

厚生労働科学研究費補助金（食品の安全確保推進研究事業）  
新たなバイオテクノロジーを用いて得られた食品の安全性確保と  
リスクコミュニケーションのための研究

総合総括研究報告書

研究代表者 近藤 一成 （昭和女子大学）

研究要旨：

リスクコミュニケーション分野では、合成生物学やフードテックと呼ばれる方法で作られた食品が注目を集めている。昆虫食、植物由来代替肉、培養肉、代替乳製品（人工乳、精密発酵乳）、微細藻類の代替タンパク質に着目して5000人規模のアンケートを実施した。培養肉、精密発酵乳の認知度が低い事が分かった。代替タンパク質に関するオンラインセミナーの結果を基に平易な冊子を作成した。代替タンパク質の日本での社会実装が進む際にリスクコミュニケーションの教材となることが期待される。

ゲノム解析分野では、ゲノム編集食品の安全性評価の一つである外来遺伝子配列のゲノム上における残存を網羅的に調べる方法として、次世代シーケンサーを利用して得られた全ゲノムシーケンサーデータを用いた標準的解析手法の開発に取り組んだ。実際に残存が想定される外来遺伝子配列（Cas9 配列など）について手法の妥当性を検討すると共に、このアセンブリ法によって解析可能な残存配列の長さや解析で必要とされる全ゲノムシーケンサーにて取得すべきデータ量（シーケンサーカバレッジ）を明らかにした。ゲノム編集食品の安全性評価の精緻化・向上のみならず、こうした新たなバイオテクノロジーを活用した食品に対する国民受容の向上にも役立つと期待される。

網羅的代謝物解析分野では、スペクトル類似性に基づいた代謝物同定・推定手法を開発した。化合物単位での解析に加え、分析データからエンリッチメント解析を通じて代謝パスウェイにおける変動を明らかにする検討を行った。開発したスペクトル類似度計算に基づく未知化合物の構造解析および可視化を行う解析ツールを、ウェブブラウザ上で動作する Docker イメージとして配布、GitHub で公開した。これにより、未知化合物の迅速な解析と可視化という従来高度な質量分析とインフォマティクス技術が必要だった解析が広範な研究者および技術者にも利用可能となった。

アレルギー性予測の分野では、既に開発したサポートベクターマシンを用いた手法である allerSTAT について、客観的性能評価のために F1 スコア、MCC で評価した結果、既存のツールよりも優れていることが確認できた。また、食物由来タンパクの主要組織適合性複合体 HLA への結合性を予測する手法を検討して、既存深層学習モデルをベースに追加の特徴量を組み合わせたトレーニングを行い、予測性能を比較、性能向上の可能性と有効な特徴量を検証した。

本研究班は、3か年に渡り研究代表者を含む以下の8名から構成され各分担課題に取り組んだ。

研究分担者	安達 玲子	(国立医薬品食品衛生研究所)
研究分担者	柴田 識人	(国立医薬品食品衛生研究所)
研究分担者	爲廣 紀正	(国立医薬品食品衛生研究所)
研究分担者	吉場 聡子	(国立医薬品食品衛生研究所)
研究分担者	小泉 望	(大阪公立大学)
研究分担者	早川 英介	(沖縄科学技術大学院大学)
研究分担者	富井 健太郎	(産業技術総合研究所)

厚生労働科学研究費補助金（食品の安全確保推進研究事業）  
新たなバイオテクノロジーを用いて得られた食品の安全性確保と  
リスクコミュニケーションのための研究

総合報告書

研究要旨：

リスクコミュニケーション分野では、（１）代替タンパク質の社会実装の動向調査、（２）日本における代替タンパク質に対する消費者の認識の定量調査を行った。昆虫食、植物由来代替肉、培養肉、代替乳製品（人工乳、精密発酵乳）、微細藻類の５つの代替タンパク質に着目したアンケートから、培養肉、精密発酵乳の認知度が低い事が分かった。植物由来代替肉、微細藻類、昆虫食は既に流通している。昆虫を食べることに対する忌避感などはあっても危険視する傾向は見られない。植物由来肉はポジティブな言葉を連想、培養肉はネガティブな言葉を連想しやすい傾向がある。５つの代替タンパク質に関するオンラインセミナーの結果を基に平易な冊子を作成した。代替タンパク質の日本での社会実装が進む際にリスクコミュニケーションの教材となることが期待される。

ゲノム解析分野では、ゲノム編集食品中のゲノムに残存する外来遺伝子配列の NGS を用いた標準的解析方法として、残存する可能性のある配列情報があれば、アセンブリ法によって残存の有無、残存配列の内容、位置を明らかできること、さらに既知の解析法との比較検討から、「k-mer 法による外来遺伝子配列の有無の検討→（検出されれば）アセンブリ法による配列内容と位置の同定」という評価スキームを提唱した。

網羅的代謝物解析分野では、意図しない代謝物の有無を明らかにするために、未知化合物の迅速な解析と可視化という従来高度な質量分析とインフォマティクス技術が必要であった問題を解決する手法を開発した。独自に開発したスペクトル類似度計算に基づく未知化合物の構造解析および可視化を行う解析ツールを、ウェブブラウザ上で動作する Docker イメージとして配布、GitHub で公開した。多くの研究者および技術者にも利用可能となった。

トランスクリプトーム分野では、mRNA レベルの変化を網羅的に検証し、ゲノムのオフターゲット解析やアレルゲン予測では捕捉できない変化を検出できる可能性がある。モデルサンプルとして、届出済みゲノム編集食品を対象とし、未編集トマトと編集済トマトの青果のトランスクリプトームから遺伝子発現の比較解析を試みた。

アレルゲン性予測の分野では、既に開発したサポートベクターマシンを用いた手法である allerSTAT について、客観的性能評価のために F1 スコア、MCC で評価した結果、既存のツールよりも優れていることが確認できた。一部の作物の予測性能のさらなる向上のために深層学習の手法でも検討を行い、これまでの方法より予測精度が向上することを見出した。また、食物由来タンパクの主要組織適合性複合体 HLA への結合性を予測する手法を検討して、既存深層学習モデルをベースに追加の特徴量を組み合わせたトレーニングを行い、予測性能を比較、性能向上の可能性と有効な特徴量を検証した。

A. 研究目的

遺伝子改変技術を応用した食品開発は、技術的には外来遺伝子導入による遺伝子組換え食品（GM 食品）から生物自身が持つ内在性遺伝子改

変で新たな形質を生み出すゲノム編集技術応用食品（ゲノム編集食品）へ、また、その生物が持たない多数の遺伝子を導入した酵母などから新規食品機能成分を産生させる合成生物学利用も諸外国を中心に進んでいる。ゲノム編集技術では DNA 2 本

鎖切断を誘導するオリジナル手法から、1本鎖切断から1塩基編集を行う塩基置換編集（Base editing）、これを発展させ数塩基の自由な組合せの塩基編集（prime editing）、さらに標的配列への制限をなくしたPAMレス（PAM配列を要求しない）編集、RNA編集など非常に多様な手法が生み出され、そこから想定される意図しない変化も一様でないことが明らかになりつつある。したがって、配列に依存しない意図しない塩基変化やそこから生じる代謝成分の変化を網羅的に検出または予測し、その変化が与える影響を正確に評価することは、食品の安全性確保において急務の課題である。また、これらゲノム編集食品や合成生物学利用食品に加えて、細胞培養によって作成される肉、魚、乳、チョコレートなど細胞性食品の研究開発が急速に進んでいる。ますます、リスクコミュニケーションの重要性が増している。国民受容とともに製品開発や普及が並行して進むことで、規制精度の整備、国民理解と受容、イノベーション推進が進むことが望まれる。

本研究では、（1）多様な遺伝子改変技術と開発に関する情報収集、（2）一様でない意図しない変化の影響解析のための手法開発（ゲノム、代謝成分、アレルゲン性）、（3）ゲノム編集食品の理解の前段階として不可欠な国内GM食品利用の現状と審査届出制度の理解に重点したリスクコミュニケーション（若手研究と連携）、（4）リスク評価側の最新技術理解と能力向上、人材育成を柱に若手研究代表者とも連携して実施する。

### リスクコミュニケーション分野での研究目的

新たなバイオテクノロジーによって作られる食品としては遺伝子組換え食品が既に25年以上利用されているが、そのリスクコミュニケーションが困難な状況が続いている。単なる科学的な説明に加えてELSI（倫理的・法的・社会的課題）が関与しコミュニケーションを複雑にしている。2021年秋に我が国においてゲノム編集技術応用食品

（ゲノム編集食品）の実用化が始まった。ゲノム編集食品に関しては、その実用化の数年前から行政、国の研究機関、大学、NPO、開発者により設立されたベンチャー企業などがコミュニケーション活動に積極的に関与してしたこともあり比較的冷静なリスクコミュニケーションが行われている。つまり、社会実装に先行したコミュニケーションが有効と考えられる。

近年、世界人口の増加と生活レベルの向上に伴う肉食の増加によるタンパク質クライシスが懸念されている。また家畜飼育による温室効果ガスやふん尿由来などの環境負荷も問題となっている。そのため代替タンパク質に注目が集まっている。注目されている代替タンパク質には植物由来代替肉、昆虫食、微細藻類、培養肉、代替乳（精密発酵乳）などがあり、本研究ではこの5つを対象とした。培養肉、代替乳（精密発酵乳）は細胞性食品とも呼ばれ日本では実用化されていない。培養肉も本研究が始まった当初には販売例は無かった。代替乳（精密発酵乳）は一部の国でアイスクリームの原料などとして使用されていた。

以上、例を挙げてきたような新たなコンセプトで作られた食品が主として国外で次々と開発され実用化されているが、その国内における認知度は低いと考えられる。近い将来こうした食品が国内でも流通する可能性が予想されることから、その効果的なリスクコミュニケーションのために1）代替タンパク質の社会実装の動向調査、2）日本における代替タンパク質に対する消費者の認識の定量調査、3）代替タンパク質に関するワークショップ（グループインタビュー）、4）代替タンパク質に関する冊子の作成を行った。

### ゲノム解析分野での研究目的

我が国では2019年よりゲノム編集食品の事前相談・届出制度が開始されており、これまでに6品種がゲノム編集食品として届出・公表され、一部はすでに流通されている。この制度では、届出の

