

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

総合総括分担研究報告書
(H30-食品-一般-002)

研究代表者 近藤一成

令和3（2021）年 5月

目 次

I. 総合総括研究報告書

近藤 一成	3
-------	---

II. 総合分担研究報告書

1. ゲノム編集に関する情報収集解析、アレルゲンタンパク分解性検討

および人工ヌクレアーゼの特異性を調べる *in vitro* アッセイツールの開発

近藤 一成、中村 公亮	9
-------------	---

2. リスクコミュニケーションに関する研究

小泉 望	13
------	----

3. メタボロームインフォマティクスによる未知化合物推定

早川 英介	19
-------	----

4. 高精度アレルゲン性予測システムの構築に必要な情報の収集に関する研究

爲廣 紀正	23
-------	----

5. 人工知能を用いたアレルゲン性評価のためのアルゴリズム開発

竹内 一郎	31
-------	----

6. ゲノム編集生物作製における現象解析と規制の進め方

木下 政人	37
-------	----

7. ゲノム編集の特性、安全性に関する調査研究

吉場 聡子	55
-------	----

III. 研究成果の刊行に関する一覧表

	57
--	----

I. 総合総括研究報告書

II. 総合分担研究報告書

III. 研究成果の刊行に関する一覧表

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

総合総括分担研究報告書
(H30-食品-一般-002)

研究代表者 近藤一成

令和3（2021）年 5月

目 次

I. 総合総括研究報告書

近藤 一成	3
-------	---

II. 総合分担研究報告書

1. ゲノム編集に関する情報収集解析、アレルゲンタンパク分解性検討

および人工ヌクレアーゼの特異性を調べる *in vitro* アッセイツールの開発

近藤 一成、中村 公亮	9
-------------	---

2. リスクコミュニケーションに関する研究

小泉 望	13
------	----

3. メタボロームインフォマティクスによる未知化合物推定

早川 英介	19
-------	----

4. 高精度アレルゲン性予測システムの構築に必要な情報の収集に関する研究

爲廣 紀正	23
-------	----

5. 人工知能を用いたアレルゲン性評価のためのアルゴリズム開発

竹内 一郎	31
-------	----

6. ゲノム編集生物作製における現象解析と規制の進め方

木下 政人	37
-------	----

7. ゲノム編集の特性、安全性に関する調査研究

吉場 聡子	55
-------	----

III. 研究成果の刊行に関する一覧表

	57
--	----

I. 総合総括研究報告書

II. 総合分担研究報告書

III. 研究成果の刊行に関する一覧表

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	(国立医薬品食品衛生研究所)
研究分担者	木下 政人	(京都大学)
研究分担者	今村 知明	(奈良県立医科大学)
研究分担者	有田 正規	(理化学研究所)
研究分担者	小泉 望	(大阪府立大学)
研究分担者	竹内 一郎	(名古屋工業大学)
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	爲廣 紀正	(国立医薬品食品衛生研究所)
研究分担者	中村 公亮	(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)

研究要旨：

本研究では、種々の手法による遺伝子改変の影響、ゲノム編集作物の開発状況や規制状況の情報収集を行い施策に反映するとともに、安全性確認で必要な項目や問題点を明らかにした。また、ゲノム編集技術や合成生物学など新たなバイオテクノロジー技術を用いた新開発食品の安全性を確認するために必要な新たな手法の開発検討を行った。ゲノム編集ではオフターゲットが課題になっていることから、配列類似性によらないバイアスのないゲノム全体の DNA 切断部位を検出する方法の開発、非アレルゲンタンパクのアミノ酸情報も加味し、既知アレルゲンタンパクとの相同性に依存しない人工知能を用いた全く新たなタンパクアレルゲン性予測アルゴリズムの開発、新開発食品試料中に出現する未知成分の質量分析インフォマティクスを用いた同定手法の開発、の検討を行った。

