

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 27 日現在

機関番号：82508

研究種目：基盤研究(C) (一般)

研究期間：2014～2016

課題番号：26350967

研究課題名(和文) RNA-Seq解析を用いた植物物質生産制御因子の単離及び最低解析数の統計学的検証

研究課題名(英文) The statistical verification of the minimum experimental datasets for the isolation of plant secondary metabolites related transcription factors using RNA-Seq analysis technique.

研究代表者

鈴木 秀幸 (Suzuki, Hideyuki)

公益財団法人かずさDNA研究所・パイオ研究開発部・グループ長

研究者番号：80276162

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：相関係数値と幾何学的要素(外側と内側の線の繋がり方)で相関ネットワーク解析を行う金平糖アルゴリズムを搭載したConfeitoGUIplus ソフトを開発した。本ソフトは多変量解析の1種である。シロイヌナズナの公開DNAマイクロアレイを活用し、フラボノール生成のモジュールを得ることを示した。除虫菊(Tanacetum cinerariifolium)が生成する天然殺虫剤成分であるピレスリン生成遺伝子情報を網羅的獲得する目的でRNA-Seq解析を行った。得られた数種のピレスリン生成遺伝子発現に対して、傷害誘導的に生じる揮発性有機化合物(VOC)のブレンド作用効果を調査した。

研究成果の概要(英文)：We established standalone ConfeitoGUIplus software for detecting local communities from a correlation network involving size sensitivity. The developed software accepts any kind of a multivariate dataset and a correlation matrix. Using the public Arabidopsis DNA array resource data (more than 9 thousands), we have shown an application of our developed software to isolate gene co-expression network module in Arabidopsis flavonol biosynthesis. We performed RNA-Seq analysis in Tanacetum cinerariifolium for gene discovery to isolate biosynthetic enzyme involved in natural pyrethrins biosynthesis. We have investigated the effects of the blend of the volatile organic compounds (VOCs) on gene expressions of seven biosynthetic enzymes.

研究分野：天然物生成研究

キーワード：RNA-Seq解析 ネットワーク解析 相関解析 生成酵素遺伝子 除虫菊 ピレスリン アントシアニン
転写因子

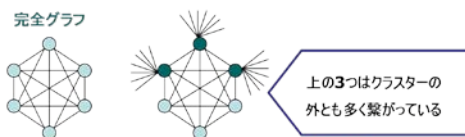
1. 研究開始当初の背景

(1) 網羅的な遺伝子発現 (トランスクリプトーム) データは次世代シーケンサーの RNA-Seq 解析で容易に得られ、植物由来の機能性代謝産物に関連する生合成酵素遺伝子情報が短時間で入手可能となった。RNA-Seq 解析は従来の DNA アレイ解析より比較的安価で行われることから、最も汎用されるトランスクリプトミクス解析として定着しつつある。モデル植物のシロイヌナズナでは、アントシアニン生合成の転写因子 (MYB75) 及びアブラナ科特有のグルコシノレート生合成の転写因子 (MYB28, 29) が遺伝子共発現ネットワーク解析より単離・機能解析され、遺伝子共発現ネットワーク解析が実用植物にも適用可能な強力なツールと成り得ることが示唆された。

(2) 我々は既に、相関係数の閾値で描画する遺伝子ネットワークの外側と内側の関係性 (トポロジー) を考慮しながら閾値を自動で設定する独自の金平糖アルゴリズム (図 1) を開発し、恣意的でないモジュール (共発現遺伝子グループ) が抽出可能で、大量の公開 DNA マイクロアレイを活用しながら利用者が簡単にアントシアニン及びグルコシノレート生合成のモジュールを得ることも可能な JAVA-GUI ソフトの開発を完成させる事を必要であると認識した。

