

令和元年6月21日現在

機関番号：12601

研究種目：基盤研究(B) (一般)

研究期間：2016～2018

課題番号：16H04858

研究課題名(和文) 遺伝資源から有用遺伝変異を高効率に発見するための新しい方法論の開発

研究課題名(英文) Development of a new method for discovering useful genetic variations from genetic germplasm efficiently

研究代表者

岩田 洋佳 (Iwata, Hiroyoshi)

東京大学・大学院農学生命科学研究科(農学部)・准教授

研究者番号：00355489

交付決定額(研究期間全体)：(直接経費) 13,000,000円

研究成果の概要(和文)：世界には1750以上のジーンバンクがあり、740万アクセスが現地外保存されている。しかし、育種に利用されているものはこのうち1%以下であると言われている。本研究では、遺伝資源について収集が進むゲノムワイドマーカー情報をもとに、有用な遺伝変異を効率的に発見するための手法を開発した。ゲノミック予測とベイズ最適化を組み合わせることで、多数の系統の中から優良な系統を効率的に発見できることが分かった。また、リモートセンシングを用いた高効率表現型計測や、環境データとの関連解析を組み合わせることで、遺伝資源内の有用変異の発見を更に効率化できることが示唆された。

研究成果の学術的意義や社会的意義

食料問題が深刻化する今後、遺伝資源のもつ有用変異を活用した育種が非常に重要となる。本研究では、ゲノム情報をもとに、有用な遺伝変異を効率的に探索するための手法を開発した。また、ゲノム情報と、リモートセンシングから得られる表現型情報、あるいは、遺伝資源採取地の環境情報を合わせて用いることで、有用遺伝変異の探索を効率化できることを明らかにした。本研究の成果は、遺伝資源のもつ有用変異の発見・利用に貢献すると期待される。

研究成果の概要(英文)：There are more than 1750 genebanks in the world, and 7.4 million plant accessions are stored in the genebanks. However, less than 1% of them have been used in plant breeding. In this study, we developed methods to discover useful genetic variations based on genome-wide marker information collected for genetic resources. By combining genomic prediction and Bayesian optimization, we can efficiently discover excellent genotypes from a large number of candidates. It was also suggested that the discovery of useful genetic variations in genetic resources can be made more efficient by combining high-throughput phenotyping using remote sensing and association studies with environmental data.

研究分野：生物測定学

キーワード：遺伝資源 ゲノミック予測 ゲノムワイドアソシエーション研究 リモートセンシング 環境適応

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

世界中のジーンバンクでは、多くの遺伝資源が活用されことなく眠っている。世界には1750以上のジーンバンクがあり、740万アクセッションが現地外保存されている(FAO 2010)。しかし、育種に利用されているものはこのうち1%以下であると言われている(Sharma ら 2013、Front Plant Sci 4:309)。遺伝資源の利用頻度の低さは様々な理由によると考えられるが、収集されているアクセッションの特性データの不足もその一つである。農業形質やストレス耐性などの特性評価には栽培試験を要するため、多数のアクセッションを評価することが難しい。こうした問題を解決するために様々な作物でコアコレクションが作成されている。これはコアコレクションを代表的な遺伝資源として評価することで遺伝資源の利用を促すものであるが、コアコレクションとして選定されなかったアクセッションについては、かえって利用頻度が低下する可能性が高まる。

近年の分子遺伝学的技術の進展により、現在では、多数のアクセッションについてゲノムワイドマーカーや全ゲノム配列のデータが得られるようになってきている。また、ドローンなどを用いたフェノタイピング技術の向上により、多数のアクセッションについて様々な特徴についての評価を行うことも可能となっている。今後、こうしてえられるゲノム情報や表現型情報を、有用な遺伝資源・遺伝変異の効率的発見・利用に結びつけることが重要な課題である。

2. 研究の目的

次世代シーケンサーの登場により、ゲノム情報を多数の系統について、素早く、かつ、安価に収集できるようになった。これにより、様々な遺伝資源についてゲノムワイドなDNA多型のデータの収集と解析が進められている。例えばイネでは、遺伝資源3,000アクセッションの全ゲノム配列データが収集された(The 3000 rice genomes project 2014、GigaScience 3:7)。また、ゲノムサイズの大きなオオムギでも、遺伝資源22,626アクセッションについて約17万の一塩基多型(SNPs)データが収集・解析された(Milner ら 2019、Nat Genet 51:319-226)。また、イネ科作物以外では、例えば、ダイズにおいて、遺伝資源12,652アクセッションについて42,509 SNPsのデータが収集されている(Xavier ら 2018)。このように、ゲノム解析技術の高速化と低価格化より、将来的には、遺伝資源に含まれる多数のアクセッションについても、ゲノム情報が得られるようになると期待される。

