

研究計画・目的

国内の亜熱帯地方において未利用となっている植物遺伝資源を調査・サンプリングすると共に、収集資源を用いた高効率な物質生産基盤の整備と次世代型ゲノム育種法の確立に向けたデータサイエンスに基づく研究課題を実施する。収集する植物サンプルについては、代謝や遺伝子発現などの高品質オミックス情報解析、画像解析などの深層学習、AI テキストマイニングを実施し、環境応答性や成長、物質生産に関わる予測技術を開発する。学術文献に対する自然言語処理と網羅的オミックス解析の融合により、生命科学における新たなデータサイエンス基盤を創出・標準化する。さらに、物質生産性に関わる因子（遺伝子、化合物など）をネットワーク化することで、ゲノム改変により物質生産能力を向上・デザインするための知見を導出する。また、取得した大規模・網羅的情報を統合した知識ベースを構築する。

これらの目標達成を加速化するために、沖縄県の本島・離島における調査・サンプリング、および、国際農林水産業研究センター（石垣島）テキサス州立大学（アメリカ）、トロント大（カナダ）、成功大学（台湾）、パヤオ大学（タイ）等との共同研究を展開する。

研究活動

獲得している外部資金である JST 未来社会創造事業（2件）、JST 産学共創プラットフォーム共同研究推進プログラム（1件）、NEDO ムーンショット、科研費・基盤（B）などの研究課題と密接に連携しつつ、我が国が保有する有用遺伝資源の探索を展開した。

研究目標達成に向けたデータサイエンスによる実施項目としては以下を展開した。AI による画像解析基盤を整備し、生体内の有用成分含量を予測するための解析プロトタイプ構築を目指すと共に、世界に先駆けて整備・活用してきた AI テキストマイニング解析基盤を拡張し、学術文献から有用遺伝子の生物学的機能を迅速かつ高精度に抽出するための解析システムの開発を目指した。また、大規模ゲノム解読技術を整備し、世界の食糧生産量の飛躍的向上に貢献し得る新たな農作物ゲノムデザインに向けたオミックス解析の確立を目指した。

また、フィールド利用による研究目標達成として、沖縄県の離島における遺伝資源の調査・サンプリングを実施した。また、国際農研（石垣島）において、イネ栽培品種を栽培・調査・交配を行い、有用形質に関わる遺伝子探索を目指した。

研究成果

データサイエンスのアプローチから、AI による画像解析基盤を整備し、生体内の有用成分含量を予測するための解析プロトタイプを構築した。また、AI テキストマイニング解析基盤を拡張し、テキストマイニングにおいて、新規の遺伝子名や化合物名にも対応して、それらの生物学的機能が記述されている文を抽出・要約するためのプロトタイプを開発した。また、スパコンを活用し、大規模ゲノム解読技術を整備し、巨大ゲノムをもつ生物種のゲノム解読を行った。

フィールド調査からは、我が国の亜熱帯地域に自生する有用植物種の調査を行った。そして、国際農研（石垣島）において、亜熱帯地域に自生する樹木の枝や果実を採取し、それらを用いた栽培試験を開始している。イネ栽培品種を栽培・調査・交配を行い、今年度も引き続き、有用形質に関わる遺伝子探索を進めている。

以上の研究活動により、査読付き論文発表（3報）、著書出版（英語、1冊）、シンポジウム主催および話題提供（2件）などを行った。著書出版は、植物バイオインフォマティクスの初学者向け教科書の位置付けであり、エディターとして内容の取りまとめを行うと共に、3章の執筆を行った。

査読付論文

1. Shizuka Koshimizu, Naoki Minamino, Tomoaki Nishiyama, Emiko Yoro, Mayuko Sato, Mayumi Wakazaki, Kiminori Toyooka, Kazuo Ebine, Keiko Sakakibara, Takashi Ueda*, Kentaro Yano* (*; Corresponding author) (2022) Phylogenetic distribution and expression pattern analyses identified a divergent basal body assembly protein involved in land plant spermatogenesis. *New Phytologist*(2022). 236(3):1182–1196. 10.1111/nph.18385 (16 July 2022).
2. Saishu Yoshida, Hideaki Yurino, Masaaki Kobayashi, Naoto Nishimura, Kentaro Yano, Ken Fujiwara, Shin-ichi Hashimoto, Takako Kato, Yukio Kato (2022) Expression and localization of tight junction-related proteins in adult rat pituitary stem/progenitor cell niches. *Journal of Reproduction and Development*. 68(3):225-231. 10.1262/jrd.2021-150 (01 June 2022).
3. Hideki Takanashi, Hiromi Kajiya-Kanegae, Asuka Nishimura, Junko Yamada, Motoyuki Ishimori, Masaaki Kobayashi, Kentaro Yano, Hiroyoshi Iwata, Nobuhiro Tsutsumi, Wataru Sakamoto (2022) DOMINANT AWN INHIBITOR Encodes the ALOG Protein Originating from Gene Duplication and Inhibits AWN Elongation by Suppressing Cell Proliferation and Elongation in Sorghum. *Plant and Cell Physiology* pcac057. 10.1093/pcp/pcac057 (30 May 2022).

