

## 科学研究費助成事業 研究成果報告書

平成 29 年 6 月 5 日現在

機関番号：63801

研究種目：基盤研究(A) (一般)

研究期間：2013～2015

課題番号：25252005

研究課題名(和文) GWASを用いた野生イネの生育、代謝形質と遺伝変異の相関解析

研究課題名(英文) Analysis of relationship between genetic variation and growth/phenotypic characters of wild *Oryza* accessions using GWAS

研究代表者

倉田 のり (KURATA, Nori)

国立遺伝学研究所・系統生物研究センター・名誉教授

研究者番号：90178088

交付決定額(研究期間全体)：(直接経費) 33,200,000円

研究成果の概要(和文)： 遺伝的解析の困難な野生イネ *Oryza rufipogon* 約300系統を用いて、代謝産物や出穂期・種子重量などの形態形質と各系統のゲノム配列変異との相関関係をGWAS(ゲノム全長相関解析)で調べた。代謝産物の解析では、比較のため交雑組み合わせ後代で量的遺伝子座(QTL)解析も行った。QTLで検出された12の有意な候補代謝産物のうち、最も高い値を示したトリゴネリンについてGWASの結果を精査した結果、機能未知遺伝子でストレス応答性に関連した遺伝子が責任遺伝子として特定できた。他の形態形質については、責任遺伝子探索のための一塩基変異検索データベースを作成し、候補領域検索を続行中である。

研究成果の概要(英文)： Using above 300 accessions of wild *Oryza* species *O. rufipogon*, we have investigated relationships between several phenotypic characters such as metabolite, heading date and grain weight, and genetic variation by GWAS analysis. QTL analysis using the crossed population of two strains has also been used for metabolite analysis. The metabolite trigonellin showed the highest QTL peak among 12 positive QTLs of above LOD score 6.0. The GWAS showed a large peak on the same region of the genome and identified two responsible SNPs in the unknown but stress reactive gene as well. To extract candidate SNPs for other phenotypes, we have constructed the SNP-searching database of 400 accessions and continue to detect responsible SNPs.

研究分野：植物遺伝育種科学

キーワード：野生イネ遺伝資源 ゲノムワイドアソシエーション解析 メタボローム解析

## 様式 C-19、F-19、Z-19 (共通)

### 1. 研究開始当初の背景

ゲノム解析分野では、イネの基本栽培種ゲノムの解読が終了し、イネにおける多様な遺伝資源の迅速なリシーケンスも可能になってきた。系統間の遺伝的多様性を1塩基多型(SNP)情報として取り出し、定量的な表現型とSNPとの相関関係を解析するGWAS解析も実現している。我々のグループは他に先駆けて多数の栽培イネおよび野生イネ系統群を用いて、ゲノム全域にわたるSNP構造解析と系統進化解析を行い、論争の的であった栽培系統の系統と栽培起源地を明らかにした。このように多数の野生イネ(*O. rufipogon*)を用いた解析は初めてであり、これらのゲノムの遺伝型情報を公開すれば貴重なリソースとなる。

一方、植物メタボローム解析の分野では、20万を超えとも言われる植物代謝産物中の有用機能物質等の探索や、代謝経路の解明を目指しており、代謝物検出機器における進展も相まって、種々の植物でメタボローム解析が進んでいる。この解析を、多様性の大きな野生集団で行い、新規な代謝産物とSNP構造との相関解析により、責任遺伝子候補を探す試みは未だ行われていない。

また、GWAS解析は、様々な栽培イネ系統群やアラビドプシスのエコタイプ系統等を用いて進められているが、いまだ解析精度や形質の種類にも限りがあり、検出力を向上させるための改善や工夫の余地は大きい。

国立遺伝学研究所には、世界各地から収集した栽培イネの祖先野生種 *O. rufipogon* の多様な系統群が保存されており、収集した国や地域、環境が大きく異なる野生集団から採種されており、大きな遺伝的多様性を含んでいる。生息地の環境に適応した多様な遺伝型と形質を比較解析することにより、多くの有用な遺伝子資源を同定、活用できると考えられる。

### 2. 研究の目的

野生イネ *O. sativa* の近縁野生種 *O. rufipogon* は、地球の様々な気候変動に適応して数十万年を生き延び、乾燥、高温、低温、病害虫などの各種ストレス耐性など優れた性質を多く保有する貴重な遺伝資源である。これまでこれら野生イネ遺伝資源は栽培イネの品種改良にほとんど利用されてこなかったが、近年解析技術の進歩により、野生イネが持つ多様な形質について遺伝解析が可能になりつつある。

