

平成 30 年 6 月 26 日現在

機関番号：82111

研究種目：挑戦的萌芽研究

研究期間：2016～2017

課題番号：16K14857

研究課題名(和文)インフォマティクス手法によるキー転写因子の推定法の確立

研究課題名(英文)Studies on estimation method of key transcription factors by bioinformatics

研究代表者

藤井 浩(FUJII, Hiroshi)

国立研究開発法人農業・食品産業技術総合研究機構・果樹茶業研究部門・主席研究員

研究者番号：00355398

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：カロテノイド代謝酵素遺伝子の発現制御には多数の転写因子が関与し、転写因子の関係性の統一的な理解が、代謝制御のキーとなる遺伝子の同定と成分育種に向けた代謝デザインにとって重要と考えられた。そこで、カロテノイド生合成系をモデルとして、カロテノイド代謝酵素遺伝子と転写因子遺伝子から構成される有向の遺伝子ネットワークを推定するソフトウェアSiGN-BNを用いて、遺伝子ネットワークを構成した。そして、構成された遺伝子ネットワークの最上流域に位置する遺伝子が代謝制御のキーとなる遺伝子として推定することができた。

研究成果の概要(英文)：A number of transcription factors are involved in the expression control of carotenoid metabolism enzyme genes. Understanding of the relationship of the transcription factor is important to identify key genes of control of metabolism and metabolism design toward component breeding. Using carotenoid metabolism as a model, gene network was configured. The gene network was comprised of carotenoid metabolism genes and transcription factors by SiGN-BN software that estimate directed gene network. The genes located in the uppermost stream were estimated as the key genes of metabolism control.

研究分野：バイオインフォマティクス

キーワード：遺伝子ネットワーク 代謝制御 転写因子 マイクロアレイ カロテノイド カンキツ

1. 研究開始当初の背景

(1) 申請者は、カンキツのカロテノイド代謝系の制御機構の解明に向けて、発現遺伝子を対象に、ゲノム解析基盤として、マイクロアレイ (Fujii et al. 2007, 2008) や、転写制御因子遺伝子領域の多型を対象とする DNA マーカー (Fujii et al. 2013)、連鎖地図 (Shimada et al. 2014) を開発してきた。そして、これらを用いて、7 種類のカロテノイド代謝酵素遺伝子の発現量を量的形質とした expression QTL 解析 (eQTL 解析) を行った。その結果、カロテノイド代謝酵素遺伝子の発現制御について、シス因子とトランス因子が複合して関与していることが明らかになった (Sugiyama et al. 2013)。トランス因子については、カンキツの交配集団 [興津 46 号 (スイートスプリング × トロピタオレンジ) × かんきつ中間母本農 6 号 (キングマンダリン × 無核紀州)] (AG 集団) の第 8 連鎖群下部に eQTL 座が集中していた (図 1)。この領域には、公開ゲノム情報の解析から、植物ホルモン、光、糖により発現が誘導される 30 個の転写制御因子遺伝子ホモログが集中して存在していた。

(2) 以上のことから、カロテノイド代謝酵素遺伝子の発現制御には多数の転写制御因子遺伝子が関与し、転写制御因子遺伝子の関係性の統一的な理解が、代謝制御のキーとなる遺伝子の同定と成分育種に向けた代謝デザインにとって重要と考えられた。

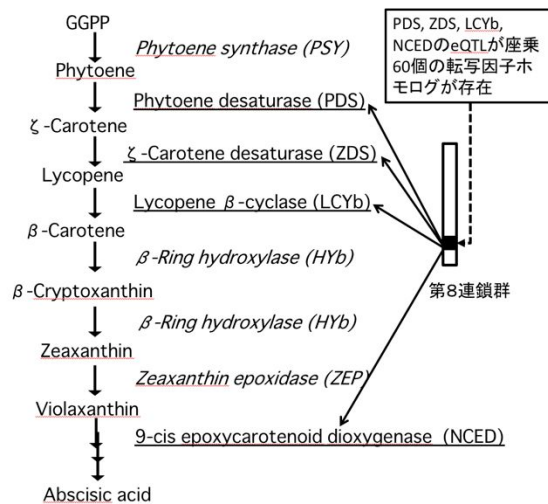


図1 カロテノイド代謝系と eQTL 解析結果

PSY、Hyb、ZEP 各遺伝子 (イタリックで示す) はシス因子によって、PDS、ZDS、LCYb、NCED 各遺伝子 (下線で示す) はトランス因子によって制御される

