

# ゲノムと画像情報を活用した効率的な果樹育種基盤の開発

南川 舞 (千葉大学 国際高等研究基幹)

minamikawa@chiba-u.jp

果物は健康で豊かな食生活を送るうえでの必需品である。消費者や生産者のニーズに対応するため、高品質な果実形質を有する新品種の迅速な育成が求められている。本研究では、カンキツ、リンゴ、ナシを対象として、ゲノムと画像情報を活用した育種法が極めて有用であることを実データで示した。芽生えの段階で将来できる果実の特性を高精度に予測して選抜できるようになり、消費者や生産者のニーズに対応した高品質な新品種の迅速かつ効率的な育成が期待される。

## はじめに

果樹は一般的に芽生えから結実までの時間が長く、交雑育種に長い年月を要する。また、個体サイズが大きいため多数の個体を選抜対象にできず、選抜基準を超える個体の獲得率が非常に低い。リンゴ・ナシなどの多くのバラ科果樹では自家不和合性を有するため、交雑可能な組合せが限定されることも育種上の障壁となっている。これらの課題を克服して効率的な果樹育種を推進するために、筆者は多くの研究者と協力し、自家不和合性機構の解明とゲノムと画像情報を活用した育種法の実装を目指して研究に取り組んできた。本稿ではこれらの研究について紹介する。

## 1. リンゴ花粉側自家不和合性因子の単離と生化学的機能解析

自家不和合性 (Self-incompatibility) とは、雌ずい・花粉共に正常な受精能力を持つにもかかわらず、自家受粉では受精しない性質である。リンゴ・ナシなどのバラ科果樹は自家不和合性を有するため、実際の栽培現場では人工授粉やハチなどの花粉媒介昆虫を利用した他家受粉によって果実を生産している。自家不和合性を人為的に制御できれば、人工授粉の省力化や花粉媒介昆虫に依存しない栽培が可能となる。

自家不和合性における自己・非自己の認識は、*S* 遺伝子座と呼ばれる単一の遺伝子座に座乗する雌ずい *S* 遺伝子と花粉 *S* 遺伝子のセット (*S* ハプロタイプ) によって支配されている。バラ科サクラ連やナス科の植物では、雌ずい *S* 遺伝子は RNase をコードする *S-RNase* であり、花粉 *S* 遺伝子は F-box タンパク質をコードする *SFB* (*S haplotype-specific F-box*) / *SLF* (*S locus F-box*) であることが知られていた。F-box タンパク質は一般に SCF (Skp-Cullin-F-box) 複合体を形成して標的タンパク質を分解することが広く知られているため、「花粉管伸長に対して細胞毒として働く *S-RNase* の認識・分解に関与しているのではないか？」といった仮説が立てられた。そこで筆者らは、バラ科ナシ連のリンゴのゲノム網羅的に花粉 *S* 遺伝子を探索して単離し、これがどのように自己・非自己を認識しているのかを明らかにすることを目的として研究を進めてきた。その結果、リンゴでは予想に反して 1 つの *S* ハプロタイプあたり 10 個以上の *F-box* 遺伝子 *MdFBXs* (*Malus × domestica F-box*) / *MdSFBBs* (*Malus × domestica S-locus F-box brothers*) が花粉 *S* 遺伝子としての性質を有することを明らかにした<sup>1,2)</sup>。当時、雌ずいおよび花粉 *S* 遺伝子は 1 つの *S* ハプロタイプあたり 1:1 で存在すると考えられてきたが、多種類の *MdFBXs/MdSFBBs* 遺伝子が花粉 *S* 因子として関与する可能性を提唱した。その後のリンゴにおけるタンパク質間相互作用解析で

は、MdFBX/MdSFBB と MdSkl1、MdSkl1 と MdCullin1 がそれぞれ相互作用することを示し、MdFBX/MdSFBB は SCF 複合体を構成して機能している可能性が示された<sup>3,4)</sup>。