本研究では、*O. rufipogon* の多数系統集団から、代謝産物量や種子形質・育成特性などの生態型を支配する遺伝子資源を同定する。GWAS (ゲノムワイドアソシエーション解析) により各形質に関連する遺伝子座

を絞り込み、関与遺伝子の同定と構造解析を行う。

野生植物は、様々な環境に適応するために特徴的な代謝産物を生合成していると考えられるため、まずメタボローム解析を試みる。野生イネでは栽培種イネの代謝プロファイルと大きく異なると考えられ、特定代謝産物の量的形質座と候補遺伝子が同定できれば、代謝育種に利用可能な画期的な育種母本候補の選抜やゲノミックセレクションを用いた高効率な育種が可能になる。さらに、形態形質 (出穂期、種子稔性、種子重量、花粉稔性、再生能力 (1年生/多年生)) についても、野生イネ (*Oryza rufipogon*) 自然集団から400系統を用いて形質調査を実施し、形質データを取得してGWASを行う。解析結果からこれら有用遺伝子資源の持つ特性の多様性を明らかにすると共に、代謝物、生育特性の育種への利用を図ることを目的とする。

### 3. 研究の方法

すでに本解析に用いる予定の *O. rufipogon* 系統及び交雑解析系統については、次世代シーケンサーによる各ゲノムのうす読みの解読を終え、系統毎のゲノム全域のSNP情報を得ており、ゲノム情報はこれを用いる。また、野生イネは種子繁殖能力の差が大きく、完全脱粒性のため種子収穫が困難な系統も多いが、メタボローム解析及び形質測定のために必要な種子は毎年度収穫を行い、実験に供する。これらを野生イネ生育条件に近い中国海南島において栽培し (メタボローム解析種子の一部は国内栽培)、計画期間3年間で以下に示した研究を実施した。

(1) メタボローム解析: 自然集団の中から400系統の *O. rufipogon* を選び材料とした。また、野生種イネ (*Oryza rufipogon*) と栽培種イネ (*Oryza sativa indica*) の戻し交雑集団の210系統も用いた。代謝産物については、理化学研究所において、戻し交雑集団210系統では1回、野生種イネ自然集団では3年間で3反復のメタボローム解析を行い、特定の代謝産物と顕著な関連性を示す、量的形質座及び候補遺伝子を探索した。各ステップには以下の手法を用いた。

メタボローム解析: 野生種イネ、交雑系統種子は系統毎に粉末化し、80% MeOH で代謝産物を抽出した後にLC-MS/MSを用いたワイドターゲットメタボロミクスを行った。ワイドターゲット解析では、予め標準化合物で最適化した高感度検出条件492種を利用し、連鎖解析系統群 (210系統) とアソシエーションマッピング系統群 (平成25,26,27年の3年間の反復で共通に種子

を使用できたのは 247 系統) の代謝産物量を調べた。

連鎖解析: メタボロームデータは、抽出溶媒のみの測定結果をノイズ、サンプル抽出物の結果をシグナルとして各代謝産物のシグナルノイズ比が 5 以上の代謝産物を 172 種選択した。このメタボロームデータと 1,793 個の多型マーカー情報を利用し、統計ソフト R の R/qtl パッケージを利用して連鎖解析を行った。

メタボロームのゲノムワイドアソシエーション解析 (GWAS): GWAS でも上記のシグナルノイズ比で代謝産物を選択した後に、日本国内で栽培したラインと中国海南島で育てたラインで共通に検出された 77 化合物を選択し、2,463,549 サイトの SNP 情報を統計ソフト R の rrBLUP パッケージを利用して GWAS を行った。

(2) 形態形質のアソシエーションマッピング: GWAS 解析集団として育成した野生種イネ *O. rufipogon* の自然集団 400 系統は、毎年 400 系統 X 各 3~5 個体を中国海南島において 3 年間育成し、出穂期、種子稔性、種子重量、花粉稔性、再生能力 (1 年生/多年生) についての形質調査を実施し、形質データを取得した。発芽不全や天候不良などによる欠失系統や、形質によるデータ取得のバラツキは大きい、毎年 300-350 の系統数のデータを収集できた。