(3) 金平糖アルゴリズムは、解析対象を点 (ノード) で表し、相関のある点同士を線 (エッジ) で結んで、ネットワーク状に関係性が表現される (図 1)。ソフトのプロトタイプは既に構築 [鈴木・科研費基盤 C (研究課題番号: 23510272) の研究成果] したが、このソフトは DNA アレイ解析用に構築したソフトであり、RNA-Seq 解析用にチューニングが必要である。また、本申請課題では、転写因子の単離を目的としているので、転写因子及び生合成酵素遺伝子を網羅的に識別する表示機能の追加が必要となる。

・ どれだけ線が繋がっているか (ネットワーク密度)



★ どれだけ特異的なノードか (ネットワーク特異率)

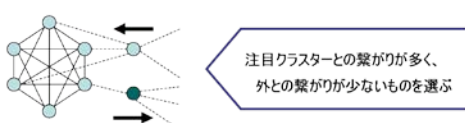


図 1 解析ソフトウェア ConfeitoGUIplus のアルゴリズム概略 (ネットワーク指標の導入)

(4) 近年、種々の分析機器のハイスループット化に伴い、いわゆるビッグデータと呼ば

れるものが氾濫している。これらのデータを有効活用するためには、多変量解析は最早必須であると言っても過言ではない。今後、さらに多くのビッグデータが生み出されることを想定すると、複数の分析機器のデータ、特に異種データの統合解析が重要な役割を担う時代が遠からず訪れるであろう。たとえば、トランスクリプトームとメタボロームのデータを同時に処理し、挙動の類似性に着目すれば、未知の代謝経路で働く遺伝子と代謝産物がペアで予測できるようになるかもしれない。来るべき時代に備え、異種データの統合解析を自在に行うためのプラットフォームの整備は急務であると言える。

(5) 以上の研究背景に基づき、我々はアルゴリズムのプロトタイプ (図 1) はすでに発表していたものの、今後起こり得るさまざまな用途 (トランスクリプトーム解析データの多様化、メタボローム解析への適用、異種データの利用拡大など) を想定し、求められる 3 つの特徴 (①入力データの自由度、②直感的なパラメーター設定、③許容幅のある関係性判定) を兼ね備えたソフトウェアへと改良することが必要であると実感した。

(6) 昆虫抵抗性物質であるピレスリンはキク科のシロバナムシヨケギク (*Tanacetum cinerariifolium*, 別名除虫菊) の胚珠及び蕾などに高濃度に蓄積している。第一菊酸 CoA (Chrysanthemoyl-CoA) とピレスロロン (Pyrethrolone) を基質とする鍵酵素ピレスリン合成酵素の遺伝子クローニングが成功し、ピレスリン生合成の酸部は非メバロン酸経路により生成し、アルコール部はオキシリピン経路により生成することが予想されている (図 2)。しかし、未だ遺伝子単離がされていないピレスロロン合成酵素 (P450) 及び Chrysanthemoyl CoA リガーゼの単離などピレスリン生合成の全貌が明らかにされていないのが現状である。

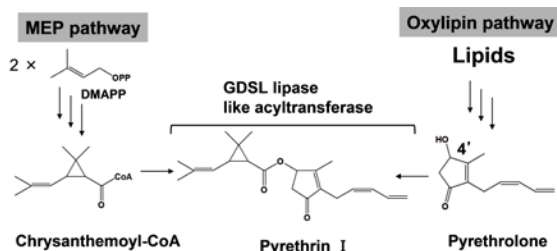


図 2 除虫菊でのピレスリン生合成経路 (鍵酵素ピレスリン合成酵素: GDSL lipase like acyltransferase の存在)

(7) ピレスリンの生産は気候変動、水の供給量、植物病原菌の感染等の自然環境の要因の影響を受ける。そのため、交配あるいは人為的な遺伝子改変によるピレスリン生合成の促進が必要である。しかし、ピレスリン生合成を支配する遺伝子に関する情報は極め