## 2. カンキツ・ナシ・リンゴにおけるゲノミック選抜とゲノムワイド関連解析

ゲノム情報を活用した育種法であるゲノミック選抜 (Genomic Selection: GS) やゲノムワイド関連解析 (Genome-Wide Association Study: GWAS) は、上述した果樹育種の障壁を取り除ける可能性がある。GS は、ゲノムワイドな DNA 多型の情報から個体の遺伝的能力を統計的に予測 (ゲノミック予測; Genomic Prediction: GP) して早期選抜する方法であり、従来のマーカー利用選抜 (Marker-Assisted Selection: MAS) では適用が難しかった多数の遺伝子が関与する果実形質などの改良も可能である。GWAS はゲノムワイドな DNA 多型と形質との関連の有無を統計的に調べる方法である。形質との関連が明らかとなった多型周辺のゲノム領域から形質を制御している候補遺伝子を同定することができる。筆者らは、農研機構果樹茶業研究部門の育種プログラムで使用されているカンキツ・ナシ・リンゴの品種群や交雑育種集団から収集されたゲノムワイドな一塩基多型 (Single Nucleotide Polymorphism: SNP) と多数の果実形質の表現型情報を用い、ゲノム情報を活用した育種の可能性を評価した。

品種群のみ、または育種集団のみよりも、品種群と育種集団を統合した集団 (予測対象のきょうだい集団は除く) をトレーニングデータとして GP モデルを構築した場合に、育種集団の GP 精度が最も高かった<sup>5,6)</sup>。カンキツの果皮色、ナシの収穫期などは特に高い精度を示した (実測値と予測値の相関係数 ( $r$ ) が 0.7 より大きい)。また、トレーニングデータに予測対象のきょうだい集団と同じきょうだいの情報も含めることで予測精度はさらに向上した<sup>6)</sup>。通常、初めての交配組み合わせで生まれた予測対象のきょうだい集団の情報を利用することはできない。GS は芽生えの段階で行うため、予測対象のきょうだい集団の果実形質の情報を集めることができないからである、しかし、有望な個体が得られた場合には、より優秀な個体の獲得を目指して同じきょうだいの個体を追加で調査することがある、この場合、予測対象のきょうだい集団と同じきょうだいの情報をトレーニングデータとして利用することが可能となる、予測対象のきょうだい集団の情報が利用可能な場合は積極的に GP モデルに利用するべきである。また、品種群と育種集団