対象となるゲノム編集食品の備えるべき要件の一つとして、ゲノム中に外来遺伝子を全く含まないことを求めているが、「ゲノム編集技術応用食品等の取扱いに関する留意事項」ではこうした外来遺伝子およびその一部が残存していないことを調べる方法として、次世代シーケンサー（NGS）による全ゲノムシーケンス（WGS）、サンガーシーケンス、サザンブロット法などを挙げている。このうち、得られるデータの網羅性や、ゲノム編集技術によって生じる意図しない遺伝子変化全般の解析にも適用できる汎用性などを考え、次世代シーケンサーによる解析が着目されており、その解析結果の届出情報への活用が今後増加すると予想される。但し解析の必要要件や標準の手順が明確になっていないことから、NGS 解析を公定試験法とする上で検討すべき課題となっている。本研究では、NGS による WGS データを用いた、ゲノム編集食品の外来遺伝子の有無を調べる標準的な解析手法の確立およびその手法に関する必要要件の検討を試みた。

### 代謝物解析分野での研究目的

本研究の目的は、ゲノム編集作物など新たな技術で作成される様々な食品において代謝の変化などで懸念される未知化合物等の様々な化合物を迅速に検出するための新たなデータ解析プラットフォームの開発にある。このプラットフォームは、液体クロマトグラフィーと質量分析技術を組み合わせたハイスループットな分析データから、構造類似性を反映したスペクトル類似度を解析することで未知化合物を含む様々な食品中の化合物の構造解析を直感的に可能にする。特に、従来のターゲット分析では検出困難な未知化合物に対して、自動的に化合物クラス情報、部分構造、毒性情報を提供し、ゲノム編集食品の安全性検証に新たなアプローチを提供することを目指している。この研究により、食品安全性の検証プロセスの高度化および効率化を図り、消費者への安全な食品供給

に貢献することを目指す。

### トランスクリプトーム分野での研究目的

本分担研究では、ゲノム編集による意図しない変化のうち、トランスクリプトーム変化（遺伝子発現の変化）に着目した。GABA トマトをモデルサンプルとして、野生型トマトと GABA トマトの遺伝子発現を比較することで、mRNA レベルの変化を予測や先入観なく網羅的に検証し、ゲノム解析やアレルゲン予測では捕捉できない可能性のある変化の掘り出しとこれらの変化とのつながりについて検証しようとした。

トランスクリプトーム比較解析で新たに得られる可能性のある情報として、次を想定した。1.オンターゲット変異による代謝経路の変化に伴う間接的な遺伝子発現の変化 2.オフターゲット予測から漏れたオフターゲット変異による遺伝子発現の変化 3.Cas9 の持続的な作用による遺伝子発現の変化（DNA 損傷シグナルの不活化等）。

本研究は、ゲノム編集食品の編集前後の想定しない変化とその影響について、トランスクリプトーム比較解析により、網羅的に検証し評価することで、ゲノム編集技術の安全性評価において新たな問題提起につながる事象を探ること、さらにゲノム編集食品のトランスクリプトーム解析に必要な要件及び課題について議論・検証することを目的とした。

### アレルゲン予測とデータベース分野での研究目的

遺伝子改変技術を応用した食品開発は、技術的には、外来遺伝子導入による遺伝子組換え食品から、内在性遺伝子の改変を行うゲノム編集技術応用食品へ、また、酵母等に多数の外来遺伝子を導入し新規食品機能成分を産生させる合成生物学の利用へと変化している。現在、ゲノム編集技術では多様な手法が生み出されており、これらの手法による意図しない塩基変化も一様ではないことが明らかになりつつある。従って、このような意図

しない変化、及びそこから生じる代謝成分の変化を検出または予測し、その変化が与える影響を正確に評価することは、食品の安全性確保において急務の課題である。

バイオテクノロジー技術を用いて開発された食品のリスクの1つに、アレルゲン性増大の可能性がある。本研究では、国立医薬品食品衛生研究所生化学部にて管理・公開している、アレルゲン性予測機能(FAO/WHO法等)を装備したアレルゲン・エпитープ情報データベース(Allergen Database for Food Safety, ADFS)について、新規アレルゲン及びエピトープ情報の収集・解析等によりアレルゲン性評価に関する検討を行い、遺伝子改変技術応用食品のリスク評価に資するデータベースとなるよう、情報を更新し内容を充実させる。

また、AIを活用した新規高精度アレルゲン性予測手法の開発を進める。令和2年度までの先行研究班では、機械学習によりアレルゲン及び非アレルゲンタンパク質から抽出した特徴的なアミノ酸配列パターンを利用してアレルゲン性を予測する手法(アレルゲン性予測手法、allerStat)を開発してきた。本研究班では、この予測システムの機能を拡充し、詳細な性能検証を行い、高精度アレルゲン性予測法としての実用化を進める。

また、医療・食品分野やリスク評価分野等におけるAIの活用実態を調査し、課題を整理する。

## アレルゲン予測 (HLA 結合性予測) の研究目的

遺伝子組み換え作物 (GMO) に対してはアレルゲン性が遺伝子改変食品の安全性と関連付けて問われているが、すべての遺伝子組換え食品のアレルゲン性を実験的に評価するのはコストの面から現実的ではない。このため、科学的根拠をもつ信頼性のある評価方法の確立が求められている。組換え DNA 技術で導入した新規遺伝子産物(タンパク質)や形質転換による意図しない新規タンパク質のアレルゲン性予測方法としては、FAO(国連食糧農業機関)/WHO(世界保健機関)が提唱しているデータベースに登録済みのアレルゲンタンパク質との相同性比較が標準的に使用されている。しかし、配列長が短い既知アレルゲン性ペプチドとの類似性に基づくため偽陽性が高いことが指摘されている。また、進化に基づくアミノ酸置換行列を用いる配列類似性比較は、オフターゲット効果による変異をもつ新規タンパク質に対するアレルゲン性の判定には十分ではない可能性がある。

本研究では、標的配列と類似した配列のオフターゲット検索に限定されている点を克服すべく、人工知能を活用して類似性がないアレルゲン性タンパク質由来ペプチド-HLA クラス II 分子間結合予測法の開発(図1)とこれを用いたアレルゲン性予測法の開発を行うことを目的に取り組んだ。

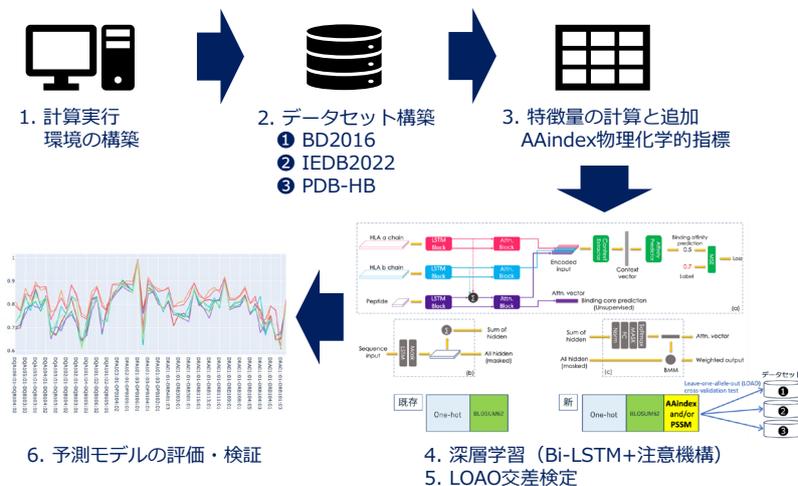


図1 結合予測用深層学習モデル

## B. 研究方法

本研究班構成では、意図しないゲノム DNA 配列の変化の解析手法開発と標準化を柴田が、ゲノム編集食品のトランスクリプトーム解析を吉場が、意図しないタンパクの生成に伴うアレルギー性の評価手法開発と実用化およびアレルギーデータベース ADFS の維持更新を、深層学習も取り入れながら安達および富井が、また、意図しない代謝物変化の網羅的開発手法の開発と Web 環境で利用できるような実用化を早川が担当した。リスクコミュニケーションについては、ゲノム編集食品、合成生物学利用食品、特に代替タンパク質に重点を置きながら小泉が担当した。

### リスクコミュニケーション分野

#### 1) 代替タンパク質の社会実装の動向調査：

複数のオンラインセミナーへの参加、フードテック専門メディア Foove (<https://foodtech-japan.com/>) への会員登録、フードテック官民協議会への会員登録、バイオインダストリー協会 Food Bio Plus への会員登録などにより適宜、情報収集を行った。三菱ケミカルリサーチに一部調査を依頼した。

#### 2) 日本における代替タンパク質に対する消費者の認識の定量調査：

5つの代替タンパク質(植物由来代替肉、昆虫食、微細藻類、培養肉、代替乳(精密発酵乳))に関する一般的なイメージや受容、その期待などを比較するために、オンラインにて質問紙調査を実施した。調査は代替タンパク質のそれぞれの短い説明文を事前に読んだ上で、全部で17問の質問項目に選択肢を選んで回答してもらった。調査は調査会社(楽天インサイト)に委託し、代替タンパク質の消費者を想定し、日本全国の20歳から65歳の男女5,000人の登録モニターを対象として実施した。

#### 3) 代替タンパク質に関するワークショップ(グ

ループインタビュー)：

2021年度に精密発酵を中心に、オンラインワークショップを実施した。多様なステークホルダーの受け止め方の調査としては2回のワークショップをオンラインにて実施した。1度目は、社会心理学、科学技術社会論、分子生物学の研究者、NPO(バイオ系)職員、生協職員、消費者団体職員、一般主婦が参加した。2回目は科学技術社会論、栄養科学、分子生物学の研究者、一般主婦が参加した。ワークショップの冒頭で代替タンパク質(特に精密発酵乳)について概説し、そうした食品が日本に導入される際に起こると考えられる問題について議論した。2022年度には計6回のワークショップを行った。毎回3名の30代~50代の一般主婦(楽天インサイトに登録しているモニター)と代替タンパク質の専門家が参加し意見交換を行った。うち1回は代替タンパク質が登場した背景を説明する総論である。それぞれのワークショップには異なるモニターが参加した。2023年度には50~60代の食に詳しい主婦、具体的には消費生活アドバイザー、生協職員、大学教員が参加し精密発酵乳に特化したワークショップをオンラインで実施した。