形質のゲノムワイドアソシエーション解析 (GWAS): 形態形質は多くの遺伝子座に支配される量的形質がほとんどで、遺伝的背景がヘテロな自然集団では明確な責任遺伝子座を同定することは難しい。また GWAS に用いる統計処理の方法によっても、観察できるピーク数が相互に一致しない場合もあり、GWAS 解析手法の改良および調整を試みた。SNP コール法の改善、欠測した遺伝子型の補完やヘテロ SNP の考慮によるゲノム全体の遺伝型の調整を行ない、GWAS 解析に適用した。解析はメタボローム同様に R の rrBLUP パッケージを利用した。

SNP ビューワーの作成: メタボローム及び形態形質の GWAS 解析で、代謝産物や形質と責任 SNP の相関解析を行うにあたり、GWAS ピークの候補領域中のすべての遺伝子座や SNP との相関を検出するには膨大な手間が必要である。よって候補領域の SNP 情報及び遺伝子の構造変異等を簡単に検出できる、ゲノムビューワーを作成し、系統毎の SNP 情報と紐付けできるツールを作成した。

#### 4. 研究成果

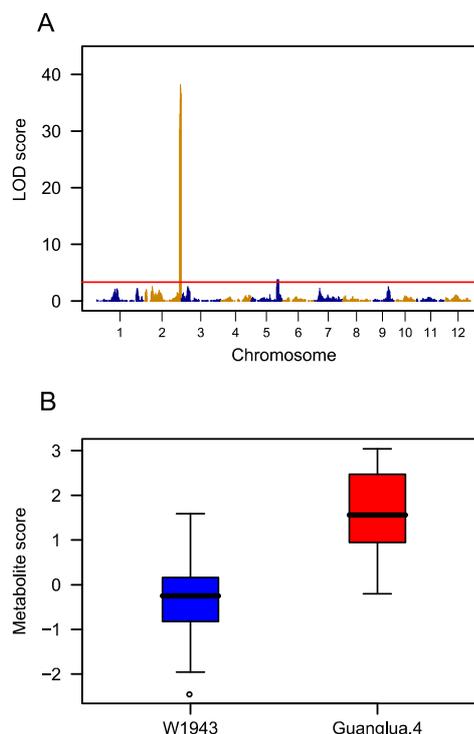
##### (1) メタボローム解析 & 代謝産物の連鎖解析

野生種イネ (W1943) と栽培種イネ (Guanglua.4 インディカ) の連鎖解析では、likelihood of odds (LOD) score が 6 以上のメタボリック QTL (mQTL) が 12 個同定できた (Table 1)。

**Table 1.** Summary of metabolome linkage mapping.

Metabolite annotation	Chromosome	Position	LOD score
Isocitrate	1	17	7
L-Allothreonine	1	5	6
AMP	2	49	6
Trigonellin	2	77	3 6
Fusaric acid	3	14	1 0
L-Glutamine	3	13	8
Methionine sulfoxide	3	20	6
Pantothenate	3	50	7
Guanosine	6	58	6
Isopentenyl adenosine	7	27	7
Betaine aldehyde	7	12	8
L-Aspartic acid		27	8

この中で、もっと高い関連性を示したトリゴネリンでは、2 番染色体 77 番で LOD core 36 を示した (Fig. 1A)。



**Fig. 1.** Linkage mapping of trigonellin.

さらに、このマーカーの遺伝型（野生種型と栽培種型）について箱ひげ図で比較した結果、トリゴネリンは栽培種で多いことがわかった（Fig. 1B）。このトリゴネリンに特徴的な mQTL の候補遺伝子を探索するためにメタボリック GWAS（mGWAS）の結果を検証した。

#### mGWAS

mGWAS では、100 か所以上の mQTL が検出されたが、これらの確度は不明である。そこで、連鎖解析で同定したトリゴネリンの mQTL 周辺の遺伝子候補を調べた。この結果、mGWAS では、2 番染色体に顕著な関連性を示す SNP が 2 個存在し（Fig. 2A）、機能未知の遺伝子 OS02G0797700 のコード領域と 5' 側の非翻訳領域（転写調節領域）に位置することがわかった（Fig. 2B）。これらのハプロタイプごとに箱ひげ図で解析した結果、いずれの SNP もマイナーハプロタイプでトリゴネリン量が増加していることがわかった（Fig. 2C）。OS02G0797700 のアノテーションは、putative endomembrane protein (emp70) であり現時点でトリゴネリンの量を制御する分子メカニズムは不明である。