2. 研究の目的

(1) 本研究では、マイクロアレイデータを用いて、カロテノイド代謝酵素遺伝子をモデルとして、発現制御における転写制御因子候補群を包含する遺伝子ネットワークを推定

し、最上流に位置する転写因子を特定する新たなインフォマティクス手法の確立をめざす。遺伝子ネットワークの解析は果実の二次代謝制御を含めた生命現象の分子生物学的・生理学的な理解のために重要である。申請者は、果樹の二次代謝系のひとつであるカロテノイド代謝系の遺伝子発現制御機構に関する研究において、既存の遺伝統計学的手法を用いて代謝系内の各代謝酵素遺伝子の転写因子候補として 30 個の遺伝子の推定を行ってきた。しかし、遺伝統計学的手法では、転写因子候補の推定は可能であるが、転写因子間の転写制御の因果関係に関する情報を得られず、キーとなる転写因子を特定できない。

(2) カロテノイド代謝は植物ホルモン、光、糖、低温など、さまざまな要因に影響を受け、シグナルネットワークのクロストークが行われていると想定される。したがって、上記で特定した 30 個の転写因子候補は相互に制御関係にあると考えられるが、eQTL 解析では転写因子間の上位性などの関係を知ることができない。そこで、マイクロアレイ解析で得た遺伝子発現データを入力データとし、数理モデルのベイジアンネットワークモデル (図 2) を用いた遺伝子ネットワーク推定ソフトウェア SiGN-BN を用いて、上記の 30 個の転写因子候補を対象に含む遺伝子ネットワークを推定し、最上位に位置する転写因子を特定することにより、代謝系を制御するキー転写因子を特定する。ベイジアンネットワークモデルでは、ノード間の方向性も情報として含むので、ノード間の因果関係が推定できる (図 2)。本研究の場合、ベイジアンネットワークにおけるノードは遺伝子相当するので、遺伝子間の因果関係が推定できる。

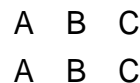


図2 ベイジアンネットワークモデル

ベイジアンネットワークモデルでは、A が起こってから B が起こり C になるのか (上)、B が起こってから、A と C が起こるのか (下) わかる。どちらのモデルかわかれば、因果関係が推定できる。

(3) 本研究により、インフォマティクス手法による果樹の二次代謝系の遺伝子ネットワークの推定法が確立され、カロテノイド代謝に関わる転写因子間の転写制御の因果関係が明らかになる。果樹における転写制御因子遺伝子研究は、シロイヌナズナやイネのようなモデル植物における転写因子のホモログを対象遺伝子と仮定して研究を進められることが多い。しかし、木本性の果樹などの場合、モデル植物と転写制御のバックグラウンドが大きく異なることが想定される。本研究

で確立される手法を用いることにより、これまで想定されなかった果樹特有の転写制御ネットワークや転写制御因子遺伝子を提示可能となる。

3. 研究の方法

(1) 遺伝子ネットワーク解析の入力データ取得に用いたマイクロアレイプラットフォーム

遺伝子ネットワーク解析ソフトウェア SiGN-BN の入力データとして、マイクロアレイ遺伝子発現データが必要である。対象とするマイクロアレイプラットフォームとして、申請者が独自に開発したカンキツ全遺伝子マイクロアレイを用いた。このマイクロアレイは、カンキツ国際ゲノムコンソーシアムにより公表されたクレメンティのドラフトゲノムデータ (<https://www.citrusgenomedb.org>) を利用して、アジレント社のマイクロアレイ設計システム eArray (<https://earray.chem.agilent.com/earray/>) を用いて設計した約 3 万の遺伝子を搭載するカンキツの全遺伝子アレイである。