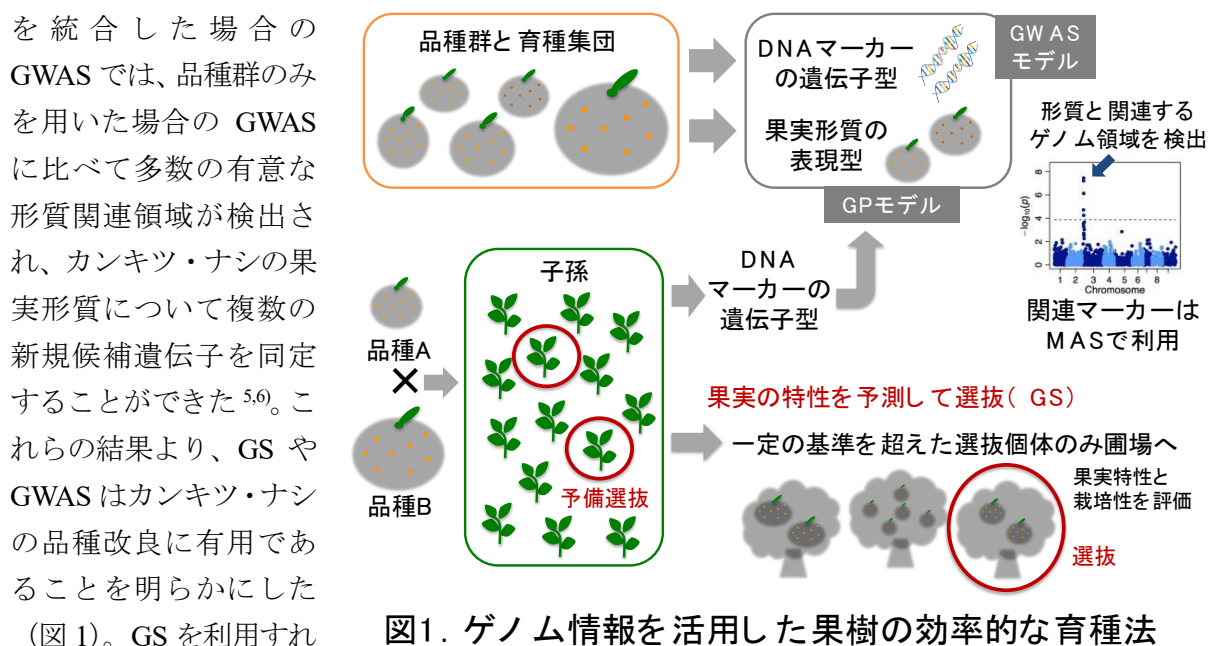


図1. ゲノム情報を活用した果樹の効率的な育種法

ば、多くの果実形質について有望な個体だけを選抜して圃場に植え付けられるため、実質的な評価個体数を大幅に増加できるものと期待される。データ収集が難しい果樹では、実際の育種集団の情報をGSやGWASに活用することが有効である。今後、交雑育種を行いながら多数の育種集団の表現型情報に加えてマーカー遺伝子型の情報も蓄積していくことで、GSやGWASの精度をさらに向上できる可能性が示唆された。

リンゴでは、7つの起源品種に由来する、14種類の起源ハプロタイプの遺伝を自動的に追跡する方法を開発し<sup>7)</sup>、その14種類の起源ハプロタイプ情報を用いたGPやGWASの精度を評価した。リンゴ酸含量ではSNPよりも起源ハプロタイプ情報を用いた場合にGP精度が高かった<sup>7)</sup>。起源ハプロタイプ情報を用いたGWASでは、‘ウースターペアメイン’という起源品種に由来する1つのハプロタイプが、果皮の着色と最も強い関連を示した。このハプロタイプの頻度が品種群内で有意に増加していることから、果皮の着色の良いリンゴの育成に利用されてきた可能性が示唆された。起源ハプロタイプ情報を活用することで、リンゴの品種改良の歴史を紐解くことができるだけでなく、精度の高いGSを行える可能性があることを示した。

### 3. 機械学習を用いた育種家の感性の理解

GSやGWASなどのゲノム育種を効果的に行うためには、非常に多くの高精度なデータを準備する必要がある。次世代シーケンス技術の進展により、多くのマーカー遺伝子型を低コストで取得できるようになってきている一方で、果実形質の多くは少数の育種家の感性により達観的に評価されており、短期間で大量かつ高精度なデータを取得することは難しい。例えば、カンキツの剥皮性や果実硬度などは育種家の触感に基づき評価されているが、その感性に関わる果実の形態的な特徴が明らかになっておらず、画像解析などによる定量的かつ自動的な評価を行うことも困難であった。そこで筆者らは、カンキツ果実の横断面の画像解析により、果実形態の特徴を客観的かつ定量的に評価する方法を検討した。そして、剥皮性・果実硬度の指標となる果実形態の特徴を明らかにするために、定性的に評価された剥皮性と、定量的に評価された果実形態の特徴との関係を、機械学習のモデルを用いて表現した。

プログラミング言語Pythonを用いることで、カンキツの果実断面の画像から、果実のさまざまな形態的特徴を定量的かつ自動的に評価する技術を開発した<sup>8)</sup>。複数の機械学習モデルを適用することで、果芯の崩壊程度は剥皮性・果実硬度の両形質と強い関連を示し、果芯の崩壊程度が大きい果実は、剥皮が容易であり、軟らかい果実である傾向が観察された(図2)<sup>8)</sup>。一方で、果実の面積に対する種子面積の割合は果実硬度でのみ強い関連を示し、種子面積の割合が大きい



