

令和 5 年 6 月 30 日現在

機関番号：81202

研究種目：基盤研究(B) (一般)

研究期間：2020～2022

課題番号：20H02962

研究課題名(和文) イネNAM集団を用いた新規ゲノム予測手法による遺伝子同定とネットワーク解明

研究課題名(英文) Genome prediction method for identifying agronomic trait genes and gene networks using rice NAM population

研究代表者

阿部 陽 (Abe, Akira)

公益財団法人岩手生物工学研究センター・ゲノム育種研究部・主席研究員

研究者番号：80503606

交付決定額(研究期間全体)：(直接経費) 13,700,000円

研究成果の概要(和文)：交雑後代(RILs)を用いて遺伝子間相互作用エピスタシスを検出する新しいモデルを構築し、RIL-StEpとしてパッケージ化した。一穂粒数に関するFZP遺伝子とエピスタティックな関係にある領域をChr.11に検出し、候補遺伝子を同定した。FZPIは、候補遺伝子の遺伝子発現を抑制する働きを持つこと、候補遺伝子のプロモーター配列の多型がエピスタシスの要因である可能性を得た。NAM集団を用いて、遺伝子型から高精度に形質を予測できるゲノミック予測モデルを構築し、相関係数0.9以上の精度で形質を予測できた。加えて、複数品種のゲノムが混ざった系統においても、同様に高精度で形質を予測することができた。

研究成果の学術的意義や社会的意義

イネにおける量的形質の変動の原因となる遺伝子とそれらネットワークを理解することは重要な課題である。本研究の成果は、イネの収量性に直結する穂の着粒構造の遺伝的改良に寄与するものである。構築したゲノミック予測モデルは、極めて精度が高く、イネのみならず多様な作物におけるゲノミック選抜に波及することが期待できる。また、目的とする形質が予測されるゲノムをデザインし、そのゲノムを交配等で再現する「ゲノムデザイン育種」への端緒となるものである。

研究成果の概要(英文)：We developed a novel model to detect epistatic interaction using recombinant inbred lines (RILs) and released an R package RIL-StEp. We identified a region on Chr.11 that is associated with grain number per panicle and has an epistatic relationship with the FZP region. A candidate gene was identified by RNA-seq analysis. We found that FZP represses the expression of the candidate gene and identified the possibility that polymorphisms in the promoter sequences of the candidate gene are involved in the epistatic effect.

A genomic prediction model was constructed using the NAM population to accurately predict agronomic traits from genotypes. The model achieved high prediction accuracy (correlation coefficient of 0.9 or higher) for leaf width, number of grains per panicle, number of panicles, and grain size. Furthermore, the model was able to predict traits with high accuracy even in lines in which the genomes of multiple varieties had been introduced through crosses.

研究分野：Plant genetics, Plant breeding

キーワード：Genomic prediction rice NAM population Gene-Gene interaction G x E interaction Epistasis

## 1. 研究開始当初の背景

地球規模の気候変動による耕作可能地の減少、栽培適地の移動などが懸念される中、世界人口は20年以内に90億人に達すると予想され、世界人口の50%以上を支える重要な主食であるイネの収量増大を達成する必要がある (Godfray et al. 2010)。この目的のため、イネ系統が進化・栽培化の過程で蓄積してきた多様な自然変異から有用遺伝子・アリル及びその機能を解明し、ゲノム情報を駆使した作物育種に取り組む必要がある。

作物育種は、農業形質に関与する QTL や遺伝子の望ましいアリルを集積するプロセスである。これまでに、イネ多収性に寄与するとされる遺伝子の単離研究は多く実施されている。例えば、シンクサイズでは、玄米サイズに関与する *qSW5* (Shomura et al. 2008) や *GS3* (Mao et al. 2010)、籾数を増やす *Gn1a* (Ashikari et al. 2005)、*WFP* (Miura et al. 2010) や *FZP* (Bai et al. 2017) を始めとして多くの遺伝子単離の報告がある。これら遺伝子は、2品種間交雑後代を利用した QTL 解析等から、大きな効果を持つ主要遺伝子として単離されている。しかしながら、作物の形態・生理生態には多数のマイナー-QTL が関わっていることは明らかであり、量的形質の変動の原因となる遺伝子とそれらネットワークを理解することは、依然として重要な課題であると考える。