(2) 遺伝子ネットワーク解析の入力データとしての既存のマイクロアレイデータ

表 1 既存アレイの培養さじょうに対する処理と培養期間 (単位は週)

品種 / 処理	暗黒	LED照射	AS6処理	ABZ処理	ABA処理
宮川早生	2,4,6	2,4	2,4	2,4	2,4
バレンシア	2,4,6	2,4			

上記のプラットフォームを用いた既存のマイクロアレイデータとして、3 シリーズ 80 サンプルのデータがある。このうち、56 サンプルはカロテノイド含有量が分離する AG 集団 56 個体の成熟果実の果肉を対象とした遺伝子発現データである。16 サンプルは、宮川早生 (*Citrus unshiu* Marc.) とバレンシアオレンジ (*C. sinensis* Osbeck) の培養さじょうを対象に、暗黒条件と LED 照射の光条件処理および植物ホルモン処理によるさじょうの遺伝子発現データである (表 1)。また、8 サンプルは、農研機構果樹茶業研究部門カンキツ研究興津拠点圃場に植栽されている宮川早生と -シトラウリンが蓄積するために果皮が赤くなる品種である山下紅を 9 月 8 日と、10 月 14 日、12 月 9 日に収穫して得たさじょうと、12 月 9 日に収穫して得た果皮から抽出したトータル RNA を対象に、マイクロアレイを用いて遺伝子発現データを取得した。

(3) 遺伝子ネットワーク解析の入力データ：新規のマイクロアレイデータ

上記のプラットフォームを用いて、新たに 32 サンプルのマイクロアレイ解析を行った。これは、SiGN-BN で遺伝子ネットワーク解析を行うためには、少なくとも 100 サンプルの

マイクロアレイデータが必要とされているためである。マイクロアレイ解析には、宮川早生とポンカン (*C. reticulata* Blanco) の培養さじょうを用いた。培養さじょうを用いたのは、遺伝子発現における環境要因を排除するためである。さじょうは、寒天培地で培養し、定温槽の中で温度管理をした。宮川早生さじょうに、植物ホルモンであるジベレリン、オーキシシン、エチレン、ブラシノライド、サリチル酸、ジャスモン酸、サイトカイニン処理をした。また低温と高温の処理区を設けた。これらを、24 時間、48 時間、72 時間静置し、トータル RNA と抽出してマイクロアレイによる遺伝子発現解析を行った (表 2)。

(4) マイクロアレイデータを入力データとする SiGN-BN プログラムによる遺伝子ネットワーク推定

遺伝子ネットワークの推定には、遺伝子ネットワーク解析ソフトウェア SiGN-BN を用いた。入力データには、(2) および (3) で示した合計 112 サンプルのマイクロアレイデータを用いた。SiGN-BN は計算時間とメモリの消費量の関係から、同時に投入できる遺伝子数が 1000 個程度に限度される。このため、入力する遺伝子数を絞り込む必要がある。そこで、クレメンティのドラフトゲノムデータにおいてカロテノイド生成に関連するとアノテーションされている 218 遺伝子 (A 群) と eQTL 解析によってカロテノイド代謝酵素遺伝子のトランス因子とされている 30 個の転写制御因子遺伝子候補 (B 群) およびクレメンティのドラフトゲノムデータにおいて転写制御因子遺伝子とアノテーションされている 2023 個の転写制御因子遺伝子候補 (C 群) の合計 2271 個の遺伝子を解析対象とした。そして、A 群と B 群の 248 遺伝子を必須遺伝子群として、合計が 1000 個になるように C 群から遺伝子を補完して、解析計算用遺伝子サブセットを作成した。遺伝子サブセットの作成にあたり、すべての遺伝子が有向グラフにおいて総当たりとなるように