本研究では、こうして収集されているゲノム情報と、ドローンなどを用いて収集される表現型情報、および、遺伝資源の採取地の気象情報を結びつけることにより、遺伝資源に含まれる有用遺伝変異を効率的に探索・検出するための手法の開発を行った。

3. 研究の方法

(1) ゲノムワイドマーカーに基づく有用遺伝変異探索手法の開発

ゲノムワイドマーカーをもとに形質を予測するゲノミック選抜のための予測(ゲノミック予測)モデルは、育種集団における優良個体の選抜だけでなく、遺伝資源に含まれる有望なアクセッションの発見にも利用できる(Pace ら 2015、Plant J 83:903-912; Yu ら 2016 Nat Plant 2:16150)。具体的には以下の手順で探索する。まず、遺伝資源に含まれる多数のアクセッションについて、あらかじめゲノムワイドマーカーのジェノタイピングを行っておく。これは上述したように、現在、様々な作物でこうしたデータが提供されつつある。しかし、多数のアクセッションの栽培試験を行うのは容易ではないので、一部のアクセッション、例えば、コアコレクションについて栽培試験を行い、ゲノムワイドマーカーから栽培試験の結果を予測するモデルを作成する。得られたモデルをもとに、遺伝資源に含まれる多数のアクセッションの栽培試験の結果を予測し、予測結果に基づいて有望なアクセッションを選抜する。こうして有望なアクセッションを探索することで、遺伝資源に含まれる有用な遺伝変異を高効率で発見して、育種に利用できる。本研究では、こうした遺伝資源探索手法について、さらにベイズ最適化とよばれる基準を応用して、有用遺伝変異の発見効率を更に向上させる手法の開発を行った。また、公開されているゲノムデータと表現型データを用いて、開発された手法を従来手法と比較して、その有効性について評価した。

(2) リモートセンシング技術を用いたハイスループットフェノタイピング技術の開発

上述したように、ゲノミック予測をもとに有用遺伝変異を効率的に探索する手法は、遺伝資源の効率的に利用するために非常に有効であると考えられる。しかし、同手法を利用するためには、一部遺伝資源アクセッションについて表現型を計測して、その計測値とゲノムワイドマーカー間の関連からゲノミック予測モデルを構築する必要がある。したがって、ゲノムを利用した有用遺伝変異探索においても、表現型計測の高効率化は重要な課題である。そこで、本研究では、タルホコムギの遺伝資源400系統の栽培試験を行い、リモートセンシングによる形質評価を行った。形質評価は、ドローンを用いたリモートセンシングと地上における近接リモート

センシングの両方を行い、それらの計測値間の比較を行った。さらに、これら遺伝資源について、DArT-Seqを用いたジェノタイピングを行った。リモートセンシングによって計測された表現型とDArT-Seqで得られたゲノムワイドDNA多型の関連をゲノミック予測モデルでモデル化し、交差検証によってその精度の検証を行った。

(3) 環境適応関連遺伝子を検出するための遺伝資源のゲノムワイドマーカーデータの解析

様々な地域から収集される遺伝資源は、各地域の環境に適応するための遺伝変異を保有していると考えられる。このような観点のもと、地理情報システム(GIS)を用いて遺伝資源から有望なアクセッションを探し出す方法として Focused Identification of Germplasm Strategy (FIGS)がある。FIGSは、地理情報システム(GIS)から得られる遺伝資源採集地の環境情報をもとに目的形質を備える有望アクセッションを選抜し、有用遺伝変異の発見確率を向上させる手法で、黒さび病抵抗性のコムギや耐乾性のソラマメの発見等で実績をあげた(Bari 2012, Genet Resour Crop Ev 59:1465-1481)。FIGSは有用遺伝変異の発見において高いポテンシャルを秘める手法だが、環境情報のみを基準とするために十分な選抜精度が得られない可能性がある。GISから得られる環境情報をゲノム情報と組み合わせ、環境適応に関連する遺伝変異を探索する方法として、Environmental Association Analysis(EAA)がある(Rellstabら 2015, Mol Ecol 24:4348-4370)。EAAはもともと、landscape genomicsとよばれる生態地理学分野で発展してきた手法である。しかし、環境適応に関連する遺伝子座の検出も可能であることから、作物の遺伝資源の解析にも用いられている。本研究では、EAAの手法、および、ゲノミック予測を用いて、環境データとゲノムワイドマーカー遺伝子型データ間の関連をもとにした環境適応関連遺伝子の探索のポテンシャル評価を行った。