て限られているのが現状である。

(8) また、除虫菊幼苗に機械傷を与えると (E)-2-hexena, (Z)-3-hexenal, (Z)-3-hexen-1-ol, (Z)-3-hexen-1-yl acetate と (E)- β -farnesene を主成分とする揮発性有機化合物 (VOC) が放出され、それらを観測値に合わせて再ブレンドしたものを除虫菊が受容するとピレスリン合成を促進することが見出されている。その研究では合成に寄与するわずか4遺伝子だけを対象としており、係る合成促進作用が4遺伝子以外のピレスリン合成遺伝子群に対して、普遍的に見られるのかどうかは不明であった。

(9) 以上のことより、次世代シーケンスによる除虫菊の EST 整備 (RNA-Seq 解析) を開始し、実用植物における二次代謝産物の遺伝子機能解明にネットワーク解析を取り入れることの有益性を予見し、これを取り入れた新しい研究手法の道筋を示すため、本申請の着想に至った。

2. 研究の目的

RNA-Seq 解析から得られる遺伝子発現データに共発現ネットワーク解析を適用し、有用二次代謝生産植物の物質生産に関与する転写因子候補を効率的に見つけ出す JAVA-GUI ソフト (相関ネットワーク解析ソフト) の改良開発を行う。実証例として、除虫菊由来のピレスリン合成経路の全貌を明らかにするために、RNA-Seq 解析の整備及び遺伝子マイニングを BALST 検索及び相関ネットワーク解析などで実行する。得られた数種のピレスリン合成遺伝子発現に対して、傷害誘導的に生じる揮発性有機化合物 (VOC) のブレンド作用効果を調査する。

RNA-seq 解析を用いて二次代謝合成酵素遺伝子から関連転写因子を選出する最低解析数の統計学的検証するために、本申請課題で行った除虫菊由来の RNA-Seq 解析或いは公開データを用いて、本研究課題で改良された相関ネットワーク解析ソフトで実行する。

3. 研究の方法

(1) 次世代シーケンサーによる RNA-Seq 解析及び EST 情報整備

ピレスリンを蓄積するキク科植物の除虫菊を実用植物の例として採用し、組織・器官の違い、ピレスリン蓄積量の増減を示す種々の処理の有無等、由来の異なる RNA を取得し、RNA-Seq 解析データを獲得した。具体的には、葉、茎、根及び蕾の計4種の器官組織と匂い成分処理を施した脱分化したカルスとそのコントロールなど、合計24個の RNA サンプルの抽出を行い、HiSeq1500 Rapid 分析

(Paired-end 100bp) にて、FASTQ ファイルデータを取得した。市販の CLC Genomics Workbench ソフトを用いて、de novo assembly 及び各組織での遺伝子発現量解析を行った。得られた contig 配列を GO (Gene Ontology) 解析と Blast 検索により遺伝子の機能アノテーション作業を行った。

(2) 揮発性有機化合物 (VOC) 投与によるピレスリン合成遺伝子群の発現量の影響

無菌除虫菊幼苗及び脱分化カルスに対して、揮発性有機化合物 (VOC) を12時間与え、上記の RNA-Seq 解析で得られたピレスリン合成遺伝子群の発現量の増減を Real time-PCR にて調査した。

(3) 相関ネットワーク解析 (ConfeitoGUIplus) ソフトの再開発

① ネットワークの描画機能の強化、② 入力データの自由度、③ 直感的なパラメーター設定、④ 許容幅のある関係性判定) を兼ね備えたソフトウェアの改良を行った。

(4) 改良された (ConfeitoGUIplus) ソフトの検証

① 開発した金平糖 JAVA-GUI ソフトの動作検証を行うために、ネットワークのモジュール形成の適正を調べるためにネットワーク F 値を利用して、他のネットワーク手法 (Louvain 法、simulating annealing 法及び fast greedy 手法) と比較解析を行った。

② シロイヌナズナ由来の公開 DNA マイクロアレイ解析データセット「遺伝子数 (行): 22746 X サンプル数 (列): 9442」を用いて、金平糖ネットワーク解析を実行した。