著書

1. *Plant Omics Advances: in Big Data Biology*. Hajime Ohyanagi, Eiji Yamamoto, Ai Kitazumi, Kentaro Yano 編集 CABI, 2022/12/14 Hardback: ISBN : 978-1-78924-751-0 • 270 pages • 2022
2. Ai Kitazumi, Isaiah C.M. Pabuayan, Kevin R. Cushman, Kentaro Yano, Benildo G. de los Reyes (2022) Plant Transcriptomics: Data-driven Global Approach to Understand Cellular Processes and Their Regulation in Model and Non-Model Plants. *Plant Omics Advances: in Big Data Biology*. Hajime Ohyanagi, Eiji Yamamoto, Ai Kitazumi, Kentaro Yano CABI, pp.10-29. 12 December 2022
3. Miyu Asari, Ai Kitazumi, Eiji Nambara, Benildo G. de los Reyes, Kentaro Yano (2022) Plant Gene Expression Network. *Plant Omics Advances: in Big Data Biology*. Hajime Ohyanagi, Eiji Yamamoto, Ai Kitazumi, Kentaro Yano CABI, pp.137-150. 12 December 2022
4. Eiji Nambara, Dawei Yan, Jing Wen, Arjun Sharma, Frederik Nguyen, Ange Yan, Karin Uruma, Kentaro Yano (2022) Plant Hormones: Gene Family Organization and

Homolog Interactions of Genes for Gibberellin Metabolism and Signaling in Allotetraploid Brassica napus. Plant Omics Advances: in Big Data Biology. Hajime Ohyanagi, Eiji Yamamoto, Ai Kitazumi, Kentaro Yano CABI, pp.151-171. 12 December 2022

今後の展望

グリーンバイオ x データサイエンスのアプローチによる SDGs の達成に向けた国策が多くの国で展開しており、我が国は、その両面において遅れを取っている。我が国が保有する植物遺伝資源は豊富であり、バイオエコノミーやバイオファウンドリー、二酸化炭素資源化、食糧増産、環境浄化などを推進するための遺伝子資源も多く含まれると期待される。この我が国の植物遺伝資源と、AI テキストマイニング、AI 画像解析、高品質オミックスデータ解析を有機的に組み合わせることにより、グリーンバイオ DX が促進されることが明らかである。今後、在外研究機関中に取得した植物資源やオミックス情報、知識情報を統合解析することにより、植物を活用した SDGs 達成を目的とした研究を展開する。

教育への効果

我が国は、データサイエンスを推進可能な人材が少なく、植物科学分野において、それらに精通する人材の教育と輩出が求められている。講義において、基礎的な AI ツールの利活用などのリテラシー教育を実践することにより、AI を積極的に活用し、さらには、開発にも貢献する人材を育成する。

特に、データサイエンスの中心的な解析手法である機械学習や AI は、従来の統計数理ではモデル化が困難であった複雑な現象を扱い得る手法として注目されている。蓄積している多様な生物種のゲノム情報をビッグデータとして用いることで、たとえば、SDGs の達成に資する革新的な農業資材（品種、植物の生長を促進する微生物利用など）の開発や石油資源からのシフトが可能となると期待されている。しかし、生物学分野において、これらの手法の適用によるめざましい成果はない。このことは、これらの手法が、従来の生物学分野で扱う以上の規模の実験データ、いわゆるビッグデータを要求することに起因する。推進している高品質ビッグデータの整備では、その蓄積量（現時点で 15 植物種・8 万 run 以上の RNA-Seq データ）は世界的にも突出した規模となっている。これらの高品質ビッグデータを AI を用いてハンドリングするための教育が展開できる。

より具体的には、人類が突破すべき喫緊の課題（SDGs など）のうち、植物科学・生命科学・農学分野が取り組むべき課題として、脱化石資源による温暖化抑制、食糧の量産と質向上、環境保全（土壌・海水・生態系）、フードロスのないサプライチェーンの確立などがあるが、これらの問題解決の糸口として、光合成植物の生産能の高度利用化が注目されている。光合成植物は、水と二酸化炭素を源材料として光エネルギーにより炭水化物を合成する能力をもつ。さらに、植物は、窒素、リン酸、カリウムを始めとする無機質肥料を吸収し、複雑な有機化合物を合成する代謝システムを進化させてきた。環境・生態系において、植物は二酸化炭素を資源化する、すなわち、二酸化炭素を原材料として活用し、合成した化合物を生体内に蓄積する機能を唯一保有する生物種である。植物の生成物質は、人類活動に不可欠な食糧や有用物質（ゴム、

燃料、医薬品、パルプなどの原料・資材)として活用されてきた。特に、これらの有用物質には、人工的に合成困難な物質も多く含まれる。このことから、植物を有用物質の生産デバイス(バイオファウンドリー)として活用するバイオエコノミーの拡大が国家戦略として多くの国で進められている。在外研究活動により探索した有用植物遺伝資源をデータサイエンスによりマイニングし、有用遺伝子の探索とバイオファウンドリー・高度利用化に向けた教育・研究が可能となる。