ゲノム解析では、SITE-seq 法を出発点にしたオフターゲット検出法を *in vivo* において検証して、その有用性を確認できた。新規アレルゲン性予測では、アレルゲンタンパクにのみ出現するアミノ酸配列パターンを抽出、特に非アレルゲンタンパクデータセットの改良を行いながら検討して、アレルゲン性予測が既報のどの方法よりも精度が高いことが確認できた。EFSA においても検討されている、アレルゲン性とも関連するタンパクの分解性試験について、pH、酵素濃度について細かく設定して検討した。その結果、分解されやすいタンパクにおいてもペプシン濃度よりも pH 変化が分解性に大きく影響することが分かった。質量分析インフォマティクスでは、食用、毒性を有する試料を用いた解析、公共データベース、および標品測定からスペクトル情報を取得することでデータベース化するとともに、結果をネットワーク化して可視化するシステムを構築できた。さらに、ゲノム編集トラフグについて、テトロドトキシンの蓄積、分布変化、およびゲノム編集トラフグのラットへの毒性試験を実施した。その結果、毒の分布に変化はなく、また、ゲノム編集トラフグの毒性もないことが確認できた。一方で、強い毒性を持った生物については、検体数を増やした詳細な検討もまた必要であると考えられる。ゲノム編集食品に対する国民理解の促進のためにリスクコミュニケーション、パンフレット作製を行った。また、近年 RNA 編集に関する論文が増えていることからその動向調査を行った結果、現在の研究の中心は疾患に関わる研究であり植物における RNA 編集に関する応用例はほとんどなかった。

A. 研究の目的

ゲノム編集技術を応用した新たな食品（ゲノム編集食品）の研究開発が国内外で活発に行なわれている。ゲノム編集食品では、従来の遺伝子組換

え食品のような外来遺伝子を導入することはなく、生物自身が本来持っている内在性遺伝子の配列を数塩基欠失により機能欠失させることで新たな形質（もち性向上、筋肉量増加、GABA 量増

加など)を付与することが期待されている。別の生物種から外来遺伝子を導入することがないため、国民受容の改善の点でも大きく期待されている。もう一つの重要な技術、合成生物学を利用した物質生産も欧米を中心に活発に研究されている。酵母などの微生物に、新たな物質生産に必要な多数の遺伝子を導入することでその生物が元来合成できない化合物の生産が可能になっている。

ゲノム編集食品では、安全性評価において重要な点は、内在性遺伝子改変に伴う塩基配列変化(オンターゲット)とそれに起因する新たな毒性やアレルギー性を有するタンパクや有害物質の出現または増加であり、加えて意図しない変化(オフターゲット)も考慮すべき点である。一方で、合成生物学利用作物では、生合成経路に関わる多数の遺伝子を導入する。従来、安全性評価対象は導入した遺伝子群とその影響であるが、組換え範囲が大きいいため従来の遺伝子組換え前後の比較による実質的同等性の考え方が適用できないことも想定されることから、絶対的な評価手法の開発なども必要と考えられる。

従来の遺伝子組換え食品における安全性確認の基本的な考え方は、十分な食経験がある組換え前の生物(の品種)に対して、新たに追加した遺伝子に対する安全性評価を行い、組換え前と実質的に同等か、リスクの程度が同定かそれ以下であることの確認、いわゆる実質的同等性確認、である。そこでは、導入遺伝子に関する分子生物学的特性、ベクターなど外来遺伝子とその断片の有無と安全性、新たなアレルギー性タンパクの有無、主要成分の変化について確認される。しかし、ゲノム編集食品では、外来遺伝子とその断片がないと仮定すると、改変されるのは内在性遺伝子上における塩基の挿入・欠失であり、標的部位(オンターゲット)での変化が十分解析されていることが重要で、その上で潜在的なリスクの一つは意図しない改変であるオフターゲットの影響(ゲノム状の塩基配列の変化と代謝産物の変化など)である。オン・オフターゲット部位での変化によって生じるリスクは、新たな毒性・アレルギー性タンパクの生成である。ゲノム解析が進んだ現在においても、ゲノム配列のみから毒性タンパクやアレルギー性タンパクが生成しないことを明らかにするのは容易ではない。また、意図しない有害成分産生の可能性があったとしても、現在の質量分

析を用いた解析では未知ピークの同定や推定は困難な現状がある。さらに、タンパクアレルギー性の確認も、現在実行可能な*in silico*解析は既知のアレルギータンパク質との相同性比較のみであり、相同性がない新規アレルギー性タンパク質の予測や非天然型アミノ酸から構成されるタンパクのアレルギー性を予測することは極めて難しい。加えて、合成生物学のような、従来の安全性評価の考え方が適用できないことも考えられる。

このような状況を鑑みて、ゲノム編集・合成生物学を利用した食品の開発状況情報収集をもとにしたケーススタディーや開発者との連携で申請側の問題点を明らかにするとともに、上記のゲノム編集食品や合成生物学利用食品の安全性確認のために必要な評価手法の新たな開発が急務と考えられた。