4. 研究成果

(1) 次世代シーケンサーによる RNA-Seq 解析及び EST 情報整備

殺虫活性を示すエステル化合物ピレスリンを蓄積するキク科植物の除虫菊 (*Tanacetum cinerariifolium*) の葉、茎、根及び蕾の計4種の器官組織と匂い成分処理を施した脱分化したカルスとそのコントロールなど、合計24個の RNA サンプルの抽出を行い、HiSeq1500 Rapid 分析での更なる RNA-Seq 解析を実行した。ピレスリン合成経路の全容を明らかにする目的で、遺伝子共発現解析によって未知の合成関連遺伝子を同定しようと試みた。まず、遺伝子配列を de novo assemble ソフト Trinity によりおよそ42万個の contig を得た。これらの contig に対して、遺伝子アノテーション作業を行

い、各組織サンプルのリードをマッピングすることによって発現頻度情報を求めた。さらに代謝経路の共発現情報をもとに、この代謝経路上で香気成分によって誘導される遺伝子を探索した。その結果、酸化酵素を中心とするピレスリン合成遺伝子の候補を効率良く絞り込むことができた(図3)。以下にピレスリン合成経路(図3)に関する関連遺伝子リストを示す。

ACX1, acyl-Coenzyme A oxidase 1;
AOC, allene oxide cyclase;
AOS, allene oxide synthase,
CAS, chrysanthemic acid synthase;
CDS, chrysanthemyl diphosphate synthase;
CoA ligase, chrysanthemic acid:coenzymeA ligase;
HPL, hydroperoxide lyase,
13-LOX, 13-lipoxygenase;
OPC-8,3-oxo-2-(2-pentenyl)-cyclopentane-1-octanoic acid;
OPDA, 12-oxo-10,15-phytodienoic acid;
OPR3, OPDA reductase 3;
TcGLIP, Tanacetum cinerariifolium GDSL lipase.

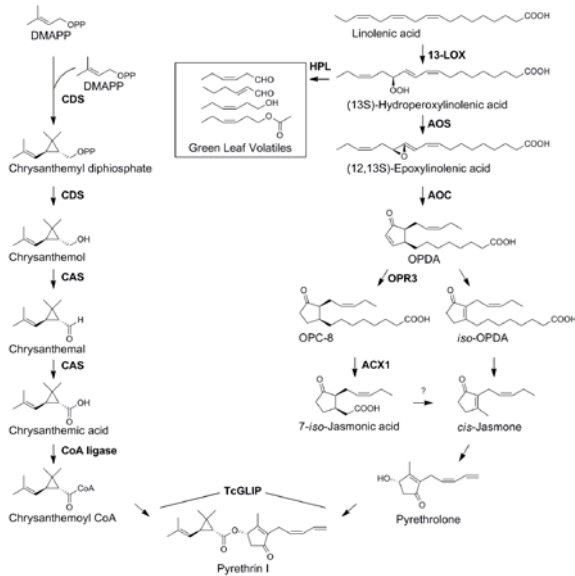
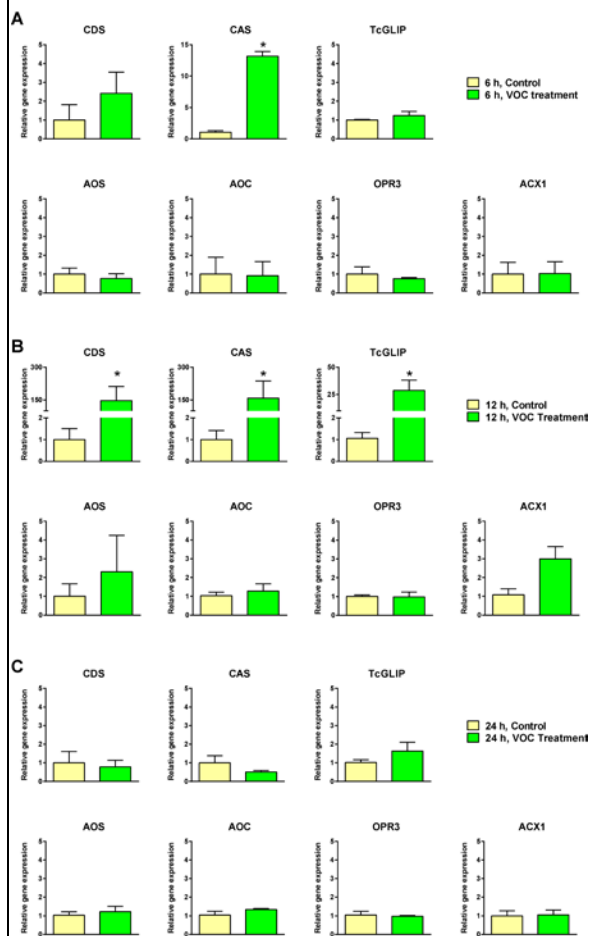


