで可視化されるゲノムワイド共発現マップの種間整列性を改善した。本年度は、この整列性の向上に基づき、複数の植物種における遺伝子を KEGG Ortholog 情報に基づいて対応付け、種を超えた視覚的な共発現関係の比較を可能とした。具体的には、NetworkDrawer に KEGG Ortholog 情報を取り込み、同一 Ortholog に属する遺伝子を色分け表示する機能を新たに実装した。これにより、ユーザーは異なる植物種における保存された遺伝子群の共発現関係を視覚的に比較・評価できるようになった。

また、共発現データのクラスタリングと共発現モジュールの抽出を通じて、モジュール間の階層構造や保存性を評価する手法の検討を開始しており、次年度以降にかけて機能としての統合を目指す。

3) 遺伝子共発現情報の構築と更新

当初計画していたインディカ米 (*Oryza sativa indica*)、ジャガイモ (*Solanum tuberosum*)、およびモロコシ (*Sorghum bicolor*) の共発現データ構築に向けて、RNAseq データの収集と定量処理を実施した。これらの生物種は公共データベースにおけるデータ数や条件のばらつきが大きく、特にデータの不均一性や品質のばらつき、加えて複数系統が混在していることによる定量化パイプラインの最適化に時間を要したため、2024 年度内での公開には至らなかった。一方で、パイプラインのマニュアル化を進め、2025 年度の目標である 20 種公開に向けた基盤構築は進んでおり、継続開発を行う予定である。

一方で、昨年度構築したコムギ・オオムギの共発現情報については、統合的なインターフェースの整備を進めた上で、2024 年 9 月に ATTED-II version 12.0 として正式に公開した。インターフェースの改良では、個々の遺伝子ページと共発現ネットワークを統合し、ユーザーがより直感的に共発現関係へアクセスできるようにした。CoexViewer による主成分空間における相関評価機能も新たに実装し、解釈性の向上に寄与している。

4)公開環境の整備と負荷増大対応

2023 年度後半に導入した新しい物理サーバを公開環境としてセットアップし、このサーバにおいて ATTED-II version 12.0 公開用仮想マシンを運用した。この仮想マシンの実体を SSD 上に配置することで応答速度の大きな向上が確認された。

一方、2024 年後半において、ATTED-II へのアクセス集中による応答遅延や一時的な停止が散発的に起きるようになった。原因調査の結果、特定範囲の IP アドレスからの過剰なリクエストがサーバ負荷の主因であることを特定し、当該範囲の IP アドレスの遮断を行うとともに、キャッシュ機構の導入と API 呼び出し制限の最適化による状況の緩和を行った。当該範囲の IP アドレスの全てが問題であるわけではないため、アクセス遮断を動的に行うなどの改善の余地は多く、次年度も対応を継続する。

§4. 成果発表等

(1) 原著論文発表

1) 論文数概要

種別	国内外	件数
発行済論文	国内(和文)	0 件
	国際(欧文)	1 件
未発行論文 (accepted, in press 等)	国内(和文)	0 件
	国際(欧文)	0 件

2) 論文詳細情報

1. Kanako Bessho-Uehara, Takeshi Obayashi, “Evolutionary approaches for narrowing down candidate genes from an unannotated gene list”, Plant and Cell Physiology, vol. 66, Issue 3, March 2025, Pages 287–290, DOI: 10.1093/pcp/pcaf003

(2) その他の著作物(総説、書籍など)

なし

(3) 国際学会および国内学会発表

1) 概要

種別	国内外	件数
招待講演	国内	0 件
	国際	0 件
口頭発表	国内	1 件
	国際	0 件
ポスター発表	国内	3 件
	国際	0 件

2) 招待講演

〈国内〉

なし

〈国際〉

なし

3) 口頭講演

〈国内〉

1. Obayashi T, ATTED-II Version 12.0: A Plant Gene Coexpression Database for Comparative Analysis Across Flowering Plants,
第 66 回日本植物生理学会年会, 金沢大学, 2025/3/14

〈国際〉
なし

4) ポスター発表

〈国内〉

1. 大林武, 植物の共発現データベース ATTED-II における種間比較機能強化の取り組み, DICP 研究交流会, JST 東京本部(サイエンスプラザ), 2024/10/4
2. 大林武, 植物の遺伝子共発現データベース ATTED-II version 12.0, トーゴの日シンポジウム 2024, 品川ザ・グランドホール, 2024/10/5
3. Obayashi T, ATTED-II Version 12.0: Plant Gene Coexpression Database with Enhanced Species Comparison for Supporting Flowering Plants, APBIC 2024[1st Asia & Pacific Bioinformatics Joint Conference], NAHA CULTURAL ARTS THEATER NAHArt, 2024/10/24,