Yu ら (2008) がトウモロコシを材料に提唱した Nested Association Mapping (NAM) は、複数の Founder 系統 (既存品種) と一つの共通親との交雑から得られた複数組み合わせの数千系統の RILs を用いて、Genome Wide Association Study (GWAS) を行う手法である。NAM は、Founder のゲノムをシャッフルすることで連鎖不平衡がキャンセルされ統計的パワーが高いこと、GWAS と同様に多数のアリルを解析できること、数千系統が持つリコンビネーションと多数の SNPs による高解像度の Mapping ができることなど、従来の QTL 解析及び GWAS に比べてメリットが大きい (Yu et al. 2008, Tian et al. 2011)。研究代表者は、イネ品種「ひとめぼれ」を共通親とする 20 組合せ 3,000 RILs (F8~F10) のイネ NAM 集団を確立し、その全系統について全ゲノムシーケンスによる genotyping を完了している。

従来の GWAS や GP は、線形回帰モデルに当てはめて解析されることが多い。近年、様々な分野において機械学習の活用が注目されている。ゲノミクス分野においてもその応用が論じられており (Eraslan et al. 2019; Zou et al. 2019)、GWAS と機械学習を組合せたダイズ遺伝資源の解析からマイナー-QTL を同定した例も報告されている。機械学習モデルを応用することで、従来より精度の高い QTL 同定も可能になると考えられる。

## 2. 研究の目的

本研究では、独自にイネで確立した NAM 集団を用いて、農業上重要形質の発現に関わる遺伝子間相互作用 (エピスタシス) を解明し、メジャー-QTL の単独効果に着目した従来育種の限界を打破するための基礎的知見を得るため、次のことを目的とした。

- (1) 穂の着粒構造に関与する遺伝子群の同定とそれらの遺伝子間ネットワークを明らかにする。
- (2) 決定木勾配ブースティング法 (Gradient-boosted decision tree) や ElasticNet などの手法を用いた Genomic prediction を検討し、高精度なゲノミック予測モデルを構築する。

## 3. 研究の方法

(1) 一穂籾数に関する GWAS 結果から、効果の大きい Chr.7 の *FZP* に着目し、着粒構造に関与する遺伝子ネットワークの一端解明を行う。新たに構築するエピスタシスを検出するモデルにより *FZP* とエピスタティックな関係にある遺伝子領域を検出する。幼穂の RNA-seq による発現量変動遺伝子の同定から、*FZP* と相互作用する候補遺伝子を探索し、遺伝学的な証明実験を行う。また、イネプロトプラストを用いたレポーターアッセイにより、*FZP* が候補遺伝子の遺伝子発現に与える影響を調べる。

(2) 遺伝子間相互作用、とくにエピスタシスを検出するため、エピスタティックな関係にある SNP ペアを検出するモデルを構築する。

(3) NAM 集団およそ 2,800 系統の形質値データ (出穂期、穂数、一穂籾数、籾サイズ、葉身幅、葉身長、SPAD 値など) と遺伝子型データを用いて、決定木勾配ブースティング法、ElasticNet など複数の Genomic prediction モデルを検証し、精度の高い形質予測モデルを構築する。NAM 集団の 90% を使ったトレーニングと残り 10% による予測をランダムに繰り返して精度を検証する。

## 4. 研究成果

(1) NAM 集団を用いた一穂籾数に関する GWAS から、Chr.7 にピークを検出した (図 1)。この領域には、籾数 (枝梗の分枝) に関与する *FZP* が座していた (Bai et al. 2017)。下記(2)で確立したエピスタシス検出モデルを用いて、*FZP* 領域とエピスタティックな関係にあるゲノム領域を探索したところ、Chr.11 の 1.6Mbp 付近の SNP (以下、この領域を *qGN11* とする) を検出した。NAM 集団のうち、「ひとめぼれ」と *FZP* が籾数増加アリルである品種「Tupa121-3」と