4. 研究成果

(1) ゲノムワイドマーカーに基づく有用遺伝変異探索手法の開発

ベイズ最適化とよばれる基準を応用して、上述した手法と異なる指標に基づく選抜を提案した。同指標では、予測値を直接用いるものではなく、予測の不確実性も考慮に入れ、予測値がある一定の基準(例えば、その時点での最良の品種の能力)を超える確率を計算して、その確率の大きなものを選抜するという指標である。この指標を用いることで、遺伝資源がもつ大きな多様性の「空間」を、効率的に探索できるようになる。予測モデルは常に真のモデルではないため、予測誤差をもつ。この状況で単に予測値の高いもの、あるいは、低いものだけを選んでいると、広い遺伝資源の「空間」の中に取りこぼされるものが生じる。不確かな遺伝資源についても、ときに探索を行うことで、こうした取りこぼしが少なくなると期待される。同指標、および、ゲノミック選抜を、公開されているゲノムデータと表現型データの4つのデータセット(イネ、トウモロコシ、コムギ×2)を用いて、有用な遺伝変異をシミュレーションを行い、ベイズ最適化に基づく探索の通常探索(予測値の大きなものを選抜する)に対する優位性を検討した。その結果、データセットおよび形質によってベイズ最適化の優位性の程度は異なったが、ほとんどの場合においてベイズ最適化を用いた探索で、より早く優良なアクセッションを発見することができた。以上の結果から、ゲノミック予測を用いて、ベイズ最適化とよばれる指標をもとにして選抜することで、より効率的に優良アクセッションを見つけ出すことができた。

(2) リモートセンシング技術を用いたハイスループットフェノタイピング技術

地上における近接リモートセンシングとドローンリモートセンシングを行い、4ブロック中1ブロック(100プロット)について植被面積を計測した結果、両手法間の計測値の相関は高く、 $r = 0.9$ 程度であった。近接リモートセンシングでは、高解像度の近赤外カメラを用いて撮影を行っており確実に植物体を抜き出せるだけでなく、高精度に面積が計測されていると考えられるため、真値として扱うことができる。このことから、ドローンを用いたリモートセンシングの精度も高いと考えられた。また、DArT-Seqを用いて346系統についてゲノムワイドSNP遺伝子型を決定した結果、49,918 SNPの多型が明らかとなった。このうち、欠測が20%未満、マイナー対立遺伝子頻度が5%以上の14,100 SNPsを抽出し、これらSNPsの遺伝子型をもとにマーカー血縁行列を計算して以下の解析に用いた。ドローンリモートセンシングの結果から、植被面積、4種の方法で計測された草丈、1プロット内の個体の体積を計算し、これらのそれぞれについてBayesian GBLUPをもとにゲノミック予測モデルを構築して、その精度を評価した。精度評価には4反復の10分割交差検証を用い、評価指標には観察値と予測値の相関係数を用いた。その結果、4種の方法で計測された草丈ではいずれも予測精度が比較的高かった($r > 0.6$ 、図1)。一方、植被面積では、予測精度が低かった(約 $r = 0.3$)。また、体積では両者の中間的な値を示した(約 $r = 0.5$)。なお、ソルガムのドローンリモセンで計測された草丈のゲノミック予測精度(Watanabe et al. 2017, Front. Plant Sci. 8:421)に比べても高い値が得られ

ており、タルホコムギの遺伝資源におけるドローンリモセンシングとゲノミック予測の有効性が示された。なお、植被面積ではゲノミック予測の精度が低かったが、上述したようにドローンリモートセンシングの計測値と近接リモートセンシングの計測値との相関は非常に高かったことから、ドローンリモートセンシングの計測精度が低いのではなく、植被面積の遺伝率が低いために、観察値と予測値の相関が低下したことが原因と考えられた。これは、各系統が、標準系統を除いて、反復なしの1個体だけで栽培・計測されたことによると考えられる。匍匐性のタルホコムギでは、1個体あたりの栽培面積が大きく、多数の個体・系統を同時栽培することが難しい。ゲノミック予測とリモートセンシングを用いた表現型計測を組み合わせることで、栽培試験を実施していない遺伝資源アクセッションについても予測値を計算し、選抜できる。タルホコムギのように多系統の同時栽培が難しい作物では、特にこうしたアプローチが有効であると考えられる。今後、上述したベイズ最適化も組み合わせることで有用遺伝資源・遺伝変異の探索を行うことで、タルホコムギの遺伝資源の有効利用が進むであろう。