表 2 新規アレイの培養さじょうに対する処理と培養期間

品種	処理	処理時間 (時間)
ポンカン	無処理 (20)	0
宮川早生	無処理 (20)	0, 24, 48, 72
宮川早生	ジベレリン (20)	24, 48, 72
宮川早生	オーキシシン (20)	24, 48, 72
宮川早生	エチレン (20)	24, 48, 72
宮川早生	ブラシノライド (20)	24, 48, 72
宮川早生	サリチル酸 (20)	24, 48, 72
宮川早生	ジャスモン酸 (20)	24, 48, 72
宮川早生	サイトカイニン (20)	24, 48, 72
宮川早生	低温 (10)	24, 48, 72
宮川早生	高温 (30)	24, 48, 72

複数の遺伝子サブセットを作成し、SiGN-BNに入力して遺伝子ネットワーク解析を行った。

4. 研究成果

(1) ホルモン処理をおこなったカンキツ培養さじょうを対象としたマイクロアレイデータ解析

宮川早生培養さじょうに、植物ホルモンであるジベレリン、オーキシシン、エチレン、ブラシノライド、サリチル酸、ジャスモン酸、サイトカイニン処理および低温と高温の処理を行い、24時間、48時間、72時間後のさじょうから取り出したトータルRNAを対象にマイクロアレイ解析を行い、32サンプルの遺伝子発現データを得た。80サンプルの既存のマイクロアレイデータとともに、合計112サンプルのマイクロアレイ発現データを遺伝子ネットワーク解析ソフトウェアSiGN-BNへの入力データとして準備することができた。

(2) ペイジアンネットワークモデルを利用したSiGN-BNによる遺伝子ネットワーク解析

カロテノイド生合成と関連するとアノテーションされている218遺伝子とeQTL解析によってカロテノイド代謝酵素遺伝子のトランス因子とされている30個の転写制御因子遺伝子の248個を遺伝子として、2023個の転写制御因子遺伝子から構成をずらしながら800個を選び、計4個の1048個からなるサブセットを作成した。これらのサブセットを順に、SiGN-BNに投入して遺伝子ネットワーク解析を行った。

その結果、それぞれの1048個の遺伝子から構成される遺伝子ネットワークを得ることができた。4つのサブセットで構成された遺伝子ネットワークの中に、子のノードを持たない遺伝子が300個存在した。この300個を排除して、同様のサブセットを再構成し、SiGN-BNによる計算を繰返した。こうした計算を繰返し、子のノードを持たない遺伝子の