図3 RNA-Seq解析で得られた除虫菊由来のピレスリン合成経路及び関連遺伝子

(2) 揮発性有機化合物(VOC)投与によるピレスリン合成遺伝子群の発現量の影響

無菌除虫菊幼苗及び脱分化カルスに対して、以下の濃度で調整した揮発性有機化合物(VOC)「90.0nM (Z)-3-hexenal, 0.5 nM (E)-2-hexenal, 5.0 nM (Z)-3-hexen-1-ol, 6.0 nM (Z)-3-hexenyl acetate and 13.0 nM (E)-β-farnesene」6、12、24時間の暴露下での上記のRNA-Seq解析で得られたピレスリン合成遺伝子群の発現量の増減をReal time-PCRにて調査した(図4、5)。

図4 無菌除虫菊幼苗の葉でのピレスリン



合成遺伝子群の発現量の影響
A(6時間処理)、B(12時間処理)、C(24時間処理)

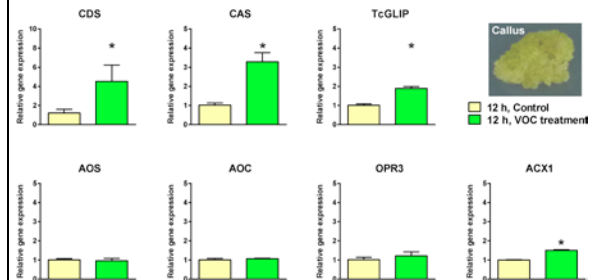


図5 無菌脱分化カルスでのピレスリン合成遺伝子群の発現量の影響

以上の結果(図4、5)より、本VOCブレンドは時間依存的に数種の遺伝子の発現を選択的に促進作用を示した。また、このようなVOCの作用は葉が脱分化することで効果が低下することも明らかにした。

(3) 相関ネットワーク解析
(ConfeitoGUIplus)ソフトの再開発

① 本研究では、ネットワーク描画機能を有する「ConfeitoGUIplus」と有しない「ConfeitoGUI」の両方を同時開発して、相關ネットワーク解析ソフト「ConfeitoGUIplus」の概略紹介のホームページを立ち上げた。

http://www.biosupport.kazusa.or.jp/sub_center3/index.php/confeitoguiplus/

本解析ソフトは、予めプログラム開発知識を必要とせず、ユーザーフレンドリーな開発(図6)に特化していることを受けて、現在、多くの植物二次代謝合成研究者に利用されるに至った。