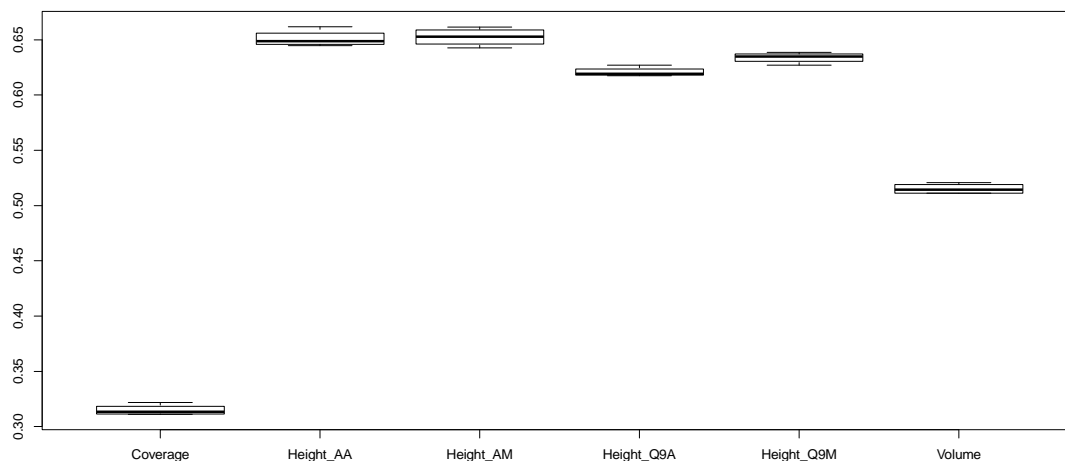


図 1. ドローンリモートセンシングによって計測された形質のゲノミック予測の精度。計測項目は、左から、植被面積、4種の計測法による草丈、体積である。

(3) 環境適応遺伝子を探索するための遺伝資源のゲノムワイドマーカーデータの解析法

ピーカン (*Carya illinoensis* (Wangenh.) K. Koch) およびその近縁種を集めた遺伝資源をもとに、GIS情報をもとに収集された環境(気象)データと、RAD-Seqによって得られたゲノムワイドマーカー間の関連の解析を行った。その結果、環境データとSNPの間に有意なアソシエーションが検出された。特に、年平均よりも月平均とのアソシエーションが高く、季節性の環境変化とSNP間のアソシエーションが高いことが分かった。さらに、環境データとのアソシエーションが高いSNPだけで開芽日の予測を行った結果、数少ないSNPである程度の予測が可能であることが示された。こうしたSNPは、ピーカン遺伝資源の環境適応に関わる遺伝子と強く連鎖している可能性があり、環境適応性の遺伝的改良のために有用なマーカーとなる可能性もある。GISを用いて遺伝資源から有望なアクセッションを探し出すFIGSでは、環境データのみに基づいて有望なアクセッションを選出する。それに対して本手法では、ゲノムワイドマーカーデータと環境データの両方に基づいて有望なアクセッションを選出するため、有望なアクセッションをより高精度に選出できると期待される。さらに、環境適応性に関連する可能性がある遺伝子の位置や効果も推定できる点は、FIGSに対する大きな利点と考えられる。今後、現在整備が進められているピーカンの全ゲノム配列も用いて、今回アソシエーションが検出されたSNPと遺伝子との関連をより詳細に検討していくことで、こうした手法の有効性がより明確になるであろう。ジーンバンクに保存されている遺伝資源には多くの場合採取地の位置情報も提供されており、採取地の環境データと、遺伝資源のゲノムワイドマーカーデータを組み合わせることで、栽培試験を行わなくとも環境適応に関わる遺伝子候補をリストアップできる可能性がある。なお、こうした候補について、実際の遺伝子の効果を確認するためには、栽培試験が必要となると考えられるが、その場合においても、マーカー遺伝子型をもとに、効率的に栽培試験を設計して実施できる。

(4) まとめ

遺伝資源においてゲノム情報の収集が進めば、コアコレクションなどをもとにしてきた従来のアプローチとは異なる、新たな戦略に基づく遺伝資源のスクリーニングが可能となる。多数の遺伝資源について収集が進むゲノム情報を、本研究で開発・検討した手法と組み合わせることで、有用な遺伝資源・遺伝変異の発見・利用が加速されると考えられる。今後は、

人口増加と気候変動により食料問題はますます深刻化すると考えられ、様々な作物種について遺伝的改良を高速に進めていくことが重要となる (Massawa ら 2016, Trends Plant Sci 21: 365-368)。そのためには、本研究で研究を行ったゲノム情報に基づく遺伝資源の戦略的活用は不可欠となると考えられる。