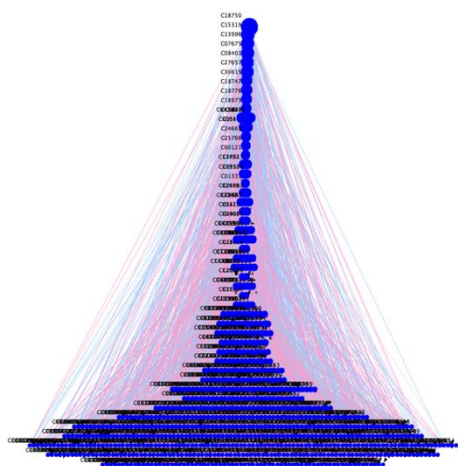


図3 カロテノイド代謝関連遺伝子と転写制御因子遺伝子の合計804遺伝子から構成される遺伝子ネットワーク

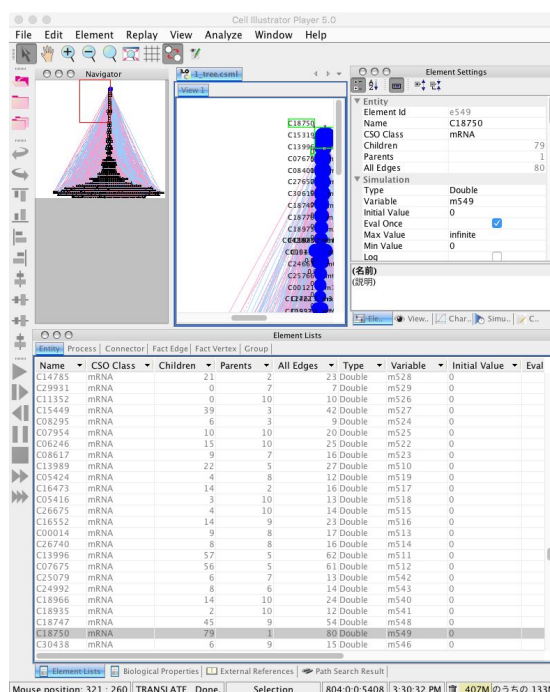


図4 Cell Illustrator Player による遺伝子ネットワークの可視化と各遺伝子の上流・下流関係および遺伝子ネットワークの解析

排除を進めた結果、4回目に804個の遺伝子からなる遺伝子ネットワークを得ることができた(図3)。このネットワークはカロテノイド代謝系を制御する転写因子のネットワークを現していると推定された。

(3) 推定された遺伝子ネットワークの評価と最上流に位置する転写因子の特定

上記の(2)で構成された遺伝子ネットワークを遺伝子ネットワーク可視化ソフトウェアであるCell Illustrator Player 5.0 (<https://cilonline.hgc.jp/cionlineserver/apps/usersman/main>)で可視化して解析を行った。その結果、4回目の計算で得られた804の遺伝子から構成される遺伝子ネットワークの最上流に位置する遺伝子はC18750であることが判明した(図4)。この遺伝子は、カロテノイド生合成系を制御するキー遺伝子である可能性がある。今後、米国の国立生物工学情報センター(NCBI)が提供しているマイクロアレイやRNA-Seqの遺伝子発現データを蓄積しているGenome Expression Omnibus(GEO)データベースを利用した機能解析を行うとともに、実験研究者と協力して実験的に機能解析を進める

(4) 各カロテノイド代謝酵素遺伝子の上流に位置する転写制御因子遺伝子候補の特定

推定された遺伝子ネットワークにおいて、カロテノイド代謝系の各酵素遺伝子であるフィトエン・シンテース遺伝子、フィトエン・デサチュレース遺伝子、 β -カロテン・デサチュレース遺伝子、リコペン-サイクラゼ遺伝子、 β -リング・ヒドロラーゼ遺

伝子、ゼアキサンチン・エポキシダーゼ遺伝子、9-シス-エポキシカロテノイド・ジオキシダーゼ遺伝子の下流に位置する転写制御因子遺伝子候補を特定することができた。

このうち、カロテノイドの含有量と遺伝子発現量が最も相関するフィトエン・システース遺伝子については、アップ・レギュレートに関連する4転写制御因子遺伝子候補(C13517, C28443, C14693, C28189)とダウン・レギュレートに関連する3転写制御因子遺伝子候補(C28456, C11177, C00506)、何らかの関係がある3遺伝子(C28369, C32746, C11726)が得られた。これらの中には、カロテノイド代謝酵素遺伝子のひとつであるリコペン・サイクラゼのホモログやアブシジン酸に関連する遺伝子ホモログが含まれていた。カロテノイド代謝系の下流は、アブシジン酸の合成と関係しており、この結果は、SiGN-BNの遺伝子ネットワーク推定の妥当性を示唆していると考えられた。

また、カロテノイド代謝系の下流に位置し、カロテノイドの分解に関わることで、総カロテノイド含有量に関係する9-シス-エポキシカロテノイド・ジオキシダーゼ遺伝子については、アップ・レギュレートに関連する3転写制御因子遺伝子候補(C19782, C15684, C11288)とダウン・レギュレートに関連する4転写制御因子遺伝子候補(C25033, C07161, C30718, C15845)何らかの関係がある3遺伝子(C11532, C20210, C25925)が得られた。この中には、カロテノイド代謝酵素遺伝子のひとつであるゼアキサンチン・エポキシダーゼ遺伝子ホモログが含まれていた。