の RILs のデータから、*FZP* と *qGN11* の両方が Tupa 型 のとき、籾数が大きく増加することが明らかとなった。「ひとめぼれ」と「Tupa121-3」の幼穂 1~2mm の RNA-seq 解析から、*qGN11* において 2 品種間で有意に遺伝子発現量が異なる遺伝子を見出した。この候補遺伝子は、*FZP* 遺伝子発現が抑制されている「Tupa121-3」において遺伝子発現量が多かった。

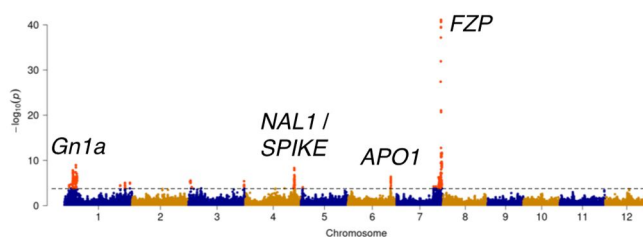


図1 NAM 集団を用いた一穂籾数に関する GWAS.

次に、*FZP* 遺伝子の発現を RNAi で抑制するエフェクターと、「ひとめぼれ」および「TUPA121-3」それぞれの候補遺伝子の上流 2kb を連結したレポーターを構築し、転写因子である *FZP* が候補遺伝子の発現を調節しているか確認するため、イネプロトプラストによるデュアルルシフェラーゼレポーターアッセイを行った。コントロールと比較し、RNAi により *FZP* の転写量を抑制した場合には、レポーターであるルシフェラーゼの発光量が増加した。このことから、転写因子 *FZP* は、候補遺伝子の発現を負に制御していることが明らかとなった。また、「TUPA121-3」のプロモーター領域を連結したレポーターの方が、発光量が増加した。このことから、プロモーター領域内における DNA 多型がエピスタシスの要因であることが示唆された。

(2) RILs を用いてエピスタシスを検出するモデルの検証例として、「ひとめぼれ / Kaluheenati」由来 235 RILs を材料に、籾色において有意な遺伝的相互作用を示す SNP ペア (chr04: 23121877 と chr04: 33353823) を検出した (図2)。RAP-DB において、検出された 2 つの領域に位置する遺伝子を調査し、籾色に影響を及ぼす可能性のある遺伝子を探索した結果、chr04:22212075-26590045 の領域に Black Hull 4 (BH4: chr04: 22969845-22971859)、chr04: 30718147-34310190 の領域では Phenol reaction 1 (Phr1: chr04: 31749141-31751604) を見出した。RIL の親品種「ひとめぼれ」と「Kaluheenati」の BH4 と Phr1 の塩基配列を比較したところ、「ひとめぼれ」は BH4 に 22bp の欠失、Phr1 に 18bp の欠失を有していた。これらの欠失は、他のジャポニカ型品種 (Fukuda et al. 2012) で報告されているものと同じであり、それぞれの遺伝子において機能喪失を引き起こすことが報告されていた (Yu et al. 2008; Zhu et al. 2011)。したがって、BH4 と Phr1 の機能は「Kaluheenati」では維持されているが、「ひとめぼれ」では失われていると考えられた。すなわち、両遺伝子がともに Kaluheenati 型の場合にのみ、籾色が黒色になると考えられた。以上をまとめ、RILs を用いてエピスタシスを検出するプログラム「RIL-StEp」として公開した (Sakai et al. 2021)。