図6 相關ネットワーク解析ソフト (ConfeitoGUIplus) の画面

本解析ソフトは FPO (False-positive-out) 解析と FNI (False-negative-in) 解析で構成

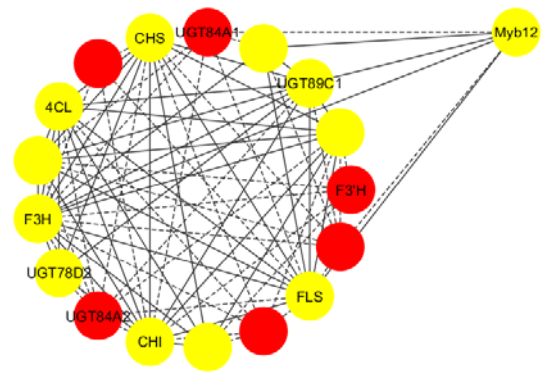
(4) 改良された (ConfeitoGUIplus) ソフトの検証

① 本解析ソフトは、ネットワークのサイズの類似性を表す F 値を用いて、他の解析手法と比較解析を行い、本解析手法がより高い F 値を示したことより、最適なネットワークのサイズを抽出するソフトであることが明らかとなった (表1)。

Method/Tool	Correlation coefficient	F-measure
ConfeitoGUI	-	0.605 ± 0.007
Louvain	0.97	0.465 ± 0.008
Simulating annealing	0.96	0.397 ± 0.008
Fast greedy	0.96	0.391 ± 0.008

表1 適切なネットワークサイズを抽出する為のネットワーク解析手法の比較検討

② 公開 DNA マイクロアレイ解析データセットを用いて、フラボノイド生合成モジュール抽出した。その結果、FNI 解析で転写因子 (Myb12) が選択されたフラボノイド生合成関連遺伝子群が全て単離 (図7) されて、本



解析ソフトの有益性が改めて、実証された。

図7 本解析ソフトの利用で得られたシロイヌナズナ由来のフラボノール生合成の転写因子 (MYB12) に関するネットワークモジュール形成

4CL, 4-coumaric acid: CoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3-hydroxylase; F3'H, flavonoid 3'-hydroxylase; FLS, flavonol synthase

③ 実用植物に遺伝子共発現ネットワーク解析を適用する場合に、9000 個近くのトランスクリプトームデータを取得することは現実的な数では無い。そこで、共発現モジュール形成と DNA アレイ実験数の検討を行うために、シロイヌナズナの DNA アレイデータを用い、ジヒドロフラボノール 4-還元酵素 (DFR) からアントシアニン生合成の転写因子 (MYB75) を選出する最低解析数の予備的な統計学的検証を行った (図8)。その結果、実用的には「10~40 数実験で注目酵素遺伝子から関連する転写因子が単離できる」ことが予想された。

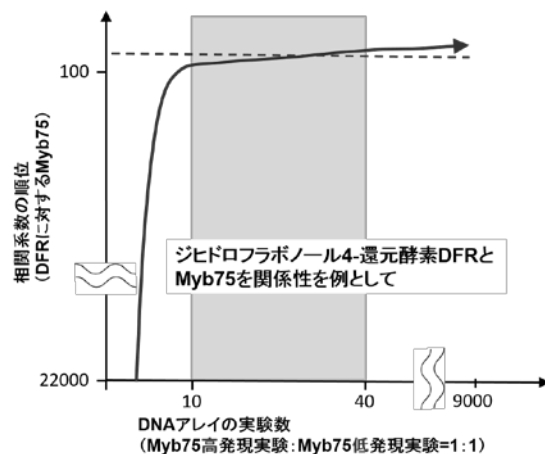


図8 二次代謝生合成酵素遺伝子から関連転写因子を選出する最低解析数の統計学的検証 DFR, Dihydroflavonol-4-reductase

以上、本研究では、有用二次代謝生産植物の物質生産に関する転写因子候補を効率的に見つけ出す有用な解析ソフトの開発に

成功した。普遍的に二次代謝生合成酵素遺伝子から関連転写因子を選出する最低解析数の統計的な提唱は未だ困難であるが、実験組み合わせを工夫すれば、50 個前後の網羅的な RNA-Seq 解析があれば、本解析ソフトを用いて、実装検証を行う価値があると思われる。

5. 主な発表論文等（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 1 件）

(1) Sakamori K, Ono N, Ihara M, Suzuki H, Matsuura H, Tanaka K, Ohta D, Kanaya S, Matsuda K
Selective regulation of pyrethrin biosynthesis by the specific blend of wound induced volatiles in *Tanacetum cinerariifolium*.
Plant Signal Behav., **11**(4): e1149675. (2016)
doi: 10.1080/15592324.2016.1149675.