5. 主な発表論文等

〔雑誌論文〕(計 3 件)

1. 岩田洋佳 (2019) 林木のゲノミック選抜: 現状と展望. 森林遺伝育種 8:32-39. (査読付き)
2. Ryohei Tanaka, Hiroyoshi Iwata (2018) Bayesian optimization for genomic selection: a method for discovering the best genotype among large number of candidates. Theoretical and Applied Genetics 131:93-105. (査読付き)
3. Alagu Macnickavelu, Tomohiro Hattori, Shuhei Yamaoka, Kazusa Yoshimura, Youichi Kondou, Akio Onogi, Minami Matsui, Hiroyoshi Iwata, Tomohiro Ban (2017) Genetic nature of elemental contents in wheat grains and its genomic prediction: toward the effective use of wheat landraces from Afganistan. PLoS ONE 12:e01694166. (査読付き)

〔学会発表〕(計 15 件)

1. 岩田洋佳 (2018) ゲノムワイドマーカーを用いたモデル化に基づく遺伝資源の戦略的活用. 日本草地学会 (招待講演)
2. 岩田洋佳 (2018) ゲノム情報をもとに植物の環境適応・応答をモデル化する. 森林遺伝育種学会 (招待講演)
3. Iwata H (2018) Modelling response of plants to environment via high-throughput phenotyping and machine learning. International Biometrics Symposium (招待講演)
4. Iwata H (2018) Selection 4.0: Next revolution of breeding will be model-based development (MBD). Plant and Animal Genome Asia (基調講演)
5. Iwata H (2018) Genomic selection in rice: potential and future directions. IRRI Seminar (招待講演)
6. Iwata H (2018) Selection 4.0: A system for model-based development (MBD) of crop. 2018 Plant GWAS Workshop (招待講演)
7. 岩田洋佳 (2017) 植物ビッグデータと育種. 植物科学シンポジウム 2017 (招待講演)
8. 岩田洋佳 (2017) ドローンリモセンを用いた作物表現型計測とそのゲノム育種への利用. 植物科学・作物育種におけるフェノーム解析 (招待講演)
9. Ishimori M, Takanashi H, Fukami K, Cervantes K, Nagano AJ, Kajiya-Kanegae H, Grauke LJ, Tsutsumi N, Rendall J, Iwata H (2017) Genome-wide RAD-Seq analysis revealed subpopulation structure of pecan (*Carya illinoensis*) germplasm collection and their relationship to geographical distribution patterns. VIII International Symposium On Walnut, Cashew and Pecan (国際学会)
10. 渡辺翔、郭威、岩田洋佳 (2017) ドローンによるフィールドセンシングで植物の成長を追う. 日本育種学会第 132 回講演会.
11. Iwata H (2016) Genomic assisted breeding and data science for expanding its possibility. Special Lecture in National Taiwan University (招待講演)
12. Iwata H (2016) Phenomics and morphometrics for genomics assisted breeding. Multi-scale Plant Phenomics Workshop.
13. Iwata H (2016) Break the phenotyping bottleneck in genomics-assisted breeding. ICRA 2016: The 9th International Conference on Intelligent Robotics and Applications (招待講演)
14. 岩田洋佳 (2016) データ科学と育種: データ科学を応用し、品種改良を加速する. 東京大学農学部公開セミナー (招待講演)
15. Chen TS, Yamasaki Y, Mahjoob M, Guo W, Tsujimoto H, Iwata H (2017) The application of image processing in quantifying the morphological traits of *Aegilops tauschii*. The 4th International Symposium of Biological Shape Analysis (国際学会)

〔図書〕(計 1 件)

1. Jennifer Spindel, Hiroyoshi Iwata (2018) Genomic Selection in Rice Breeding. In "Rice Genomics, Genetics and Breeding" (Takuji Sasaki Motoyuki Ashikari Eds.), Springer, pp. 473-496.

〔産業財産権〕

出願状況（計 0 件）

名称：
発明者：
権利者：
種類：
番号：
出願年：
国内外の別：

取得状況（計 0 件）

名称：
発明者：
権利者：
種類：
番号：
取得年：
国内外の別：

〔その他〕
ホームページ等

6 . 研究組織

(1)研究分担者

研究分担者氏名：辻本壽
ローマ字氏名：Hisashi Tsujimoto
所属研究機関名：鳥取大学
部局名：乾燥地研究センター
職名：教授
研究者番号（8桁）：50183075

(2)研究協力者

研究協力者氏名：Jennifer Rendall
ローマ字氏名：Jennifer Rendall

研究協力者氏名：LJ Grauke
ローマ字氏名：LJ Grauke

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。