〈国際〉
なし

(4) 知的財産権の出願（国内の出願件数のみ公開）

1) 出願件数

種別		件数
特許出願	国内	0 件

2) 一覧

2)-1. 国内出願

なし

(5) 受賞・報道等

1) 受賞

なし

2) メディア報道

なし

その他の成果発表

なし

§5. 主要なデータベースの利活用状況

(1) アクセス数

1) 実績

表 1 研究開発対象の主要なデータベースの利用状況

名称	種別	2024 年度(月間平均値)
ATTED-II	訪問者数	1,225
	訪問数	2,207
	ページ数	13,236

2) 分析

上記のアクセス数は Google Analytics によるものであり、2023 年度(訪問者数 1,067、訪問数 2,244、閲覧ページ数 14,588)と比較して、同程度の水準であった。

一方、AWStats によるアクセス解析では、以下のとおり、はるかに多いアクセス数が観測されている。

表:AWStats によるアクセス状況(抜粋)

	訪問者数	訪問数	ページ数
Aug-24	5,686	92,007	1,287,983
Sep-24	16,625	271,562	899,116
Jan-25	988,465	3,215,038	8,690,354
Feb-25	22,498	195,427	1,235,613
Mar-25	278,581	570,098	4,509,594

また、仮想マシン上の httpd ログに基づく API アクセスの解析では、月間約 30,000 件のアクセスが確認されている。その多くは、シロイヌナズナの世界的ポータルである Araport からのアクセスであり、これにより、API の提供がシロイヌナズナの情報を参照する植物研究者の利便性向上に寄与しているものと解釈している。

Google Analytics と AWStats のアクセス数に大きな差がある理由としては、Google Analytics が JavaScript を無効にしているブラウザからのアクセスや API へのアクセスを検出できないのに対し、AWStats ではそれらも含めて集計している点が挙げられる。

一般的に、通常のユーザが JavaScript を無効にしてアクセスするケースは少ないため、これらのアクセスはプログラムによるものであると推測され、ロボットによる自動アクセスが含まれている可能性は十分にある。ATTED-II では、基本的に robots.txt によりクローラーの巡回を拒否しているが、近年の大規模言語モデルなどに伴う自動データ収集の増加が影響している可能性がある。

今後は、過剰なロボットアクセスが通常のサービス提供の妨げにならないよう、JavaScript 実行を前提としたページ構成や、さらなるアクセス制限の導入についても検討する必要がある。

(2) データベースの利用状況を示すアクセス数以外の指標

Google Scholar による毎年の ATTED-II 引用論文数を示す。

2019 年: 108
2020 年: 108
2021 年: 103
2022 年: 108
2023 年: 80
2024 年: 81
2025 年: 45 (5 月末日時点)

2023 年より若干引用数が下がっている可能性もあるが、大方これまで通りの水準で利用されている。

(3) データベースの利活用により得られた研究成果(生命科学研究への波及効果)

1. Nocchi ら(2024)は、*Arabidopsis thaliana* および *Medicago truncatula* を含む 17 種の被子植物において、生物種を超えて正の選択を受ける遺伝子の特徴を解析した。比較ゲノム解析により正の選択を受ける遺伝子を抽出し、ATTED-II の共発現データを用いて、それらの遺伝子が共発現ネットワーク中で有意に中心性が低いことを明らかにした。これは、正の選択を受ける遺伝子が発現特異性が高く、機能的多面性が低いことを意味する。ATTED-II で提供する複数種にわたる共発現情報が、植物全般における適応進化の遺伝的基盤に関する知見の発見に貢献しており、本申請課題が究極的に目指す「任意の植物種への共発現情報の提供」という構想の妥当性を裏付ける成果と位置づけることもできる。

G. Nocchi, J. R. Whiting, S. Yeaman, Repeated global adaptation across plant species. *Proc. Natl. Acad. Sci. U. S. A.* 121, e2406832121 (2024).

2. Zhou ら(2024)は、*Arabidopsis thaliana* における脂肪酸合成の制御因子を探索するため、ATTED-IIを用いた共発現ネットワーク解析を実施した。hetACCaseのサブユニットや既知の調節因子(BCCP、BADC、CTI、PII など)をベイトに用いて、共発現 Z スコア 3 以上の遺伝子として RFS1 が特定された。RFS1 は脂質代謝関連遺伝子との共発現性が高く、遺伝子ノックアウト実験を通じて、hetACCase 活性と脂肪酸合成速度の上昇が確認された。

L. Zhou, et al., REGULATOR OF FATTY ACID SYNTHESIS proteins regulate de novo fatty acid synthesis by modulating hetACCase distribution. *Plant Cell* 37 (2024).