図2. カンキツ剥皮性・果実硬度と関連がある果実の形態的特徴

い果実は硬い果実である傾向が観察された。ベイジアンネットワークによる解析では、果芯の崩壊程度は剥皮性・果実硬度の両形質に、また、種子面積については果実硬度に対して、直接的な影響を及ぼしている可能性があることが示された。次に、機械学習の手法の一つである深層学習を用いて剥皮の難易または果実の硬軟を判別させ、その判別に寄与する果実形態の特徴を Grad-CAM 法を用いて可視化した。その結果と果実形態の情報を組み合わせることで、果芯やアルベドの崩壊領域の特徴が、剥皮しやすい、かつ、軟かい果実の分類に寄与している一方で、果肉やアルベドの領域の特徴は、剥皮しづらい、かつ、硬い果実の分類に寄与していることを明らかにした。種子の領域の特徴は、硬い果実の分類でのみ寄与していることがわかった。

以上の結果より、画像解析とさまざまな機械学習の手法を組み合わせることで育種家の感性を紐解き、カンキツの剥皮性・果実硬度に関連する果実の形態的な特徴を明らかにすることができた。果芯の崩壊程度や種子面積を改良することで、望ましい剥皮性や果実硬度を有する新品種の開発が期待される。また、カンキツ果実の横断面の画像から、剥皮性・果実硬度に関連する果実の形態的な特徴のデータを自動的かつ大量に収集できるようになり、高精度なゲノム育種を実現できる可能性がある。

## おわりに

筆者らがこれまでカンキツ・ナシ・リンゴを対象に開発してきたゲノム育種の枠組みは、他の果樹育種へも応用が可能である。これまでの研究により GS や GWAS の有効性が育種学の研究レベルでは明らかになったが、これらの技術を育種現場へ実装するためには課題も多い。それぞれの樹種における育種プログラムに合わせ、実際の育種に携わる人が直感的に使いやすい GS や GWAS のシステム作りも必須であろう。

## 謝辞

本賞の受賞にあたり、千葉大学大学院園芸学研究院よりご推薦を賜りました。松岡延浩研究院長ならびにご支援いただきました関係者の皆様に心より御礼申し上げます。本研究の実施にあたり、東京大学の岩田洋佳教授と千葉大学の佐々英徳教授には、終始懇切なご指導を賜り、心より感謝申し上げます。果樹のデータをご提供いただきました農研機構果樹茶業研究部門の山本俊哉博士ならびに多くの共同研究者の皆様にも貴重なご助言を多数いただきました。末筆ながら、統計遺伝学を学ぶ機会を授けてくださった大澤良博士に深く感謝申し上げます。

## 引用文献

- 1) Minamikawa M., Kakui H., Wang S., *et al.*: *Plant Mol. Biol.* 74: 143–154 (2010).
- 2) Sassa H., Kakui H., Miyamoto M., *et al.*: *Genetics* 175: 1869–1881 (2007).
- 3) Minamikawa M. F., Fujii D., Kakui H., *et al.*: *Plant Biotechnol.* 30: 119–123 (2013).
- 4) Minamikawa M. F., Koyano R., Kikuchi S., *et al.*: *PLoS One* 9: e97642 (2014).
- 5) Minamikawa M. F., Nonaka K., Kaminuma E., *et al.*: *Sci. Rep.* 7: 4721 (2017).
- 6) Minamikawa M. F., Takada N., Terakami S., *et al.*: *Sci. Rep.* 8: 11994 (2018).
- 7) Minamikawa M. F., Kunihisa M., Noshita K., *et al.*: *Hortic. Res.* 8: 49 (2021).
- 8) Minamikawa M. F., Nonaka K., Hamada H., *et al.*: *Front. Plant Sci.* 13: 832749 (2022).