〔学会発表〕（計 6 件）

(1) 萬年一斗、尾形善之、鈴木秀幸 植物代謝研究に役立つ関連ネットワーク解析（金平糖解析）第 68 回日本生物工学会大会シンポジウム 植物代謝工学研究最前線～新産業創出に向けて～ 2016 年 9 月 28-30 日（招待講演、富山国際会議場、富山）

(2) 鈴木秀幸 高精度・高分解能質量分析計を用いた植物由来機能性代謝研究ートランスオミックス解析に役立つ関連ネットワーク解析（金平糖解析）の紹介ー 質量分析フォーラム 2016（サーモフィッシャーサイエンティフィック主催） 2016 年 8 月 2 日、8 月 4 日（招待講演、ザ・リッツ・カールトン大阪、大阪：東京コンファレンスセンター・品川、東京）

(3) 萬年一斗、尾形善之、永島良樹、細内敦、柴田大輔、鈴木秀幸 植物オミックス研究における関連ネットワーク解析ソフト“金平糖 Java-GUI”の実用例 第 56 回日本植物生理学会年会 2015 年 3 月 17 日（口頭、東京農業大学・世田谷キャンパス、東京）

(4) 萬年一斗、尾形善之、鈴木秀幸 金平糖アルゴリズムによる新たなデータ解析手法の確立 第 67 回日本生物工学会大会 2015 年 10 月 26-28 日（ポスター、城山観光ホテル、鹿児島）

(5) 萬年一斗、尾形善之、鈴木秀幸 非モデル植物におけるトランスクリプトームデータの新たな活用法 ConfeitoGUI BMB2015（第 38 回日本分子生物学会年会、第 88 回日本生化学会大会）2015 年 12 月 1-4 日（ポスター、神戸国際会議場、兵庫）

(6) 鈴木秀幸（食品）機能性代謝研究に役

立つオミックス解析 千葉県バイオ・ライフサイエンス・ネットワーク会議 平成 26 年度シーズ発表会 2015 年 1 月 30 日（招待講演、ホテルグリーンタワー幕張、千葉）

〔図書〕（計 1 件）

(1) 萬年一斗、尾形善之、鈴木秀幸
プロジェクトバイオ：多変量解析に一石を投じるーConfeitoGUIplus の開発
「生物工学会誌」95-7 号 印刷中

〔産業財産権〕

出願状況（計 1 件）
名称：関連ネットワーク解析プログラム
発明者：鈴木秀幸、萬年一斗、柴田大輔、尾形善之
権利者：（公財）かずさ DNA 研究所、公立大学法人大阪府立大学
種類：特許出願
番号：2015-226403
出願年月日：H27 年 11 月 19 日
国内外の別：国内

〔その他〕

関連ネットワーク解析ソフト
「ConfeitoGUIplus」の概略紹介のホームページ
<http://www.biosupport.kazusa.or.jp/sub-center3/index.php/confeitoguiplus/>

6. 研究組織

(1) 研究代表者

鈴木 秀幸 (Suzuki Hideyuki)
（公財）かずさ DNA 研究所・バイオ研究開発部・機器分析グループ・グループ長
研究者番号：80276162

(2) 分担研究者

松田 一彦 (Matsuda Kazuhiko)
近畿大学・農学部・教授
研究者番号：00199796

(3) 分担研究者

尾形 善之 (Ogata Yoshiyuki)
大阪府立大学・大学院生命環境科学研究科・准教授
研究者番号：90446542

(4) 分担研究者

水谷 正治 (Mizutani Masaharu)
神戸大学・農学研究科・准教授
研究者番号：60303898

(5) 連携研究者

萬年 一斗 (Mannen Kazuto)
（公財）かずさ DNA 研究所・バイオ研究開発部・特任研究員
研究者番号：60744216