#### 4) 代替タンパク質に関する冊子の作成：

3)で述べたオンラインセミナーの文字起こしを行い、内容を簡潔にまとめるとともに写真、イラストを入れた冊子を作成した。

### ゲノム解析分野

#### (1) 外来遺伝子残存モデルサンプルにおけるWGSの実施と残存性の検出方法の検討

外来遺伝子残存モデルサンプルとして、除草剤耐性遺伝子組換えダイズRRS2系統およびチョウ目害虫抵抗性トウモロコシMIR162系統を選択し、大規模に塩基配列を解読するNGSとして、ショートリードシーケンサーではHiSeq X Ten(Illumina)(150bp × 2のペアエンドシーケンス)を、ロングリードシーケンサーでは

PromethION (Oxford Nanopore Technologies) にて WGS を実施した。得られた WGS データをもとに、論文などで報告のある種々の変異検出ツールを用いた構造変異 (Structural variant, SV) 解析を実施した。なおマッピングツールについても種々の方法が報告されているが、本研究では、ショートリードシークエンスデータには BWA (ver. 0.7.17-r1188) の MEM アルゴリズム、ロングリードシークエンスデータには Minimap2 (ver. 2.21-r1071) にて実施した。またこの WGS データをもとに、アセンブリツールとして、ショートリードシークエンスデータでは Velvet (ver. 1.2.10) または SPAdes (ver. 3.15.5)、ロングリードシークエンスでは Flye (ver. 2.9) を用いた残存配列の再構成を試みた。

(2) ゲノム編集食品で残存する可能性のある配列を組み込んだシミュレーションサンプルによる、アセンブリ法の適用

ゲノム編集技術による標的ゲノム配列の改変では部位特異的ヌクレアーゼが使用されることから、ゲノム編集食品で残存する可能性の高い外来遺伝子配列は、この部位特異的ヌクレアーゼおよびその導入に使用されたバックボンプラスミドである。そこで、ゲノム編集で利用されることが多い化膿性連鎖球菌の Cas9 ヌクレアーゼ (SpCas9) の遺伝子配列を含む植物向けゲノム編集ベクター pDGE64 (#79446) の全長またはその部分配列をダイズ参照ゲノム配列に組み込んだ人工配列を作成し、それを基に、ショートリードシークエンスおよびロングリードシークエンス各々の疑似 WGS データを生成した。この疑似 WGS データを用いて、アセンブリ法に供することで、残存性検出の可否、検出できる配列の長さ、検出に必要な WGS のデータ量 (シークエンスカバレッジ) などを調べた。

(3) 既存の解析方法との性能比較

ゲノム編集食品に残存した外来遺伝子配列を NGS によって解析する方法として、WGS データについて残存が想定される配列の断片と照合する k-mer 法と、残存が想定される配列の断片と相補的な DNA 断片を集めて NGS によって解析する capture-based target enrichment 法が報告されている。このうち、本研究課題で我々が検討しているアセンブリ法と同様に、k-mer 法は WGS データの取得までに特別な作業を必要とせず、同じデータセットを用いた解析が可能であったことから、アセンブリ法との性能比較を試みた。解析には Cas9 ヌクレアーゼの活性ドメイン HNH ドメインの部分配列である 21 または 31 塩基の配列が挿入された疑似 WGS データおよび野生型の疑似 WGS データを供した。解析ではどちらも 30×カバレッジにダウンサンプリングしたデータセットを使用した。

## 代謝物解析分野

質量分析技術を利用して化合物から得られる質量スペクトルの類似性を基に、試料中の未知化合物の迅速な構造情報の取得及び可視化を可能にする新たな手法とその解析ツールを開発した。具体的には、化合物の質量スペクトルデータから、構造の類似性を反映するスペクトル類似性に基づき、ネットワーク構造を生成するアプローチを採用。このネットワーク試料中の未知化合物を試料レイヤー、標準物質のスペクトルライブラリ由来のスペクトルをリファレンスレイヤーに配置することで、未知化合物を含む複雑な試料中の化合物間の関係を明確にし、迅速な化合物クラスといった構造情報の推定を実現する。

## トランスクリプトーム分野

### 概要

ゲノム編集トマト (GABA トマト) をモデルサンプルとして、野生型および GABA トマトから total RNA を抽出し、RNA-seq による遺伝子発現比較

解析を行った。

#### サンプル

GABA トマト青果はサナテックシード社より購入した。野生型トマト(cv. シシリアンルージュ CF) は神奈川県内のスーパーで購入した。

#### Total RNA の抽出

生鮮サンプルを-30°Cで凍結した後、凍結乾燥機 (EYELA FDU-1200; 東京理化機械株式会社) により凍結乾燥した。凍結乾燥サンプルは、液体窒素を入れた乳鉢で粉碎後、Fruit-mate for RNA purification (タカラバイオ株式会社) による前処理と NucleoSpin RNA plant (Macherey-nagel GmbH & Co.KG) を組み合わせて total RNA を回収した。回収した total RNA は分光光度計 (NanoDrop; Thermo Fisher Scientific) および TapeStation (Agilent Technologies Japan, Ltd.) を用いて、濃度測定及び品質検定を行った。

#### RNA-seq

シーケンスは、次世代シーケンス (NGS) 解析サービス (タカラバイオ株式会社) により行った。作業内容は下記の通りである。

SMART-Seq v4 Ultra Law Input RNA Kit for Sequencing を用いて、RNA 増幅、Nextra XT DNA Library Prep Kit を用いて DNA ライブラリ作製。ライブラリ作製は、Smart-Seq2 法 (参考文献3) を用いている。シーケンスは、NovaSeq システム (illumina) を用いて、150base 両末端解析 (6Gb/4,000万リード;2,000万リードペア)をWT トマト及びGABA トマトそれぞれ3サンプルずつ (合計6サンプル) 実施。得られたリード配列をリファレンスゲノム配列にマッピング後、遺伝子ごとに正規化された発現量、リードカウント値及び TPM (transcript per million) 値を算出している。なお、マッピングに用いたトマトのリファレンスゲノム配列および遺伝子アノテーション情報

はそれぞれ、SL3.0、ITAG2.3、遺伝子数の合計は 35,825 である。結果のリード配列は FASTQ 形式、マッピング結果は bam 形式、遺伝子発現量データは csv ファイルで納品された。

#### 遺伝子発現解析

再現性の確認 (technical replicate) 及び3反復2群間比較による遺伝子発現量変動解析を行った。A.リードカウント値からトランスクリプトームデータ分析ウェブプラットフォーム iDEP (integrated Differential Expression and Pathway analysis; iDEP.96, 参考文献4) を用いて解析する方法と、B.TPM 値から手動で発現変動遺伝子を抽出する方法を平行して行った。

A では、遺伝子発現量データからリードカウントデータを作成し、iDEP を用いて、データの前処理と検証後、発現量変動解析、Gene Ontology (GO)エンリッチメント解析などを行った。

B では、正規化された各サンプルの TPM 値を用いて t-検定 (関連なし二群比較) を行い、下記の基準で WT トマトと GABA トマトで発現量が有意に異なる遺伝子を抽出しリスト化した。

FoldChange:  $|\log_2 TPM_{GABA} - \log_2 TPM_{wt}| \geq 1$

t-test p-value  $\leq 0.05$

t-test q-value  $\leq 0.05$

さらに、抽出した遺伝子リストを用いて、gProfiler による GO 解析を行った。

#### アレルギー予測とデータベース分野

(1) アレルギー性予測手法 (allerStat)の機能拡充及び性能検証

現システムの分類器である線形 SVM (Support Vector Machine)に加えて、非線形 SVM、ロジスティック回帰、LightGBM (Gradient Boosting Machine)、LSTM (Long short-term memory)、proteinBERT (Bidirectional Encoder Representations from Transformers) を分類器と

して追加実装し、それぞれの分類器を用いた場合の予測性について検証した。検証においては、AlgPred2、Allerdicator、Allertop、及びMEMEを比較対照とし、Leave-Category-Out Cross-ValidationにおけるROC曲線のAUC、AUC-10%、F1スコア、MCC (Matthews Correlation Coefficient)を評価指標とした。

## (2) ADFS エピトープ情報の追加

令和2年6月から令和5年5月までの3年間にNCBI PubMedに収録された論文から、キーワード検索により、エピトープ配列決定に関するものを抽出した。キーワードとしては、IgE、epitope、linear、conformational、sequence、recognition等々のワードを使用し、これらを複数組み合わせで6通りの検索式を作成して検索を行った。この検索により抽出されてきた論文についてピアレビューを行った。その結果エピトープ情報を報告していると判断された論文について、そのエピトープ情報を整理し、ADFSのデータに追加した。

(3) FAO/WHO法アレルギー性予測ツール改修  
ADFSに搭載しているFAO/WHO法アレルギー性予測ツールについて、従来使用していたFAO/WHO法の改変法から本来のFAO/WHO法に改修し、アレルギー性予測ツールとしてより精密な解析情報を得られるよう改良した。

(4) AIのリスク評価分野への応用に関する調査  
Web of Scienceにて、AI (artificial intelligence)、machine learning、deep learning、neural networkのキーワードを用いて、2020年以降に出版された総説を検索し、ヒットした14,626報の中から、引用数及び内容を考慮して48報(食品分野18報、医療分野20報、生物学分野10報)を選択し、その内容を精査した。続いて、精査した総説において出現頻度及び重要度の高い機械学習・深層学習関連の15種の用語 (Neural network、Machine

learning、artificial intelligence、Random forest、Support Vector Machine、Deep Learning、Logistic regression、Ensemble、Principal component analysis、Boosting、k-nearest neighbor、Decision tree、Naive Bayes、Long Short Term Memory、Natural language processing)のいずれかを含み、かつ“risk assessment”または“safety assessment”をキーワードとして含む2020年以降に出版された論文を、Web of Scienceにて検索したところ、5,181報がヒットした。その中から、引用数及び内容を考慮して65報(食品分野7報、医療分野58報)を選択して、その内容を精査した。

また、国際機関や各国規制当局 (世界保健機関 (WHO)・国際連合食糧農業機関 (FAO)・コーデックス委員会・医薬品規制調和国際会議 (ICH)・薬事規制当局国際連携組織 (ICMRA)・国際医療機器規制当局フォーラム (IMDRF)・米国食品医薬品局 (FDA)・米国農務省 (USDA)・欧州食品安全機関 (EFSA)・欧州医薬品庁 (EMA)・英国食品基準庁 (FSA)・英国医薬品・医療製品規制庁 (MHRA)・中国国家衛生健康委員会 (NHC))におけるAI関連の取り組みについても調査を行った。