3. Shirakawa ら(2025)は、*Arabidopsis thaliana* において、気孔の分化を制御する転写因子 FAMA の下流標的である WSB(WASABI MAKER)遺伝子を特定し、その機能を解明した。WSB と FAMA の共発現関係は、ATTED-II の共発現ネットワークに基づいて示され、さらに WSB が TGG1 や TGG2 などのミロシナーゼ遺伝子とも強く共発現していることから、WSB が気孔細胞とミロシン細胞の両方の分化に関与するという仮説が導出された。この仮説は遺伝子破壊変異体や転写因子の ChIP-qPCR、レポーター解析等により検証された。

M. Shirakawa, et al., Co-option and neofunctionalization of stomatal executors for defence against herbivores in Brassicales. *Nat. Plants* 11, 483–504 (2025)

4. Lucchin ら(2025)は、*Arabidopsis thaliana* におけるカルシウムセンサー CML8 の機能解析において、ATTED-II を活用することで、CML8 と共発現する約 140 個の遺伝子を同定した。これらの共発現遺伝子の多くが成長や発生に関連しており、特に根組織での機能が示唆された。また CML8 がブラシノステロイド早期応答性遺伝子であることと合わせて、ブラシノステロイドの調節を通じて成長制御に関与しているという仮説が導出された。実験により CML8 が BR 受容体 BRI1 とリガンド依存的に相互作用す

ることが示され、この共発現情報に基づいた仮説が検証された。これにより、CML8 がブラシノステロイドシグナル伝達経路に新たな調節機構を提供していることが明らかとなった。

A. Lucchin, et al., The calcium sensor AtCML8 contributes to Arabidopsis plant cell growth by modulating the brassinosteroid signaling pathway. *Plant J.* 121, e17179 (2025).

5. Li ら(2024) は、イネの分げつ数を制御する新規遺伝子を探索するために、ATTED-II の共発現データを活用した。具体的には、ATTED-II を用いて、分げつに関与する既知遺伝子 OsFON1 の共発現遺伝子を探索し、機能未知の糖輸送体 OsSTP15 が候補遺伝子として着目した。OsSTP15 のノックアウトにより、シュート基部での糖(グルコース、スクロース、Tre6P)蓄積が促進され、サイトカニン合成・シグナルの活性化を介して分げつ数が増加し、結果として収量が 20%以上向上することが実験により検証された。ATTED-II の共発現情報は、OsSTP15 が分げつ制御に関与するという仮説を導き出した。

M. Li, et al., Knockout of the sugar transporter OsSTP15 enhances grain yield by improving tiller number due to increased sugar content in the shoot base of rice (*Oryza sativa* L.). *New Phytol.* 241, 1250–1265 (2024).

6. Liu ら(2025) は、トウモロコシの非リグニン性フェニルプロパノイド代謝に関与する酵素 Zm4CL5 の機能を明らかにした。ATTED-II の共発現データを用いて、Zm4CL5 が既知のフラボノイド生合成遺伝子と強く共発現することが示されると同時に、Zm4CL1 や Zm4CL2 とは異なる代謝経路に関与していることが示唆された。得られた仮説に基づき、Zm4CL5 ノックダウン変異体の代謝プロファイルや生育特性、さらに昆虫食害や病原菌感染に対する応答が詳細に解析され、Zm4CL5 がフラボノイド等の代謝産物の蓄積や植食者応答に重要であることが実証された。

H. Liu, et al., The herbivore-inducible coumarate-CoA ligase 5 (Zm4CL5) in maize contributes to the production of non-lignin phenylpropanoids. *Planta* 261, 101 (2025).

7. Feng ら(2025) は、ダイズのソヤサポニン代謝に関与する β -グルコシダーゼ(GmSSBG1)を同定した。まずシロイヌナズナとのホモロジーから、ダイズの β -グルコシダーゼを 66 遺伝子同定した。この 66 遺伝子について ATTED-II を用いて探索したところ、特にそのうち 1 つがソヤサポニン代謝に関与することが示唆された。その後、発現解析、リコンビナントタンパク質の機能解析、ノックアウト実験により、GmSSBG1 は A0 型および B0 型ソヤサポニンに特異的に作用すること。GmSSBG1 遺伝子の欠損により、ソヤサポニンの前駆体が蓄積し、莢実内の害虫 *Leguminivora glycinivorella* への感受性が増加することが確認された。

C. Feng, et al., Soyasaponin β -glucosidase confers soybean resistance to pod borer (*Leguminivora glycinivorella*). *Abiotech* (2025). <https://doi.org/10.1007/s42994-025-00214-7>.