カロテノイド代謝酵素遺伝子とは別に、すでにカロテノイド代謝系の上流域におけるカロテノイド代謝酵素遺伝子の制御に関連することが知られている転写制御因子遺伝子 Tf0271 遺伝子についても、遺伝子ネットワークにおける上流と下流の遺伝子について調査した。その結果、Tf0271 遺伝子のアップ・レギュレートに関連する1転写制御因子遺伝子候補(C25178)と、何らかの関係がある1遺伝子(C18750)が得られた。

また、Tf0271 遺伝子の下流に位置して、Tf0271 遺伝子がアップ・レギュレートする遺伝子として14遺伝子(C30832, C04461, C08617, C18747, C06943, C27613, C14807, C15951, C27619, C28603, C01530, C25577, C11134, C11312)が検出された。さらに、Tf0271 遺伝子がダウン・レギュレートする転写制御因子遺伝子候補として14遺伝子(C26387, C20139, C13910, C13908, C26675, C18402, C31390, C00742, C28775, C27181, C20064, C29007, C09824, C28391)何らかの関係がある遺伝子として3遺伝子(C16473, C15449, C02305)が得られた。Tf0271 遺伝子がアップ・レギュレートする遺伝子の中には、カロテノイド代謝酵素遺伝子のひとつとして、カロテン・イソメラーゼ遺伝子ホモログが含まれていた。また、何らかの関係がある

遺伝子の中には、同じくカロテノイド代謝酵素遺伝子のひとつであるフィトエン・デサチュレース遺伝子ホモログが含まれていた。この結果からも、本解析方法の妥当性の一端が示された。

これらの遺伝子についても、NCBI が提供しているマイクロアレイやRNA-Seqの遺伝子発現データを蓄積しているGEOデータベースを利用した機能解析を行うとともに、実験研究者と協力して実験的に機能解析を進める。

<引用文献>

Fujii H, Shimada T, Sugiyama A, Endo T, Nishikawa F, Nakano M, Ikoma Y, Shimizu T and Omura M、Profiling ethylene - responsive genes in mature mandarin fruit using a citrus 22K oligoarray、173 巻、2007、340 348

Fujii H, Shimada T, Sugiyama A, Nishikawa F, Endo T, Nakano M, Ikoma Y. and Omura M、Profiling gibberellin (GA3) - responsive genes in mature mandarin fruit using a citrus 22K oligoarray、Scientia Horticulturae、116 巻、2008、291 298

Fujii H, Shimada T, Nonaka K, Kita M, Kuniga T, Endo T, Ikoma Y and Omura M、High - throughput genotyping in citrus accessions using an SNP genotyping array、Tree Genetics & Genomes 9 巻、2013、45 153

Shimada T, Fujii H, Endo T, Ueda T, Sugiyama A, Nakano M, Kita M, Yoshioka T, Shimizu T and Omura M.、Construction of a citrus framework genetic map anchored by 708 gene-based markers、Tree Genetics & Genomes、10 巻、2014、1001 1013

Sugiyama A, Omura M, Shimada T, Fujii H, Endo T, Shimizu T, Nesumi H and Ikoma Y、Expression Quantitative Trait Loci Analysis of Carotenoid Metabolism-related Genes in Citrus、Japanese Society for Horticultural Science、83 巻、2014、32-43

5 . 主な発表論文等

[雑誌論文](計1件)

Sugiyama A, Ikoma Y, Fujii H, Endo T, Nesumi H, Shimada T and Omura M、Allelic diversity of phytoene synthase gene influences the transcription level in citrus fruit among a citrus F1 hybrid population、Breeding Science、67 巻、2017、382 392、査読有

6 . 研究組織

(1)研究代表者

藤井 浩(FUJII, Hiroshi)

国立研究開発法人農業・食品産業技術総合研究機構・果樹茶業研究部門・主席研究員
研究者番号：00355398