## アレルギー予測 (HLA結合性予測)

### (1) 公的データベース

機械学習に使用するためのMHCクラスII分子の配列情報およびアレル情報を収めたデータベースからデータファイルをダウンロードした。UniProt、PDB両データベース間のエントリ情報を関連づけた。IEDBから必要な情報を含むデータをダウンロードした。

### (2) 分子間結合予測の先行研究

先行研究として、既に提案されている主な予測法について調査した。さらに、これらのうち、5つの予測法が使用したデータセットについて調査した。

### (3) 抗原ペプチドとMHCクラスII分子間相互作用解析

ペプチドと MHC クラス II 分子間相互作用の関係を調べるために、IMGT において立体構造データを検索した。さらに、PDBe PISA (Proteins, Interfaces, Structures and Assemblies) を用いて、IMGT への検索で得られた複合体構造における抗原ペプチドと MHC クラス II 分子間の相互作用状態について調査した。

#### (4) データセットの構築

以下の 3 種類のデータセットを用意した。

既存 DeepSeqPanII の学習に使用された「BD2016 データセット」を用いた。

IPD-IMGT/HLA からデータファイルをダウンロードし、対象としないエントリを取り除いた。IEDB からアレルゲン関連情報を含むデータをダウンロードした。これらをフィルタリング処理し「IEDB2022 データセット」を構築した。

IMGT ツールを用いて PDB に登録されているペプチド結合状態の HLA クラス II 分子を、検索・抽出しデータを選別した。同じく IMGT を用いて HLA クラス II 分子と結合ペプチド間において水素結合の形成に関する情報を抽出し、「PDB-HB データセット」を構築した。

#### (5) 特徴量

予測法に組み込む特徴量は以下の 5 種類を用いた (One-hot encoding と BLOSUM62 に関しては、DeepSeqPanII で採用されている)。

- One-hot encoding : 20 種類のアミノ酸を 20 次元の bit ベクトルで表現。
- BLOSUM62 : アミノ酸置換行列の値を 23 次元のベクトルで表現。
- AAindex 物理化学的インデックス : 20 種類のアミノ酸に KEGG AAindex データベースに登録されている 566 種類のインデックスの値を割り当てる。インデックス全てを次元圧縮したものと、類似したものを相関係数に基づき非冗長な 62 種類に削減したうえで次元圧縮したものを別々に特徴量として追加した。
- PSSM : マルチプルアラインメント処理によって得られたカラム重複部分の重み付きアミノ酸出現頻度として、20 次元のベクトルで表現。
- AAindex 残基間ポテンシャルインデックス : KEGG AAindex データベースに登録されて

いる 20 種類のアミノ酸残基間におけるポテンシャルのスコア行列 (47 種類) を用いて、立体構造既知の HLA クラス II 分子とペプチド間で 3.5 オングストローム以内に近接する残基ペアに対してスコアを割当。

#### (6) 深層学習ベースモデル

既存予測法の DeepSeqPanII をベースモデルとして、前節で述べた 5 種類の特徴量を組み合わせ、3 種類のデータセットごとに LOAO 交差検証テストを行った。

DeepSeqPanII は LSTM (long short-term memory) を用いているため、時系列データの学習や予測 (回帰・分類) に有利であり、モデルが複雑にならないことが期待できる。また、注意機構 (attention mechanism) を採用しており、ネットワークの内部を可視化できる長所をもつ。プログラムが GitHub 上に公開され、MIT License で再利用が認められているため、新たに別の特徴量を組み込んで利用することは問題にならない。

#### (7) LOAO 交差検証テストと AUC の算出

LOAO (leave-one-allele-out) 交差検証では、データセットに含まれるアレル分子群から 1 つのエントリだけを抜き出してテスト事例とし、残りをトレーニングとする。この処理を全てのエントリが一回ずつテスト事例となるよう検証を繰り返す。

各 HLA クラス II 分子に対し、

$$FPR = FP / (FP + TN)$$

$$TPR = TP / (TP + FN)$$

を算出し、横軸に FPR (偽陽性率)、縦軸に TPR (真陽性率) をとって描かれる ROC 曲線下面積 (AUC : area under the curve) 値に基づき予測性能を評価した。

#### (8) 新規アレルゲン性評価手法開発の基盤研究と AI のリスク評価への応用

新規アレルゲン性評価手法の開発に向けた、関連データと機械学習を活用した基盤的研究実施のために、アレルゲンと抗体タンパク質間における分子間相互作用予測法 DeepSeqPanII をもとに、アミノ酸の物理化学的インデックス AAindex を特徴量に加えた改良版深層学習モデルを構築した。このモデルをアレルゲン性予測に活用するため、予測手法の一つである NetAllergen に組み込

むことで予測性能に及ぼす影響を解析した。

## C. 研究成果

### リスクコミュニケーション分野

#### 1) 代替タンパク質の社会実装の動向調査：

世界人口の増加によるタンパク質クライシスに備えるために肉食の多い欧米（西洋）を中心に代替タンパク質の研究開発が行われており、米国ではバーガー向けなどの消費が多い。日本国内でも大豆ミートと呼ばれる食品が研究開始後も増加しているが、健康志向の側面が大きい。日本には大豆を食べる文化があることもあり、消費者の大豆ミートの受容はスムーズである。昆虫食は国内外で複数のベンチャーが手掛けているが、国内では2022年11月に学校給食にコオロギパウダーを使用したことに対しネット等で批判も見られた。微細藻類も国内外で研究開発が進んでいるが、食品としての利用は限定的である。細胞性食品と呼ばれることも多い、培養肉と代替乳（精密発酵乳）は海外では社会実装が進みつつあるが、国内では実用化されていない。

オランダはフードテックで世界をリードしており、世界20大食品企業のうち15社がオランダに拠点を置いている。産学官連携、先端技術を国家が支援している。米国はそれ以前からフードテックが進んでおり、精密発酵で作られた Perfect Day 社などによる代替乳製品の導入が進んでいる。シンガポールは食料自給率が低いこともあり、2030年までに国民に必要な栄養の30%を国産にする「30 by 30」が国策として進められている。最初に培養肉（チキンナゲット）が販売されたのもシンガポールである。イスラエルも代替タンパク質の研究開発が盛んである。アジアでも韓国は昆虫食を国家として2014年から推進している。上述の事例は世界のフードテックのごく一部に過ぎない。フードテックのカオスマップは数カ月単位で更新必要なほど研究開発動向は目まぐるしく変化している。

代替タンパク質の社会受容を一概に論ずることはできない。植物由来代替肉、微細藻類、昆虫食に関しては既に流通している。昆虫を食べることに対する忌避感などはあっても危険視する傾向は見られない。培養肉に関しては2018年、2019年に調査が行われているが、未だ消費者の意識が十分に把握されているとは言えない。実用化されていないこともあり、その実態が消費者に分からないことが理由と考えられる。精密発酵による代替乳製品については、さらにイメージすることが難しく社会受容についての評価は現実的ではない。

国内では代替タンパク質の導入あるいは肉食を避けることへの意識が欧米と比べて進んでいないように思われる。ビーガンやアニマルウェルフェアの観点も異なるが、地球環境問題への影響についての認識が大きく異なる。酪農国であるニュージーランドでも牛や羊を飼育する酪農家にげっふ税を課することが提案されるなど、世界の多くの国で地球温暖化を中心とする地球環境の悪化に対する懸念が高まっている。国内ではこうした代替タンパク質が注目されている背景が明確に伝えられていない。

日本で代替タンパク質のリスクコミュニケーションを行う際に、いわゆる食品安全だけを議論の対象とするのか、環境問題あるいはタンパク質危機への対応の概念も含めた上での議論するかにより結果は異なると思われる。

#### 2) 日本における代替タンパク質に対する消費者の認識の定量調査：

認知度に関しては植物由来代替肉>昆虫食>微細藻類>培養肉>代替乳の順に高かった。培養肉、代替乳は流通しておらず妥当な結果となった。食べたいかどうかという設問については、植物由来肉>微細藻類>人工乳>培養肉>>昆虫食の順に受容度が高い結果となった キーワードとの関連では、やはり昆虫食についてはネガティブな言葉がより選ばれ、悪いイメージが連想されやすい傾

向にあった。培養肉は昆虫食に次いで「気持ちが悪い」という言葉が連想されており、ネーミングについて注意が必要である。

栄養価については「栄養価が高い」という言葉は、モニターの4割以上が昆虫食、植物由来代替肉、微細藻類のイメージとして選んでいた。多重対応分析の結果では昆虫食は他の食品とは別のグレードに入りネガティブな語句と関連が強い結果となった。技術と伝統（自然）、動物と植物の二つの大まかな軸がみられた。植物由来肉はポジティブな言葉を連想しやすく、微細藻類のイメージと最も近い傾向にあった。培養肉はネガティブな言葉を連想しやすく、代替乳のイメージと最も近い傾向にあった。科学的関心が代替タンパク質への嗜好と関連していることも示された。科学的関心が高いと培養肉や代替乳製品といった新しいバイオテクノロジーに対して寛容である傾向が示された。このような科学的関心が、代替タンパク質に対する態度を決める共通の重要な属性であることが分かった。

### 3) 代替タンパク質に関するワークショップ（グループインタビュー）

#### 3-1. 代替タンパク質全般

植物由来は動物由来よりも健康に良いイメージがあるかもしれないが、栄養不足による健康被害は無いのか。

世界の環境を考えると、22世紀を生きていくために、必ずしも自分が食べたいからという事ではなく考え方そのものを育てる教育も必要ではないか。自分が食べたいからとか既存の食を守るという観点からだけではなくて、SDGs等を考えたうえで食を選ぶという考え方を入れていく教育ツール（コミュニケーション）にもなり得る話題だと思う。リスクリテラシーの教材としてこの話題を使うのも面白いのではないか。例えば、「あなたが今日豆腐ハンバーグを食べることだって一つのSDGsかもしれない」という言い方もできる。

一般的に消費者は新しい技術のものは不安に思う傾向がある。新しい科学でいろいろな食品が出て、添加物などが知らないうちに出回ってしまう事に対する不安を抱く人たちに対して、その不安解消をすることが目的ならば、このような背景（新しいバイオ食品が食糧の安定供給を支え得ること等）を説明しておくことも、リスクミではないか。

