



PRESS RELEASE

岡山大学記者クラブ

文部科学記者会

科学記者会

御中

令和 4 年 3 月 4 日

岡山大学

報道解禁：令和 4 年 3 月 8 日（火）午後 9 時（新聞は 9 日朝刊より）

AI が覗き込むトマトゲノム：果実の遺伝子の動きを見抜く

～果実が「熟れる」仕組みの緻密なデザイン～

◆発表のポイント

- ・ トマトの全ゲノム配列情報に AI を適用することで、遺伝子発現（遺伝子の「動き」）を予測する技術を開発しました。
- ・ 果実が「熟れる」際の遺伝子発現にとって鍵となる DNA 配列を予測することができました。
- ・ AI が予測した鍵となる DNA 配列を自由に改変することで、果実が熟れる仕組みの緻密なデザインが可能になると期待されます。

果実の色や甘さ・香りなどの特徴はそれぞれの作物のゲノムに存在する特定の遺伝子の発現（遺伝子の「動き」）によって決定されています。しかし、それら遺伝子の発現パターンはゲノム内の複雑な要素の組み合わせで決まるため、たとえ全ゲノム配列の情報が明らかになっていても予測することは困難なものでした。岡山大学学術研究院環境生命科学学域（農）の赤木剛士研究教授、増田佳苗院生（D3）、栗田恵理子院生（M1）は、このたび国立研究開発法人農業・食品産業技術総合研究機構、筑波大学大学院生命環境系、九州大学大学院システム情報科学研究院の共同研究者とともに、本来は写真などの画像・映像や言語に使用されることが多い AI 技術である深層学習（ディープラーニング）をトマトの全ゲノム情報に適用することで、果実が「熟れる」過程で重要となる遺伝子の発現変化を予測する技術を開発しました。さらに、「説明可能な AI (X-AI)」と呼ばれる技術を活用することで、AI が予測した果実の遺伝子発現にとって重要な鍵となる DNA 配列を特定することができました。今後、さらなる「AI との協働」によって、私たちヒトでは判断が難しい膨大な全ゲノム情報の中にある仕組みを紐解き、その鍵となる DNA 配列を自由に改変することで、果実の様々な特徴についての緻密なデザインが可能になると期待されます。

本研究成果は、日本時間 3 月 8 日午後 9 時（米国東部標準時間：3 月 8 日午前 7 時）、米国の科学雑誌「*The Plant Cell*」オンラインアドバンス版に掲載されます。