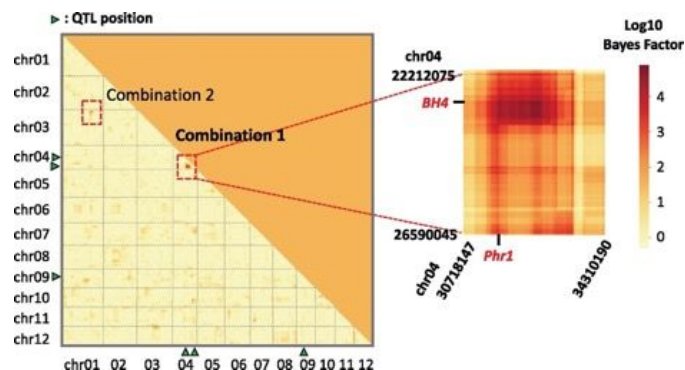


図2 籾色に関する SNP ペアのベイズ因子を示すヒートマップ。左のヒートマップは全ゲノムにおける SNP ペアのベイズ因子を示している。Combination 1 および 2 では、2 つの領域の間に位置する SNP ペアのベイズ因子は 100 以上であった。縦軸および横軸の三角形は主要な QTL の位置を示している。右のヒートマップはゲノム領域を拡大したものである。(Sakai et al. 2021)

(3) ゲノミック予測モデルの構築では、遺伝子型データとして、全ゲノム領域の SNP をハプロタイプブロックにまとめ、親系統ごとに区別した遺伝子型情報を使用することで精度が高くなった。また、予測に用いるハプロタイプブロックの数を段階的 (10~1000) に変えて、その予測精度を比較すると、200 ブロックで予測精度は  $r^2 > 0.9$  となった。本研究の材料では、選択した SNP を用いた 200 ハプロタイプブロックで十分な予測精度になることが判った。統計モデルには、複数の手法を用いて予測結果を比較し、解釈性と精度を考慮して、線形モデル (Elastic Net) が最適と判断した。Elastic Net では、各ハプロタイプブロックの重要度を数値として知ることができるため、解釈性や実際の育種選抜において優れていると考えられる。葉身幅・一穂籾数・穂数・籾サイズ (面積) の形質に関して、テストデータにおける計測形質値と予測形質値間

の相関係数が0.9以上(5-fold CV)を示す予測モデルの構築に成功した(図3)。一方で、出穂期や葉身長では、相関係数が0.7~0.8程度で予測精度はやや低かった。これら形質は環境の影響も強く受けていると推察される。

また、モデルの汎用性を確認するため、モデルの構築・精度の確認に用いたNAM集団とは独立に育成された系統に対し、構築した

モデルを適用した。なお、これら系統は、複数のNAM集団 Founder 親のゲノム断片が混ざっている系統である。これら系統をリシーケンスして得た遺伝子型情報から形質値を予測した結果、一穂粒数・葉身幅サイズに関して相関係数0.9以上の高い精度が得られた。エピスタシスや遺伝子環境間相互作用(G×E)(Sakai et al. 2023)を加味したゲノミック予測モデルの構築、ゲノミック予測モデルを応用したゲノムデザイン(ゲノミックデザイン育種の実証) 予測モデルからのQTL同定などが今後の課題である。

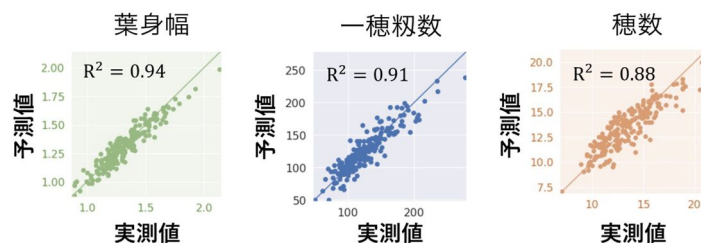


図3 NAM集団を用いた Genomic prediction.