培養肉の認知度が低い中で、リスクミとしては、まずは培養肉について知ってもらうこと、次に実際手に取って見てもらえるかということの二つの段階がある。リスク認知が効いてくるのが前者。後者については、環境負荷軽減、おいしい、新しい食品、などの理由で手に取る人がある。それぞれの人に合ったコミュニケーションが求められる。

#### 3-2. 人工乳/精密発酵乳に特化

精密発酵による代替乳自体の安全性に関しては官の積極的な関与が必要、社会受容に関しては民の自主的な行動が求められる。例えばカニカマは模倣食品だが受容されている。複数の企業が協力して取り組みことが重要。規制に関しては官民が協議していく必要がある。SDGsに代表されるサステナブルという観点は一見、受容にポジティブに働くように見えるが酪農家とのせめぎあい等の問題点も多い。ニュージーランドは酪農先進国であるとともに環境保護を大切にしており精密発酵を推進している。日本でも国が推進するかどうか重要なポイント。日本でもやはり、酪農との関連は複雑な問題。表示は重要な懸案事項。

#### 4) 代替タンパク質に関する冊子の作成

図1に示す冊子を作成し、冊子体を関係者に送付するとともに、研究分担者（小泉）の研究室のホームページからダウンロード可能とした。

## ゲノム解析分野

### (1) 外来遺伝子残存モデルサンプルにおける WGS の実施と残存性の検出方法の検討

まず、外来遺伝子残存モデルによる WGS データについて、宿主となる生物の参照配列に対してマッピングさせたのち、SV 解析を実施した。種々の変異検出ツールを検討したが、ショートリードシークエンスデータであれば GRIDSS (ver. 2.12.1)のみで、ロングリードシークエンスデータであれば Sniffles (ver. 1.0.12) および cuteSV (ver. 1.0.11) というツールにおいて、モデルサンプルの外来遺伝子の挿入部位を検出することができた。しかしながら、いずれのツールにおいてもモデルサンプルで想定している挿入部位以外にも多数の変異部位を検出しており、本研究の目的に適った検出方法とは言えないことが分かった。

次にこのモデルサンプルに挿入されている遺伝子配列に対して WGS データをマッピングさせ、マッピングリードを抽出し、アセンブリ解析に供した。その結果、ショートリードシークエンスおよびロングリードシークエンス WGS データともにほぼ 100%の正確性で挿入配列を再現できることが分かった。さらに WGS データをダウンサンプリングすることで、検出できる配列長や検出に必要なデータ量について検討し、ショートリードシークエンサーによる WGS では 25×以上のシークエンスカバレッジデータがあれば 40～4000 塩基程度の長さの標的配列を、ロングリードシークエンサーによる WGS では、30×程度のシークエンスカバレッジのデータがあれば、200～4000 塩基程度の長さの標的配列をほぼ 100%の確率で再現できることが分かった。他方で標的配列にマルチコピーな配列や対象生物に内在性の配列が含まれる場合には、少なくともショートリードシークエンス WGS データを利用したアセンブリ法では、解析が困難であることが分かった。なお上述の必要なデータ量の検討はダウンサンプリングにて行っているが、4000 塩基程度が挿入されたモデルサンプルにおけるショートリードシークエンサーに

よる WGS データにて、実際に 30×程度のシークエンスカバレッジのデータであればこの挿入配列を検出できることも確認した。

### (2) ゲノム編集食品で残存する可能性のある配列を組み込んだシミュレーションサンプルによる、アセンブリ法の適用

実際にゲノム編集で残存する可能性がある Cas9 遺伝子配列や薬剤選抜に使われる耐性遺伝子配列およびそれらのバックボーンプラスミドの配列を挿入させた人工ゲノム配列をもとにした疑似 WGS データをアセンブリ法に供した結果、これらの配列においてもアセンブリ法による挿入配列の検出が可能であることが分かった。WGS データからマッピングリードを集める際、前項のモデルサンプルを用いた検討では挿入している配列に対してマッピングしていたが、実際には残存が想定される配列の一部しか残存していないと考えられる。そこで本項では Cas9 遺伝子配列の一部のみを組み込んだ人工ゲノム配列をもとにした疑似 WGS データについて、Cas9 遺伝子配列などを含むプラスミド配列全体に対してマッピングし、このマッピングリードをアセンブリ法に供した。その結果、このような実際に即したケースでも本アセンブリ法は有効であること、および少なくともショートリードシークエンス疑似 WGS データについては、検出可能な最短の配列長は 31 塩基であり、その検出には少なくとも 25×カバレッジ以上の疑似 WGS データが必要であることが分かった。

なお本研究課題で検討しているアセンブリ法では複数のバイオインフォマティクスツールを使用している。従って、この解析を各実験者が実施するには、各自の解析環境にこれらのツールを導入し、適切に解析環境が構築されているか確認する必要がある。こうした確認には解析のコントロールになるような標準サンプルが使用されるが、今回作成した疑似 WGS データが利用できる可能性がある。本研究によって、ショートリードシーク

エンスであれ、ロングリードシークエンスであれ、Cas9 部分配列 (101 塩基) が挿入されたゲノムからシミュレートされた 30×疑似 WGS データより、Cas9 断片配列の検出が可能であることが分かっている。そこでこの疑似 WGS データより Cas9 断片配列の検出が検出可能かどうかを各実験者の解析環境で確認することで、この点の検証が可能であると考えられる。

### 3. 既存の解析方法との性能比較

Cas9 の HNH ドメインの部分配列である 21 または 31 塩基の配列が挿入された疑似 WGS データセット (30×シークエンスカバレッジ) を用いて、k-mer 法とアセンブリ法との性能比較を行った。その結果、

- 短い配列 (21 塩基) の検出について、k-mer 法は正確かつ迅速に検出できたが、アセンブリ法ではこれを検出できなかった。
- アセンブリ法は残存配列の有無と内容そしてそのゲノム上の位置を明らかにできるが、k-mer 法のみではゲノム上の位置を特定することはできない。
- アセンブリ法はゲノム編集作物の WGS データのみで解析できるが、k-mer 法はバックグラウンドとなった野生型作物の WGS データも必要となる。よってシークエンスのコストが 2 倍になる。

以上の留意点を踏まえ、ゲノム中に残存した外来遺伝子配列の検出にあたっては、まず k-mer 法を実施し、検出された場合にはアセンブリ法を実施して具体的な配列やゲノム上の位置を特定する、こうしたスキームで解析することが望ましいと考えられた。

### 網羅的代謝物解析分野

本研究で開発したツールは、質量分析データから取得したスペクトル類似性を活用して、標準物質・標準試料由来のリファレンススペクトルライ

ブラリと試料の分析データの未知化合物をスペクトル類似度に基づいて多層レイヤー上にネットワーク構造を生成する。各レイヤーは化合物クラスや毒性情報など情報を反映し、膨大で複雑なノンターゲット質量分析データの直感的な可視化と化合物構造情報の効率的な取得を可能にした。(図 2) このツールは比較定量データの可視化、未知化合物周辺のサブネットワークの抽出など、多様な機能を備えており、Python、Django、Plotly を組み合わせた実現したユーザーフレンドリーなインターフェースを通じて提供されている。Docker を利用することで、異なる環境下でも容易にデプロイメントと使用が可能となり、広範な研究者がアクセスしやすいように設計された (図 3)

本研究によって開発された解析ツールは、ゲノム編集作物を含む様々な食品の安全性検証に有効な手段を提供する。このツールによる分析は、従来のターゲットアプローチでは捉えられない、食品中の未知化合物や代謝変化の詳細な解析を可能にし、食品の安全性評価における新たな視点を提供する

このツールをゲノム編集トマトと非ゲノム編集トマトとの比較解析に適用し、多数の代謝物の変動と共に、優位に量的に変動している未知化合物の化合物クラスの推定および分布の可視化を行うことに成功した。さらにエンリッチメント解析を併用し関与する代謝パスウェイの解析を組み合わせることで、これらの化合物が生じる遺伝学的・生化学的背景の解明につながると期待された。

### トランスクリプトーム分野

(1) マッピング及び遺伝子アノテーション  
6 つすべてのサンプルにおいて、input reads に対し、ペアドリッド (ペアでマッピングされたリード) の合計はそれぞれ 4,000 万リード以上 (99% 以上)、また MAPQ が 10 以上のリードについても 4,000 万リード以上 (88% 以上) だった。これ

らのリードに対して、EnsemblID を用いて 35,825 遺伝子がアノテーションされた。

## (2) 遺伝子発現解析の結果

リードカウント値を用いた iDEP による解析

遺伝子発現量のリードカウントデータを iDEP に入力、最小 CPM (count per million) 値を 0.5 に設定したところ、EnsemblID でアノテーションされた 35,825 遺伝子のうち、17,597 遺伝子がパスした。この 17,597 遺伝子に対して、発現量変動解析を行った。方法は DESeq2 を選択し、FDR cutoff: 0.1、fold-change (GABA/WT):  $\geq 2$  で発現変動遺伝子の絞り込みを行ったところ、3,210 遺伝子で発現量が上昇 (Upregulated)、3,525 遺伝子で減少 (Downregulated) していた。さらに抽出された遺伝子を用いて、GO エンリッチメント解析を行ったところ、Up-regulated genes では ribonucleoprotein の生合成や chitin の活性に関わる GO、一方 down-regulated genes では Carbohydrate metabolic/catabolic process, plastid, chloroplast に関わる GO などが抽出された。

## TPM 値を用いた t-検定

マッピングの結果から算出された TPM 値に対して t-検定を行い、WT トマトと GABA トマトで発現量が有意に異なる遺伝子を抽出しリスト化した。EnsemblID でアノテーションされた 35,825 遺伝子のうち、研究方法に示した基準で、発現量が有意に上昇していた遺伝子 (up-regulated genes) は 932、減少していた遺伝子 (down-regulated genes) は 4,154、合計 5,086 遺伝子が抽出された。

## gProfiler による GO 解析

上記で抽出されたそれぞれの遺伝子リストを用いて、GO 解析を行ったところ、up-regulated genes では、chitin や cell wall の代謝プロセス、ribonucleoprotein の生合成など、down-regulated genes では、catalytic activity, small molecule

metabolic process, plastid に関係する GO が抽出された。これらの結果は、iDEP の結果と傾向が一致した。

## GAD 遺伝子の発現量変化

GABA トマトでは、グルタミン酸脱炭酸酵素 (glutamate decarboxylase; GAD) をコードする遺伝子の一つ、SIGAD3 が改変されている。当該酵素は、グルタミン酸のカルボキシル基を除去することで GABA を合成する。トマトは 5 つの GAD 遺伝子を有しているが、GABA 合成に寄与するのは GAD2 と SIGAD3 であり、特に SIGAD3 が主要な働きをすることがわかっている (参考文献 6,7)。GABA トマトは、CRISPR/Cas9 による変異導入 (一塩基挿入) により、SIGAD3 の C 末端の自己阻害領域が除去されており、その結果 SIGAD3 の活性が上昇することで、GABA の蓄積量が增大する。