果実の「食べごろ」を一緒に探してみませんか？あなたのアイデア次第で、新しい果実をデザインできるかも！



栗田院生（M1）



増田院生（D3）



赤木
研究教授



PRESS RELEASE

■発表内容

<現状>

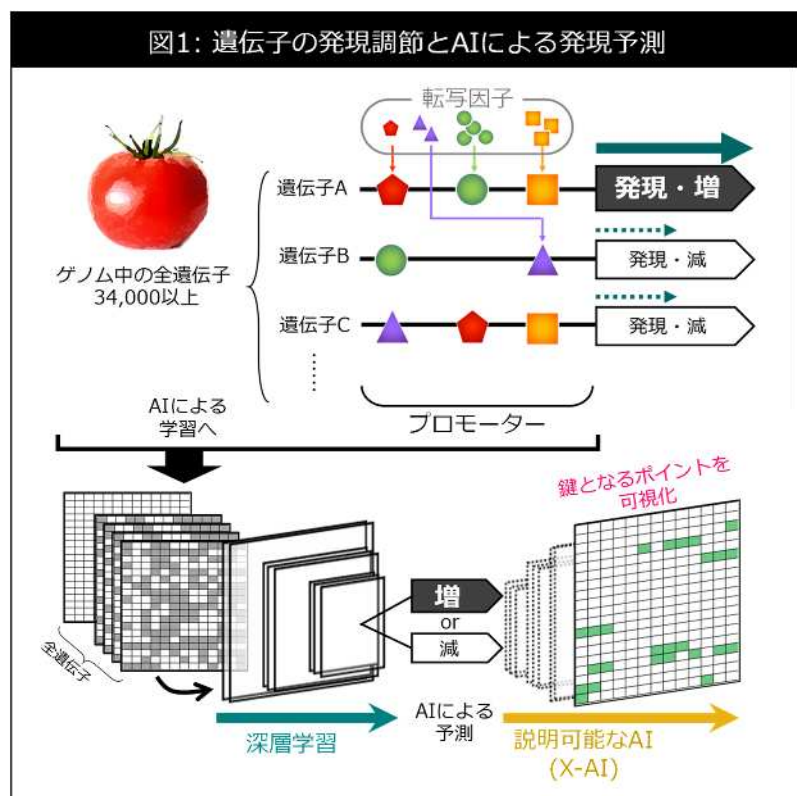
果実の大きさや色・甘さ・香りなどを決定する生育プロセスは、それぞれの個体の遺伝情報である「ゲノム」に存在する数万にもおよぶ遺伝子の発現（すなわち遺伝子の動き）の組み合わせで決定されます。しかし、それぞれの遺伝子の発現は、多くの複雑な要素で決まってくるため、たとえ全ゲノム配列の情報が明らかになっていても、それを予測することは極めて難しいという現状がありました。近年、話題となっているAI（深層学習）の技術は、様々な分野で取り入れられており、特に画像解析・診断の分野で利用される機会が多くなってきています。今回、私たちの研究ではAI技術をトマトの全ゲノム情報にも応用して、トマト果実が熟れていく過程での遺伝子発現を予測し、その予測の鍵となったDNA配列を特定する研究を行いました。

<研究成果の内容>

遺伝子には「プロモーター」と呼ばれる場所が存在し、この部分に「転写因子」と呼ばれる様々な種類のタンパク質が結合することで遺伝子の発現は調節されています（図1）。しかし、1つの転写因子が結合するプロモーターのDNA配列は一様ではなく揺らぎを持つ複数のパターンが存在するうえ、遺伝子の発現パターンは結合する転写因子の複雑な組み合わせによって変化するため、遺伝子発現を予測することは容易ではありません。そこで本研究では、最近のAI技術の代表としても知られる深層学習（ディープラーニング）を2つの段階で活用することでプロモーターDNA配列から遺伝子発現の予測を試みました。まず、モデル植物（注1）であるシロイヌナズナの転写因子が結合するDNA配列情報を集めたデータベースを学習することで、トマトの全遺伝子（34,000以上が存在）のプロモーターにおいて転写因子が結合するポイントを網羅的に予測しました。次の

ステップとして、トマト果実が熟れていく過程での全遺伝子発現パターンを学習することで、遺伝子発現の増加・減少を予測するAIモデルを構築しました。さらに、「説明可能なAI（X-AI）」と呼ばれるツールを活用して、深層学習モデルをさかのぼり、「AIが判断した理由を可視化する」ことで、予測した遺伝子発現の鍵となるDNA配列を1塩基レベルで明らかにする技術を開発しました。この予測遺伝子発現の鍵となるDNA配列を改変した遺伝子を実際のトマト果実に導入すると、AIによる予想と同じ結果が得られ、AI目線での予測が私たちヒトでは到底