#### <引用文献>

- Godfray HCJ. et al. Food Security: The Challenge of Feeding 9 Billion People. *Science* 327, 812-818 (2010).
- Shomura A. et al. Deletion in a gene associated with grain size increased yields during rice domestication. *Nat Genet* 40, 1023–1028 (2008).
- Mao H. et al. Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *PNAS* 107, 19579-19584 (2010).
- Ashikari M. et al. Cytokinin Oxidase Regulates Rice Grain Production. *Science* 309 (5735), 741-745 (2005).
- Miura, K. et al. OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nat Genet* 42, 545–549 (2010).
- Bai, X. et al. Duplication of an upstream silencer of FZP increases grain yield in rice. *Nature Plants* 3, 885–893 (2017).
- Yu J. et al. Genetic Design and Statistical Power of Nested Association Mapping in Maize, *Genetics* 178, 1, 539–551 (2008).
- Tian, F. et al. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat Genet* 43, 159–162 (2011).
- Eraslan G. et al. Deep learning: new computational modelling techniques for genomics. *Nat Rev Genet* 20, 389–403 (2019).
- Zou J. et al. A primer on deep learning in genomics. *Nat Genet* 51, 12–18 (2019).
- Fukuda A. et al. Complementary genes that cause black ripening hulls in F1 plants of crosses between Indica and Japonica rice cultivars. *Plant Prod. Sci.* 15, 270–273 (2012).
- Yu Y. et al. Independent losses of function in a polyphenol oxidase in rice: differentiation in grain discoloration between subspecies and the role of positive selection under domestication. *Plant Cell.* 20, 2946–2959 (2008).
- Zhu BF. et al. Genetic control of a transition from black to straw-white seed hull in rice domestication. *Plant Physiol.* 155, 1301–1311 (2011).
- Sakai T. et al. RIL-StEp: epistasis analysis of rice recombinant inbred lines reveals candidate interacting genes that control seed hull color and leaf chlorophyll content. *G3 Genes|Genomes|Genetics* 11, jkab130 (2021).
- Sakai T. et al. Whole-genome analysis of recombinant inbred rice lines reveals a quantitative trait locus on chromosome 3 with genotype-by-environment interaction effects. *G3 Genes|Genomes|Genetics* 13, 6, jkad082 (2023).

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件/うち国際共著 1件/うちオープンアクセス 3件）

1. 著者名 Sugihara Yu, Young Lester, Yaegashi Hiroki, Natsume Satoshi, Shea Daniel J., Takagi Hiroki, Booker Helen, Innan Hideki, Terauchi Ryohei, Abe Akira	4. 巻 10
2. 論文標題 High-performance pipeline for MutMap and QTL-seq	5. 発行年 2022年
3. 雑誌名 PeerJ	6. 最初と最後の頁 e13170
掲載論文のDOI（デジタルオブジェクト識別子） 10.7717/peerj.13170	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

1. 著者名 Sakai Toshiyuki, Abe Akira, Shimizu Motoki, Terauchi Ryohei	4. 巻 11
2. 論文標題 RIL-StEp: epistasis analysis of rice recombinant inbred lines reveals candidate interacting genes that control seed hull color and leaf chlorophyll content	5. 発行年 2021年
3. 雑誌名 G3 Genes Genomes Genetics	6. 最初と最後の頁 jkab130
掲載論文のDOI（デジタルオブジェクト識別子） 10.1093/g3journal/jkab130	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Sakai Toshiyuki, Fujioka Tomoaki, Uemura Toyokazu, Saito Shinichi, Terauchi Ryohei, Abe Akira	4. 巻 13
2. 論文標題 Whole-genome analysis of recombinant inbred rice lines reveals a quantitative trait locus on chromosome 3 with genotype-by-environment interaction effects	5. 発行年 2023年
3. 雑誌名 G3 Genes Genomes Genetics	6. 最初と最後の頁 jkad082
掲載論文のDOI（デジタルオブジェクト識別子） 10.1093/g3journal/jkad082	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計1件（うち招待講演 0件/うち国際学会 0件）

1. 発表者名 堺 俊之, 寺内 良平, 阿部 陽
2. 発表標題 イネNAM集団を用いたGenomic Predictionモデルの構築および育種への応用
3. 学会等名 日本育種学会第142回講演会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担者	石川 和也  (Ishikawa Kazuya)  (40804703)	公益財団法人岩手生物工学研究センター・ゲノム育種研究部・研究員   (81202)	

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 協力者	堺 俊之  (Sakai Toshiyuki)  (50911682)	京都大学・農学研究科・助教   (14301)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------