SIGAD3 を含む 5 つの GAD のゲノム編集前後の遺伝子発現量の変化について、SIGAD3 及び GAD2, GAD4 の遺伝子発現量に有意な差は見られなかった。一方 GAD の発現量は 1/3 以下 (logFC: -1.8) を示し、有意に減少が認められた。SIGAD3 の活性の上昇により間接的に影響した可能性はあるが、関連は不明である。

## その他の GABA 代謝系に関わる遺伝子の発現量変化

GABA はグルタミン酸を基質として GAD により合成された後、GABA アミノ基転移酵素 (GABA-T) により Succinic semialdehyde (SSA) へと代謝される。SSA は SSA 脱水素酵素 (SSADH) により Succinate に代謝された後、TCA 回路へ流入する (GABA shunt)。これらの GABA 代謝経路に関わる可能性のある分子について、遺伝子の発現量を比較した。その結果、WT と比較して GABA トマトでは、GABA-TP1 の発現量が有意に増加 (logFC: 1.7)、GABA-TP3 と SSADH の発現量が

有意に減少 (logFC:-4.9,-1.3) していた ( $p < 0.05$ )。これらに関しても、SIGAD3 の活性の上昇により遺伝子発現が間接的に影響した可能性はあるが、関連は不明である。

WT と GABA トマトでは、有意に遺伝子発現のパターンの違いが見られたが、その原因として、ゲノム編集による影響が考えられるが、トマトの生育環境、収穫時期、生果の保存状態の違いなどその他の影響についても可能性は無視できない。特に当初に想定した中で、オンターゲット変異による代謝経路の変化に伴う間接的な遺伝子発現の変化、オフターゲット予測から漏れたオフターゲット変異による遺伝子発現の変化を明確にするためには、今回得られた遺伝子発現の変化のデータを、遺伝子配列や発現制御領域などゲノムの変化に落とし込む必要がある。3つめに想定した、Cas9 の持続的な作用による遺伝子発現の変化に関して、特に DNA 損傷シグナルに関連する遺伝子は抽出されなかった。

#### 参考文献：

- 1) ゲノム編集技術応用食品及び添加物の食品衛生上の取扱要領に基づき届出された食品及び添加物一覧 ([https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou\\_iryuu/shokuhin/bio/genomed/newpage\\_00010.html](https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryuu/shokuhin/bio/genomed/newpage_00010.html))
- 2) グルタミン酸脱炭酸酵素遺伝子の一部を改変し GABA 含有量を高めたトマトに関する届出情報 (<https://www.mhlw.go.jp/content/11120000/000828873.pdf>)
- 3) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Picelli et al., Nat Methods, 10 1096-1098, 2013  
doi:10.1038/nmeth.2639.
- 4) iDEP: an integrated web application for differential expression and pathway analysis of

RNA-Seq data. Ge SX et al., BMC Bioinformatics, 19:534, 2018

doi: 10.1186/s12859-018-2486-6

5) g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update) Nucleic Acids Research 2019; doi:10.1093/nar/gkz369

6) Tomato Glutamate Decarboxylase Genes SIGAD2 and SIGAD3 Play Key Roles in Regulating  $\gamma$ -Aminobutyric Acid Levels in Tomato (*Solanum lycopersicum*). Takayama et al., Plant Cell Physiol. 56:1533-45, 2015 doi: 10.1093/pcp/pcv075

7) Activating glutamate decarboxylase activity by removing the autoinhibitory domain leads to hyper  $\gamma$ -aminobutyric acid (GABA) accumulation in tomato fruit. Takayama et al., Plant Cell Rep. 36:103-116, 2017 doi: 10.1007/s00299-016-2061-4

#### アレルギー予測とデータベース分野

(1) アレルギー性予測手法 (allerStat) の機能拡充及び性能検証  
allerStat について、線形 SVM に加えて新たに非線形 SVM、ロジスティック回帰、LightGBM、LSTM、proteinBERT を分類器として追加実装し、それぞれの予測性について比較検討した。20 品目の食品のアレルギー及び非アレルギーデータを用いて Leave-Category-Out Cross-Validation を行い、予測性の指標となる ROC (Receiver Operating Characteristic) 曲線の AUC (Area Under Curve) を算出した。各分類器における AUC を図 1 に示す。ロジスティック回帰及び LightGBM では線形 SVM とほぼ同程度の AUC が得られた。非線形 SVM では線形 SVM と比較して AUC が小さい品目が多く、また品目ごとの AUC の変動が大きかった。LSTM では線形 SVM とほぼ同程度の AUC が得られた。一方 proteinBERT では線形 SVM と

比較して AUC が大きい品目が多く、品目ごとの AUC の変動が小さかった。

allerStat の予測性の検証においては、AlgPred2 (固定長 k の連続したアミノ酸配列解析(k=1, 2))、Allerdicator(固定長 k の連続したアミノ酸配列解析(k=6))、Allertop (アミノ酸の物理化学的性質に基づくアプローチ)、及び MEME (Multiple Expectation maximizations for Motif Elicitation、期待値最大化を利用した配列パターンマイニング法) の 5 種の予測法を比較対照とし、Leave-Category-Out Cross-Validation における AUC、AUC-10%、F1 スコア、MCC の 4 種を評価指標として、性能検証を行った。結果を表 1 に示す。4 種の評価指標の全てにおいて、allerStat で最も大きな値となり、allerStat が他の予測法と比較して高い予測性能を有することが示された。

## (2) ADFS エピトープ情報の追加

令和 2 年 6 月から令和 5 年 5 月までの 3 年間で、キーワード検索により抽出された論文は 74 報であった。その中からエピトープ情報が記載されていると思われる 47 報を選択し、ピアレビューを行った。その結果、28 報の論文から 42 種のアレルゲンについて、総数 166 のエピトープ情報を新たに追加した。

上記のアレルゲン及びエピトープ情報更新作業により、最終的に、ADFS のアレルゲン及びイソアレルゲンのアミノ酸配列情報は 2,408、エピトープ既知のアレルゲン数は 286、構造既知のアレルゲン数は 194、糖鎖付加アレルゲン数は 127 となった。

(3) FAO/WHO 法アレルゲン性予測ツール改修  
FAO/WHO のアレルゲン性予測法は、1)クエリタンパク質を N 末端側から 80 残基のウィンドウで順に区切ってゆき (sliding window 方式)、FASTA アラインメントにて既知アレルゲンと 35% 以上のアミノ酸が一致する場合、あるいは、2)クエリタ

ンパク質の連続する 6-8 アミノ酸が既知アレルゲンと完全に一致する場合にアレルゲン性が疑われる、とするものである。このうち 1)の方法は、ウィンドウ単位に細分化された配列を大量に処理する必要があり計算速度の遅延を招くと予想されたことから、ADFS ではこれまで改変法を用いてきた。改変法では、クエリタンパク質全長に対して FASTA アラインメントにより既知アレルゲンとの相同性比較を行い、80 残基以上の領域でオーバーラップが認められ、かつその 35% 以上のアミノ酸が一致する場合を陽性と判定する。しかし、オーバーラップ長全体では一致率 < 35% であっても、80 残基 sliding window では一致率 > 35% となる場合が見られることが分かった。そこで、現在の ADFS サーバは十分なスペックを有していることも考慮し、FAO/WHO 法の 1)を本来の方法に改修し公開した。この改修により、ADFS の FAO/WHO 法は、80 残基 sliding window 毎のより詳細な解析情報を得られるアレルゲン性予測ツールとなった。

(4) AI のリスク評価分野への応用に関する調査  
各分野において AI 技術の活用が進められる中で、特に医療分野においては、画像解析やプログラム医療機器 (Software as Medical Device: SaMD) 等で AI の導入が進められている。また、様々な疾病のリスク予測への AI 活用事例の論文報告があった。食品 (農業) 分野における AI の活用は医療分野ほどには進んでいないが、これまでに、土壌特性や気象パターンの予測、作物収量予測、スマート灌漑、インテリジェント収穫、需要・購買行動予測等に関する報告があった。また、家畜の病気や化学物質汚染に関するリスク予測等の論文報告があった。一方、AI 活用における課題としては、データセットの入手・内容・質、サンプルサイズ、モデルの性能評価・実用化・説明可能性等が挙げられていた。

国際機関や各国規制当局の取り組みやリスク評

価への活用について、医療・健康分野に関しては、WHO では、デジタルヘルスとして AI 技術の活用について、ガイダンス文書“Ethics and governance of artificial intelligence for health”作成等、多くの検討が進められている。ICH では医薬品承認申請におけるデジタル関連技術活用等が進められている。ICMRA では、Informal Innovation Network が AI に関するホライズン・スキニングとして、2 件のケーススタディ（中枢神経系アプリ、ファーマコビジランスにおけるシグナル管理）を実施し、その結果を踏まえて提言を行っている。IMDRF では、AI 関連の医療機器に関するワーキンググループが 2013 年に結成され、イノベーションを支援するガイダンス作成等を目指して活動が行われている。FDA では、デジタルヘルスイノベーション行動計画策定、デジタルヘルスソフトウェア事前認証プログラム（Pre-Cert プログラム）の試験運用等が行われている。EMA では、AI と新しいデジタル技術を活用した医薬品規制プロセスの効率化等が進められている。MHRA では、ソフトウェアと AI に対する規制要件を明確にした医療機器規制の改革プログラムを 2021 年 9 月に公表している。食品・農業分野に関して、FAO では、リモートセンシングと GEO-AI を用いた農作物季節学と農事暦の生成や水ストレス（乾燥期や干ばつ）の検出等、AI 技術を活用した多くの取り組みが進められている。Codex では、AI を食品不正に対抗する革新的な手段の 1 つであるにとらえ、2022 年 9 月に国際会議を開催している。USDA では、AI の使用事例の目録を作成しており、2022 年 12 月時点で 26 件が登録されている。EFSA では、「リスク評価におけるエビデンス管理のための AI 関連活動のためのロードマップ」プロジェクトを発足させ、2027 年までにリスク評価プロセスの信頼性を高めることを目標にしたロードマップを作成している。FSA では、2016 年にデータ戦略アクションプランを公表し、AI を地方自治体による食品施設衛生検査の効率化のスキーム等に活用している。