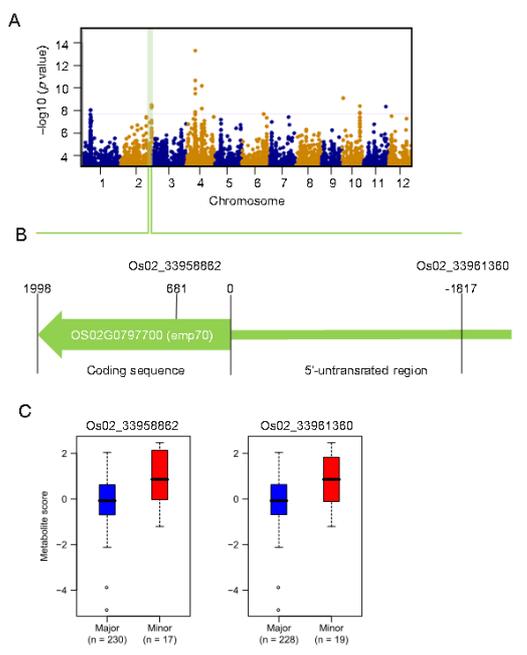


Fig. 2. Association mapping of trigonellin.

#### トリゴネリンの育種利用の検討

トリゴネリンは4級アミンを含むピリジンアルカロイドであり、古くはニコチン酸の貯蔵物質と考えられていたが、先行研究ではストレス応答性（酸化、UV、塩）や細胞周期において機能することがわかっている。さらに、マメ科植物では根粒形成や休眠運動にかかわることも知られており、植物に特徴的な生理機能が予測されている。本研究では、トリゴネリンの量が特定の

mQTL で制御され、栽培種でも存在することが確認されたことから、今後トリゴネリンを標的とした代謝育種を行うことで各種のストレス耐性を増強したイネの育種が可能になると期待される。

現在、上記 ~ の結果に関して論文執筆中である。

#### (2) 形態形質のアソシエーションマッピング

GWAS 解析： 野生イネ *O. rufipogon* の 400 系統から、3 年間にわたり、毎年 300-350 の系統からデータを収集し、SNP コールの改良やヘテロ遺伝子型の調整を加えて、GWAS 解析によりアソシエーションマッピングを行った。しかし、年次間の環境の変化が大きく、データのバラツキが大きかった。解析の結果、5 つの形質それぞれで p 値が  $10^{-6}$  を下回る SNP が多数（種子稔性の 7 から出穂期の 437）検出された。しかし関与が特定できるほど明瞭なピークを示した SNP はなく、またピークが検出された領域も数十-100Kb と幅があり、このままの状態では責任遺伝子を特定することは困難であった。そこで、示すツールを用いて、現在責任 SNP の同定を進めており、特定できた遺伝子については、系統間の遺伝子構造多様性を解析し、形質発現との相関等も確認する。これらの結果から、新規な農業形質遺伝子とその対立遺伝子の発掘を行う。

#### SNP 検索ビューワーの作成：

アソシエーションマップされたピーク領域の中の候補 SNP で、実際に特徴的の形質を持つ系統と相関する SNP を特定できるか直接検証する必要がある。多数の系統で長い領域内の SNP を効率よく検証するため、レファレンスゲノム配列に沿って、遺伝子コード領域と系統毎の SNP 配列を記載した SNP ビューワーを作成した。すでに一部作成済みの栽培イネゲノム全域の配列と遺伝子アノテーションを示したデータベース（DB）OryzaGenome に野生イネゲノム全長の配列を重ね、用いた 400 全系統及び精密配列を持つ追加 50 系統の SNP を表示した。さらに SNP に起因するアミノ酸変異の有無も確認できるゲノムビューワー DB OryzaGenome2 を構築した。

同時に、この DB は、新たに比較解析したいイネゲノムの配列を検索サイトにロードすることにより SNP 検索もでき、本研究で用いた系統群の SNP との比較解析もできる形にした。また、本研究で調査した 5 形質及び独自に調査した 10 以上の形質について 3 年間のデータで GWAS 解析を行なった結果も示している。この DB は、<http://211.5.106.18/oryzagenome2detail/downloads/index.xhtml> のサイトで利用できるが、現在パスワード付きの未公開で

ある (Fig.3)。最終調整し、論文作成すると同時 2017 年中には公開予定である。

また、本研究で用いた *O. rufipogon* 400 系統のうち、国立遺伝学研究所が NBRP として保有する 300 系統については、リクエストに応じて種子の分譲を行う。

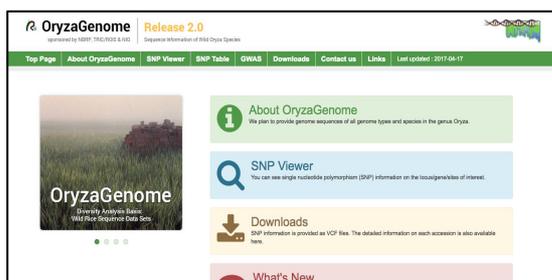


Fig. 3. Top page view of the OryzaGenome Release 2.