8. Kang ら(2025) は、表皮タンニン含量が異なる 2 つのブドウ品種『Tano Red』と『Ruby Seedless』の果皮において、果実発達初期におけるタンニン生合成遺伝子群の発現を解析した。*Vitis vinifera* を対象に、ATTED-II の共発現データを用いてタンニン生合成関連遺伝子のネットワーク解析を実施し、構造遺伝子(ANR, LAR)と修飾酵素(SCPL-AT)を中心とした強い共発現関係が確認された。この共発現解析により、タンニン合成が果実発達初期に集中的に起こる分子的基盤が明らかとなった。また、共発現ネットワークに含まれる輸送体遺伝子も同様に解析され、タンニンの蓄積・輸送に関わる多遺伝子群の協調的発現が示唆された。ATTED-II の共発現情報は、*Vitis vinifera* のタンニン合成に関する仮説の構築に活用され、その後の qPCR 実験により仮説検証が行われている。

Y. Kang, et al., Early peak of tannin content and gene expression related to tannin biosynthesis in table grape skin during berry development. *Hortic. Environ. Biotechnol.* (2025). <https://doi.org/10.1007/s13580-025-00692-0>.

9. ブロッコリーはカリウム欠乏に対する耐性がシロイヌナズナよりも高い。そのメカニズムを解明するため、Bueso ら(2024) は、ブロッコリーとシロイヌナズナの比較トランスクリプトーム解析を行った。根における BoHAK5 の高発現と、サリチル酸(SA)生合成関連遺伝子の迅速かつ強い誘導が、耐性に寄与する候補として挙げられている。特に、SA 生合成経路の誘導と SA の蓄積は、ATTED-II の共発現ネットワー

クにより得られた *Arabidopsis* における共発現遺伝子群の情報をもとに、ブロッコリーにおける転写因子候補 (WRKY29, WRKY72) 同定に活用された。これにより、ATTED-II が NO_3^- 輸送系や SA 応答経路に関与する転写調節ネットワークの仮説構築に貢献し、それがブロッコリーの耐性機構の実験的検証へとつながった。

E. Bueso, et al., Constitutive BoHAK5 expression and early robust induction of salicylic acid biosynthesis genes as candidates to explain the tolerance of broccoli to limiting potassium. *Discov. Plants* 1 (2024).

10. Hawkins ら (2025) は、植物代謝データベース PMN の最新版である PMN 16 を構築し、155 種の植物・藻類の代謝パスウェイ情報を統合した。*Arabidopsis thaliana*, *Oryza sativa*, *Glycine max*, *Zea mays* など 9 種の主要モデル植物において、パスウェイページ上に ATTED-II の共発現データを統合的に表示できる「Pathway Co-expression Viewer」が搭載されており、これにより、各代謝パスウェイに関与する遺伝子群の共発現関係を可視化し、当データベースにおける仮説生成や遺伝子機能予測に貢献している。

C. Hawkins, et al., Plant Metabolic Network 16: expansion of underrepresented plant groups and experimentally supported enzyme data. *Nucleic Acids Res.* 53, D1606–D1613 (2025).

- (4) データベースの利活用によりもたらされた産業への波及効果や科学技術のイノベーション(産業や科学技術への波及効果)

なし

§6. 研究開発期間中に主催した活動(ワークショップ等)

(1) 進捗ミーティング

年月日	名称	場所	参加人数	目的・概要
2024 年 4 月 10 日	チーム内ミーティング(非 公開)	東北大学 大林研究室	4 人	研究進捗報告のためのミーテ ィング

(2) 主催したワークショップ、シンポジウム、アウトリーチ活動等

(3) なし

以上

別紙 1 既公開のデータベース・ウェブツール等

No.	正式名称	別称・略称	概要	URL	公開日	状態	分類	関連論文
1	植物の遺伝子共発現データベースATTED-II	ATTED-II	代表的な植物の遺伝子共発現情報を提供するデータベース	https://atted.jp	2004/11/14	維持・発展	データベース等	Obayashi T, Hibara H, Kagaya Y, Aoki Y, Kinoshita K. (2022) ATTED-II v11: A Plant Gene Coexpression Database Using a Sample Balancing Technique by Subagging of Principal Components. Plant Cell Physiol. 15;63(6):869-881. doi: 10.1093/pcp/pcac041.