一方、課題としては、デジタルデバイス（情報格差）、法令や規制体制の整備、患者や公衆の安全確保、透明性・公正性の確保、データ利用に関する責任等が挙げられている。

AI の活用状況の調査では、現状では医療分野への活用や規制の取り組みが多いこと、食品・農業分野やリスク評価分野への活用は今後進んでいくと考えられることが示された。また AI 導入の問題点としては、各分野に共通のものとして、データセットの入手・内容・質、サンプルサイズ、モデルの性能評価・実用化・説明可能性等が挙げられる他、規制当局側の問題点としては、法令や規制体制の整備、患者や公衆の安全確保、透明性・公正性の確保、デジタルデバイスへの対応等が挙げられる。

## アレルゲン予測（HLA 結合性予測）

### （1）公的データベース

UniProt/Swiss-ProtKB からヒト MHC クラス II $\alpha$  鎖（計 6 配列：DRA、DQA1、DQA2、DPA1、DMA、DOA）、 $\beta$  鎖（計 9 配列：DRB1、DRB3、DRB4、DRB5、DQB1、DQB2、DPB1、DMB、DOB）を抽出し、構造情報（PDB： $\alpha$  鎖=216、 $\beta$  鎖=193）とを関連付けた。さらに IPD-IMGT/HLA からアレル遺伝子由来タンパク質配列情報合計 8,331 配列をダウンロードした。

### （2）先行研究予測法

提案されている 11 種類の MHC クラス II-ペプチド結合予測法に関して、採用しているアルゴリズムを調査した。これらのうち 5 種類の予測法が開発の際に使用した学習用データセットは、共通のリソースを使用し、データセットの継承/統合により量と質の確保に努めていた。MixMHC2pred では著者らが LC-MS/MS を用いた実験によって新規ヒト MHA クラス II 結合ペプチド（77,189 リガンドペプチド）を同定し、これを評価用データセットとして使用していた。

学習用データとして利用するペプチドデータについて、T 細胞受容体が結合し T 細胞応答が確認できているかを IEDB の登録情報を基に精査した。

### (3) 抗原ペプチドと MHC クラス II 分子間相互作用解析

IMGT のウェブサイトにてペプチド-MHC クラス II 分子-T 細胞受容体の複合体を形成している立体構造データを検索した結果、29 の複合体が得られた。

### (4) 構築したデータセット

LOAO 交差検定テストに使用する「IEDB2022 データセット」および「PDB-HB データセット」を構築した。

### (5) LOAO 交差検証テスト結果

One-hot encoding と BLOSUM62 を特徴量に用いているオリジナルの予測法 DeepSeqPanII よりも、AAindex を特徴量に加えた場合が全体的に高い AUC 値を示した。特に、AAindex (次元圧縮) が 54 アレル中 50 アレルで DeepSeqPanII よりも高く、2 アレルは同じ AUC 値であった。AAindex (次元圧縮) のほうが AAindex (非冗長 62 インデックス+次元圧縮) よりも、2 アレルを除いて、優れていることが分かった。

「IEDB2022 データセット」は「BD2016 データセット」と同様に AAindex (次元圧縮) は DeepSeqPanII よりも AUC 値が高い傾向にあり (15/19 アレル)、AAindex (非冗長 62 インデックス+次元圧縮) よりも全アレルで優れていることが分かった。AAindex を次元圧縮せずに 566 インデックスを特徴量に加えても AAindex (次元圧縮) の AU 値を全アレルで超えられていなかった。

PDB-HB データセット」では AAindex (次元圧縮) + PSSM による AUC 値が 8 アレル中 5 アレルにおいて最も高い AUC 値を示した。1 アレル (DRA\*01:01-DRB1\*01:01) のみ、DeepSeqPanII のデフォルトと同値であるが、それ以外の 7 アレルでは顕著な差 (AUC :  $6.9 \pm 3.2$ ) があつた。この特徴量の単独と比較すると、組み合わせによる効果が 5/8 アレルで有意であることが分かった。アミノ酸残基間ポテンシャルを加えた結果は、AAindex (次元圧縮) + PSSM よりも高い AUC が得られた一方で、1 アレル (DQA1\*03:01-DQB1\*03:02) では他の特徴量に比べて顕著に低い値となった。

### (6) 新規アレルゲン性評価手法開発の基盤研究と AI のリスク評価への応用

次元圧縮した AAindex を特徴量に加えることで改良した分子間相互作用予測法を、新たなアレルゲン性予測法に活用するため、既存予測手法の一つである NetAllergen 中の分子間相互作用予測部分を改良した相互作用予測法で置き換え、予測性能に及ぼす影響を、ベンチマークデータセットを用いて評価し、オリジナルと同等の結果が得られることが分かった。

以上、先行研究の調査、データセット整備などの基盤構築を経て、深層学習を用いた DeepSeqPanII をベースモデルに、複数の特徴量を組み合わせてトレーニングさせた。交差検定による予測性能評価では、物理化学的インデックスを用いた場合、AUC 値で平均 2.1% 向上した。ペプチド結合残基位置が明らかな立体構造情報データセットに対する性能評価結果からは、特徴量に PSSM と残基間ポテンシャルインデックスを組み合わせることによって平均 6.9% の AUC 値を向上できた。予測手法 NetAllergen 中の分子間相互作用予測部分を改良した相互作用予測法で置き換え、予測性能に及ぼす影響を、ベンチマークデータセットを用いて評価し、オリジナルと同等の結果が得られることが分かった。

## D. 研究発表・業績

### リスクコミュニケーション分野

論文 (解説、総説、書籍を含む) ○は本研究と関係が深い

1. ○Shineha R., Takeda K.F., Yamaguchi Y., Koizumi N. (2024) A comparative analysis of attitudes towards genome-edited food among Japanese public and scientific community, PLOS ONE (in press).
2. Ryohei Yamamoto, Seigo Higuchi, Yuji Iwata, Satomi Takeda, Nozomu Koizumi, Kei-ichiro Mishiba (2024) High  $\beta$ -carotene accumulation in transgenic eggplant fruits grown under artificial light.
3. ○小泉望 ゲノム編集食品をどう伝えるか:生活協同組合研究. 577, 34-41 (2024)
4. ○Takeda, K.F., Yazawa, A., Yamaguchi, Y., Koizumi, N., & Shineha, R. (2023) Comparison of

- public attitudes toward five alternative proteins in Japan. *Food Quality and Preference*, 105, 104787
5. ○小泉望、四方雅人 (2023) ゲノム編集食品に関する取扱いルールを経緯とこれから:ゲノム編集技術を応用した製品開発とその実用化. p565-572 技術情報協会
  6. ○小泉 望 (2023) リスクコミュニケーションのために求められること:ゲノム編集技術 ~ 実験上のポイント/産業利用に向けた研究開発動向と安全性周知 p295-p301 情報機構
  7. ○小泉 望、山口 夕、標葉隆馬 (2022) OECD 加盟国におけるゲノム編集作物に関するパブリック・エンゲージメント事例集、科学コミュニケーション研究所
  8. ○小泉望、四方雅人 (2022) ゲノム編集食品の取り扱いに関するルール 化学と生物 60, 150-153
  9. Sho Takeda, Taisuke Togawa, Kei-ichiro Mishiba, Katsuyuki T. Yamato, Yuji Iwata, Nozomu Koizumi (2022) IRE1-mediated cytoplasmic splicing and regulated IRE1-dependent decay of mRNA in the liverwort *Marchantia polymorpha*. 39, 303-310
  10. Masato Nakamura, Mamoru Nozaki, Yuji Iwata, Nozomu Koizumi, Yasushi Sato (2022) THESEUS1 is involved in tunicamycin-induced root growth inhibition, ectopic lignin deposition, and cell wall damage-induced unfolded protein response. *Plant Biotechnology*, 39, 129-138
  11. Rikako Hirata, Tomoya Makabe, Kei-ichiro Mishiba, Nozomu Koizumi, Samir M. Hamdan, Yuji Iwata (2022) Unpaired nucleotides on the stem of microRNA precursor are important for precise cleavage by Dicer-Like1 in *Arabidopsis*. *Genes to Cells* (印刷中) doi: 10.1111/gtc.12927
  12. ○小泉望 (2021) 遺伝子組換え食品とゲノム編集食品 食の安全安心通信 41、1
1. 標葉隆馬、武田浩平、小泉望「代替タンパク質をめぐる ELSI」、社会科学技術論学会 2023年12月9日、大阪大学(豊中市)
  2. 武田浩平、小泉望、標葉隆馬、「5種類の代替タンパク質に関する一般市民の態度」、社会科学技術論学会 2023年12月10日、大阪大学(豊中市)
  3. 小泉望「ゲノム編集食品のリスクコミュニケーション」(招待講演)、リスク学会 2022年11月12日 ハイブリッド
  4. 小泉望 「ゲノム編集食品のリスクコミュニケーション」(招待講演) 日本食品衛生学会 2022年12月2日 立命館大学(草津市)
  5. 2022年小泉望「ゲノム編集食品に関するコミュニケーションの振り返りと今後」12月4日 日本サイエンスコミュニケーション協会

### ゲノム解析分野

#### 論文発表

1. Soga K, Taguchi C, Sugino M, Egi T, Narushima J, Yoshida S, Takabatake R, Kondo K, Shibata N: Investigation of genetically modified maize imported into Japan in 2021/2022 and the applicability of Japanese official methods. *Food Hyg. Saf. Sci.* 2023; 64: 218-225
2. Takabatake R, Egi T, Soga K, Narushima J, Yoshida S, Shibata N, Nakamura K, Kondo K, Kishine M, Mano J, Kitta K: Development and interlaboratory validation of a novel reproducible qualitative method for GM soybeans using comparative Cq-based analysis for the revised non-GMO labeling system in Japan. *Anal. Chem.* 2022; 94: 13447-13454
3. Soga K, Nakamura K, Egi T, Narushima J, Yoshida S, Kishine M, Mano J, Kitta K, Takabatake R, Shibata N, Kondo K: Development and validation of a new robust detection method for low-content DNA using  $\Delta\Delta Cq$ -based real-time PCR with

学会発表

optimized standard plasmids as a control sample. Anal. Chem. 2022; 94:14475-14483.