## 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10件)

Ohyanaki H, Ebata T, Huang X, Gong H, Fujita M, Mochizuki T, Toyoda A, Fujiyama A, Kamminuma E, Nakamura Y, Feng Q, Wang ZX, Han B, Kurata N. OryzaGenome: Genome Diversity Database of Wild Oryza Species. Plant Cell Physiol. 査読有. 57:e1, 2016. DOI: 10.1093/pcp/pcv171.

Shenton M, Iwamoto C, Kurata N, Ieko K. Effect of wild and cultivated rice genotypes on rhizosphere bacterial community composition. Rice(N Y). 9:42, 2016. 査読有 .DOI: 10.1186/s12284-016-0111-8.

Horiuchi Y, Harushima Y, Mochizuki T, Fujita M, Ohyanagi H, Kurata N. Global expression differences and tissue specific expression differences on rice evolution result in two contrasting types of differentially expressed genes. BMC Genomics. 16:1099, 2015. 査読有. DOI: 10.1186/s12864-015-2319-1.

Shenton M, Ohyanagi H, Wang Zx, Toyoda A, Fujiyama A, Nagata T, Feng Q, Han B, Kurata N. Rapid turnover of antimicrobial-type cysteine genes in closely related Oryza genomes. Mol Genet Genomics. 290:1753-70, 2015. 査読有. DOI:10.1007/s00438-015-1028-4.

〔学会発表〕(計9件)

Onogi A, Phenotypic diversity and genetic dissection of adaptive traits of wild rice (*Oryza rufipogon*). 13th International Symposium on Rice Functional Genomics (ISRFG). September 23, 2015, 華中農業大学 (中国武漢市)

倉田のり、野生イネ*Oryza*属200ゲノムの解読と多様性・系統関係の解析、日本育種学会第128回講演会、2015年9月11日、新潟大学(新潟県新潟市)  
澤田有司、メタボロミクスを利用した環境応答の解析と育種への応用利用、日本育種学会第126回講演会 第56回シンポジウム 作物環境応答の網羅的計測とモデリング~その育種への活用~、2014年9月26日、南九州大学(宮崎県都城市)

Sawada Y. Integrated metabolomics for large scale rice bioresources. Plant and Animal Genome XII. January 13, 2014, San Diego (アメリカ)

〔その他〕

本研究に使用した400系統(追加50系統含む)に関するゲノムSNP情報データベース

OryzaGenome: <http://viewer.shigen.info/oryzagenome/mapview/Top.do>

本研究で作成したOryzaGenome2:

<http://211.5.106.18/oryzagenome2etail/downloads/index.xhtmll>

(現在PW付き限定公開)

## 6. 研究組織

### (1)研究代表者

倉田 のり (KURATA, Nori)

国立遺伝学研究所・系統生物研究センター・名誉教授

研究者番号: 90178088

### (2)研究分担者

久保 貴彦 (KUBO, Takahiko)

国立遺伝学研究所・系統生物研究センター・助教

研究者番号: 00370148

平井 優美 (HIRAI, Masami)

国立研究開発法人理化学研究所・環境資源科学研究センター・チームリーダー

研究者番号: 90415274

### (3)連携研究者

澤田 有司 (SAWADA, Yuji)

国立研究開発法人理化学研究所・環境資源科学研究センター・研究員

研究者番号: 00415176

### (4)研究協力者

岩田 洋佳 (IWATA, Hiroyoshi)

東京大学農学生命科学研究科・准教授

研究者番号: 00355489

小野木 章雄 (ONOGI, Akio)

東京大学農学生命科学研究科・研究員  
研究者番号：60760501

鐘ヶ江 弘美(KAJIYA-KANEGAE, Hiromi)  
東京大学農学生命科学研究科・特任研究員  
研究者番号：30758470