図1: 遺伝子の発現調節とAIによる発現予測

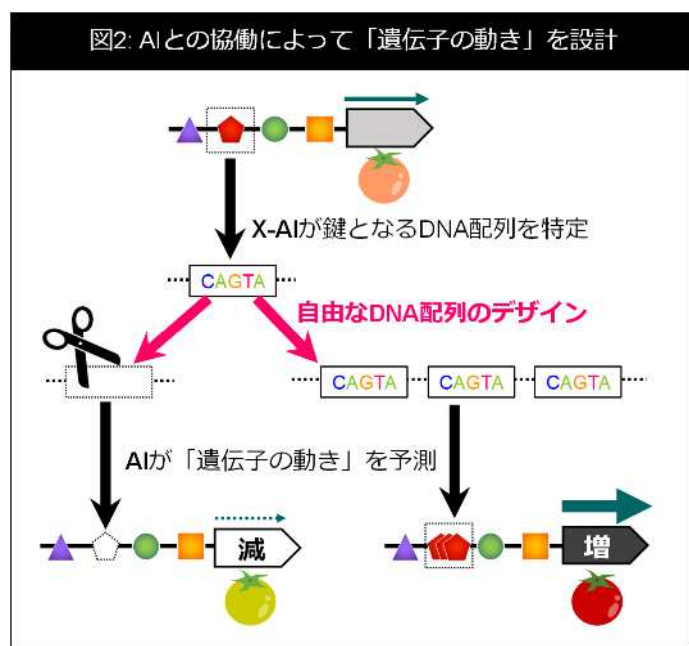


PRESS RELEASE

判断のつかない複雑なトマトゲノム情報の中の仕組みを正確に読み解いていることが示されました。この技術によって、果実が熟れる過程で遺伝子発現の制御を担う重要な DNA 配列の存在を明らかとすることが可能であり、トマトなどの果実が「熟れる」仕組みの緻密なデザインを目指すための重要な知見を生み出すと考えられました。

＜社会的な意義＞

ゲノム情報に AI 技術を適用するという発想は、トマトに限った話ではなく、あらゆる作物・植物について応用することが可能です。今回は一つのモデル研究として果実の「食べごろ」にとって重要な過程に着目した研究を行いました。多様な作物の全ゲノム情報を対象にして、色や形、美味しさ、香りなど、様々な特徴やプロセスに関わる遺伝子の発現予測に使っていくことが出来ると思われれます。また、今回の技術では「説明可能な AI」の概念を適用して、予測した遺伝子の発現にとって重要な鍵となる DNA 配列を見つけ出すこともできるため、近年になって注目を集めている遺伝子編集（注2）を応用することで、最適な遺伝子発現パターンを人工的に創出して、自由な果実のデザインを目指す（図2）ことも可能となることでしょう。



■論文情報

論文名：Genome-wide *cis*-decoding for expression design in tomato using cistrome data and explainable deep learning

掲載紙：The Plant Cell

著者：Takashi Akagi*, Kanae Masuda*, Eriko Kuwada* (*Contributed equally), Kouki Takeshita, Taiji Kawakatsu, Tohru Ariizumi, Yasutaka Kubo, Koichiro Ushijima, Seiichi Uchida

DOI：10.1093/plcell/koac079

■研究資金

本研究は、科学技術振興機構（JST） 戦略的創造研究推進事業 さきがけ「植物分子の機能と制御（研究総括：西谷 和彦）」における研究課題「ゲノム・遺伝子倍化が駆動する植物分子の新機能の探索とデザイン（JPMJPR20Q1）」（研究者：赤木剛士、研究期間：2020年12月～2024年3月）、新学術領域「植物新種誕生の原理」における「植物における性表現の揺らぎを成立させる進化機構（19H04862）」（研究者：赤木剛士、研究期間：2019年4月～2021年3月）、先端バイオイメージング支援プラットフォーム（ABiS）（JP16H06280）（研究者：内田誠一、研究期間：2019年～2022年



PRESS RELEASE

3月)の支援を受けて実施しました。

■補足・用語説明

(注1) モデル植物

植物学において、多様な種の全てを研究することが難しいため、それらを代表して共通した生命現象を解明する「モデルとなる研究」のために用いられる植物。

(注2) 遺伝子編集

生物のゲノム配列上にある「特定のDNA配列を狙って変化」させる技術。従来は、DNA配列に「ランダムに変化を起こす技術(あるいは自然に起こったもの)」が作物の育種で使われてきたが、遺伝子編集技術の開発によって飛躍的に標的遺伝子の改良効率が向上した。

<お問い合わせ>

岡山大学 学術研究院環境生命科学学域
研究教授 赤木 剛士 (アカギ タカシ)
(電話番号) 086-251-8337



岡山大学は持続可能な開発目標(SDGs)を支援しています。