4. Shibata N, Soga K, Sugino M, Nakamura K, Narushima J, Yoshiba S, Egi T, Takabatake R, Kondo K: Evaluation of conversion factor for rapid quantification of authorized genetically modified maize and soybean in Japan. BPB Reports. 2022; 5:115-120.
5. Narushima J, Kimata S, Shiwa Y, Gondo T, Akimoto S, Soga K, Yoshiba S, Nakamura K, Shibata N, Kondo K: Unbiased prediction of off-target sites in genome-edited rice using SITE-Seq analysis on a web-based platform. Genes to Cells. 2022; 27: 706-718.

#### 学会発表

1. 成島純平、杉野御祐、曾我慶介、吉場聡子、近藤一成、柴田識人:ゲノム編集イネにおける *in vitro* オフターゲット予測法 SITE-Seq を用いたオフターゲット予測性能の評価、日本ゲノム編集学会 第8回大会、東京、2023年6月6日-8日
2. 成島純平、吉場聡子、細川葵、曾我慶介、杉野御祐、田口千恵、安達玲子、近藤一成、柴田識人: Cas9 targeted long-read sequencing による食品中の外来性遺伝子配列の同定、第46回日本分子生物学会年会、兵庫、2023年12月6日-8日
3. 曾我慶介、江木智宏、成島純平、吉場聡子、真野潤一、高畠令王奈、橘田和美、中村公亮、柴田識人、近藤一成: 新遺伝子組み換え表示制度に向けた試験法開発 ~ 「遺伝子組み換えでない」表示の今後~、第8回次世代を担う若手のためのレギュラトリーサイエンスフォーラム、東京、2022年8月26日
4. 近藤一成、曾我慶介、成島純平、吉場聡子、柴田識人、田口千恵、坂田こずえ、加藤怜子: ゲノム編集によって発生する意図しない変異はどこから来るのか、NGS EXPO 2022、大阪、2022年10月18日
5. 成島純平、木俣真弥、志波優、権藤崇裕、秋元智、曾我慶介、吉場聡子、中村公亮、柴田識人、近藤一成: ゲノム編集におけるオフターゲット予測法 SITE-Seq の Galaxy ベース新規解析パイプラインの開発とその作物への応用、NGS EXPO 2022、大阪、2022年10月18日
6. 曾我慶介、吉田光範、成島純平、吉場聡子、柴田識人、近藤一成: ナノポアリードで構築したスギヒラタケアセンブリ結果からみたアセンブラーツールの選択について、NGS EXPO 2022、大阪、2022年10月19日
7. 吉場聡子、成島純平、曾我慶介、杉野御祐、柴田識人、近藤一成: 遺伝子組換え食品の同定に資する新たなゲノム解析技術の検討、第59回全国衛生化学技術協議会年会、東京、2022年10月31日-11月1日
8. 柴田識人、成島純平、曾我慶介、吉場聡子、近藤一成: ゲノム編集食品における外来性遺伝子の残存を評価する全ゲノムシーケンスデータ解析の標準化に向けた取り組み、第59回全国衛生化学技術協議会年会、東京、2022年10月31日-11月1日
9. 成島純平、杉野御祐、曾我慶介、吉場聡子、柴田識人、近藤一成: NGS を用いた網羅的なオフターゲット変異候補部位予測法の高 GABA 産生ゲノム編集トマトへの適用と検証、第59回全国衛生化学技術協議会年会、東京、2022年10月31日-11月1日
10. 柴田識人、成島純平、曾我慶介、吉場聡子、近藤一成: ゲノム編集食品における外来性遺伝子の安全性評価における全ゲノムシーケンスデータを用いた解析の標準化に向けた課題、第118回日本食品衛生学会学術講演会、長崎、2022年11月10日-11日
11. Soga K, Yoshiba S, Narushima J, Shibata N, Kondo K: Genome analysis of deadly

- poisonous mushroom *Amanita virosa* using nanopore sequencing technology, Cell Bio 2022-An ASCB/EMBO Meeting, Washington DC, USA, 2022/12/5
12. Kondo K, Fukuda N, Soga K, Yoshiba S, Narushima J, Shibata N, Taguchi C, Sakata K, Kato R: G2/M synchronization with CDK1 inhibitor suppresses genome-wide mutations and genome rearrangements during genome editing, Cell Bio 2022-An ASCB/EMBO Meeting, Washington DC, USA, 2022/12/5-6
  13. 高島令王奈、大西真理、峯岸恭孝、布藤聡、曾我慶介、柴田識人、中村公亮、近藤一成、真野潤一、橘田和美：アグロバクテリウム法によって作出された遺伝子組換え植物中に存在する超短配列 25 bp の検出法の開発、第 64 回日本植物生理学会年会、仙台、2023 年 3 月 15 日-17 日
  14. 曾我慶介、成島純平、吉場聡子、柴田識人、近藤一成：猛毒キノコドクツルタケのドラフトゲノムを用いた新規毒性ペプチドの探索、日本薬学会第 143 年会、札幌、2023 年 3 月 27 日
  15. 曾我慶介、中村公亮、成島純平、吉場聡子、木俣真弥、江木智宏、岸根雅宏、真野潤一、橘田和美、高島令王奈、柴田識人、近藤一成：改正食品表示基準に対応した遺伝子組換えとうもろこし混入の判定に係る定性 PCR 検査法の開発、日本食品衛生学会第 117 会学術講演会、東京、2021 年 10 月 26 日
  16. 高島令王奈、江木智宏、曾我慶介、峯岸恭孝、成島純平、吉場聡子、柴田識人、中村公亮、近藤一成、岸根雅宏、真野潤一、橘田和美：改正食品表示基準に対応した遺伝子組換えダイズ混入の判定に係る定性 PCR 検査法の開発、日本食品衛生学会第 117 会学術講演会、東京、2021 年 10 月 26 日
  17. 吉場聡子、成島純平、曾我慶介、柴田識人、近藤一成：安全性未承認の遺伝子組換えナタネの試験法開発、第 58 回全国衛生化学技術協議会年会、名古屋、2021 年 11 月 25-26 日
  18. 柴田識人、曾我慶介、成島純平、吉場聡子、江木智宏、高島令王奈、近藤一成：遺伝子組換えダイズ・トウモロコシ定量のための内標比の算出、第 58 回全国衛生化学技術協議会年会、名古屋、2021 年 11 月 25-26 日
  19. 曾我慶介、成島純平、吉場聡子、江木智宏、真野潤一、高島令王奈、橘田和美、中村公亮、柴田識人、近藤一成：「遺伝子組換え出ない」表示確認に係る新定性検査法の試験室間共同試験による妥当性評価、第 58 回全国衛生化学技術協議会年会、名古屋、2021 年 11 月 25-26 日
  20. 成島純平、木俣真弥、志波優、権藤崇裕、秋元智、曾我慶介、吉場聡子、中村公亮、柴田識人、近藤一成：ゲノム編集作物におけるオフターゲット予測法の検討、第 44 回日本分子生物学会年会、横浜、2021 年 12 月 2 日
  21. 曾我慶介、吉田光範、吉場聡子、成島純平、柴田識人、近藤一成：ナノポアシーケンシング技術を用いたスギヒラタケの全ゲノムアセンブリの検討、第 44 回日本分子生物学会年会、横浜、2021 年 12 月 3 日
  22. 近藤一成、成島純平、曾我慶介、吉場聡子、柴田識人、坂田こずえ、田口千恵、加藤怜子：ヒト細胞 TK6 を用いた CRISPR/Cas による構造変異(SV)解析の検討、第 44 回日本分子生物学会年会、横浜、2021 年 12 月 3 日

## 網羅的代謝物解析分野

### 論文発表

- 解析・可視化ツールの投稿論文準備中
- 開発された解析ツールの Docker イメージ配布レポジトリ：<https://github.com/mass-spec-info/SNVis>

### 学会発表

1. 日本食品化学学会 第 27 回総会・学術大会 (2021.6.10)、オンライン開催 「質量分析イ

ンフォマティクスによる食品中の未知化合物の解析プラットフォーム」(早川英介, 渡邊寛, 近藤一成) (ポスター発表)

2. 日本食品衛生学会第 118 回学術講演会(2022年 11 月 10 日-11 日)出島メッセ長崎「質量分析インフォマティクスによる食品試料の迅速・ノンターゲットな未知化合物解析プラットフォーム」(早川英介, 渡邊寛, 近藤一成) (ポスター発表)
3. 日本食品衛生学会第 119 回学術講演会 (2023年 10 月 12 日-13 日) タワーホール船堀「未知化合物解析支援のための質量スペクトルネットワーク可視化ツール」( 早川英介、平田香南子、近藤一成、有田正規) (ポスター発表)

### アレルギー予測とデータベース分野

#### 論文発表

1. Goto K, Tamehiro N, Yoshida T, Hanada H, Sakuma T, Adachi R, Kondo K, Takeuchi I. AllerStat: Finding Statistically Significant Allergen-Specific Patterns in Protein Sequences by Machine Learning. J Biol Chem. 2023 Jun;299(6):104733.
2. Okazaki F, Momma K, Hirakawa Y, Kawai N, Yamaguchi-Murakami Y, Adachi R, Mori Y, Kondo Y, Narita H. Determination of severe peach allergens, gibberellin-regulated protein, and lipid transfer protein, using monoclonal antibodies. J Nutr Sci Vitaminol, 2022;68(3):221-227.

#### 学会発表

1. 安達玲子. 医療分野等における AI 技術の利用状況とリスク評価分野への適用性. 日本薬学会第 143 年会 (2023 年 3 月 26 日、札幌)
2. 爲廣紀正. 機械・深層学習を利用した新たなアレルギー性予測手法の開発. 日本薬学会第 143 年会 (2023 年 3 月 26 日、札幌)
3. 爲廣紀正. 新開発食品の安全性評価～アレルギー誘発性～. 2023 年 12 月 1 日、オンライン開催、令和 5 年度東京農業大学総合研究所研究会【食の安全と安心部会】第 6 回シンポジウム

### アレルギー予測 (HLA 結合性予測)

なし

### E. 健康危険情報

該当なし

### F. 知的財産権の出願・登録状況

1. 特許取得

該当なし

2. 実用新案登録

該当なし