本研究では、安全性確認のための新たな手法開発において、*in silico*解析では標的配列と類似した配列のオフターゲット検索しかできない点を克服すべく、全ゲノム解析をすることなく潜在的なDNA2本鎖切断部位を網羅的に検出する手法、意図しない新たな成分が産生した場合の質量分析インフォマティクスを用いた成分同定あるいは基本構造推定手法、人工知能を活用して相同性がないアレルギー性タンパクの予測や非天然型アミノ酸から構成されるタンパクのアレルギー性を予測する手法、の開発検討を行う。また、ゲノム編集技術を用いた動物、植物の研究開発動向、技術動向の調査を行い、注視すべき生物種や技術動向から課題点抽出をする。また、ゲノム編集食品を対象としてリスクコミュニケーションを適切に行うために、Webアンケート調査からの詳細な統計的解析、国内外におけるリスクコミュニケーション事例の包括的調査、多様なステークホルダーによる座談会に基づく冊子の作成を目的とした。

B. 研究の実施

(1) ゲノム編集技術を用いた動植物の研究開発動向調査およびアレルギー分解性の検討

本課題は、国立医薬品食品衛生研究所・近藤一成が分担して実施した。植物・動物を主な対象に、ゲノム編集技術(ZFN、TALEN、CRISPR/Cas)を用いた動物、植物の研究動向について、PubMedを中心にキーワードを組合せて検索を行い、整理

した。増加している生物種、技術の動向などを図式化した。アレルギー解析について、モデルタンパクとして代表的なアレルギータンパクとして、ピーナッツアレルギーや卵白アレルギーについて、pH や分解酵素の濃度を変化させてその分解性の変化と抗原性の保持状況について、非アレルギータンパクの変化とも比較しながら解析した。

人工ヌクレアーゼの特異性を調べる *in vitro* アッセイツールの開発

本課題は、国立医薬品食品衛生研究所・中村公介 (H30~R1) および近藤一成 (R2) が分担して実施した。細胞から抽出した DNA を用いて、潜在的な DNA 二本鎖切断部位を予測する手法として SITE-Seq 法を検討した。本方法で予測した、切断候補部位について、*in vivo* 条件で検証するためにイネカルスを用いて検討し、SITE-Seq 法で予測した候補部位は、実際の生物においても確認できるかどうか確かめた。本アプローチが、ゲノム編集食品の開発において、非意図的な切断の影響評価に役立つことが期待される。

(2) リスクコミュニケーションに関する研究

本課題は、奈良県立大学・今村知明、および、大阪府立大学・小泉望が分担して実施した。GM 食品のリスクコミュニケーションのキーファクターを盛り込んだ有効な説明手法を活用し、厚生労働省の GM に関するパンフレットについて提案した。また、NBT に関する正しい理解と判断を促すための新たなコンテンツについても検討した。

また、一般モニターと専門家を対象にした Web アンケートを実施してそれを解析した。また、リスクコミュニケーションの手法について検討する為、諸外国におけるリスクコミュニケーション事例 (38 事例) を調査行った。情報提供のための資料として、消費者、事業者、開発者、行政などの多様なステークホルダーが参加する座談会を元にした冊子の作成を行った。

(3) 質量分析インフォマティクスによる化合物同定

本課題は、理化学研究所・有田正規、および、沖縄科学技術大学院大学・早川英介が分担して実施した。実測スペクトルを収集し、そこに含まれる MS/MS フラグメント・ライブラリを構築した。

マススペクトルを入手できない物質について、フラグメンテーション予測によるフラグメント・ライブラリを構築するための検討を行った。これまで、遺伝子改変など何らかの改変によって生じた意図しない成分変化とその化合物同定は、特に複雑な食品抽出試料からは非常に困難であると考えられてきた。そのため、通常は限られた標品のある成分においてメタボローム解析を実施してその変化を解析している。本研究では、ゲノム編集操作で、想定しない成分の変化を捉え、その化合物の同定または推定を行い、毒性についても考察できるシステムの構築を検討した。

(4) アレルギーデータベース ADFS のアップデート、および、人工知能を用いたアレルギー性評価のためのアルゴリズム開発

本課題は、名古屋工業大学・竹内一郎および国立医薬品食品衛生研究所・爲廣紀正が協力連携分担して実施した。

アレルギー情報について、アレルギーデータベース (AllergenOnline) を参考にアレルギーデータベース (ADFS) を更新した。エピトープ情報について、NCBI PubMed に掲載された論文を査読して必要なエピトープ情報を含む場合はその情報を整理し、ADFS にデータを追加した。

新規アレルギー予測手法の検討のために、アレルギーデータベース COMPARE からアレルギータンパク情報を、非アレルギータンパク情報は、UniProt からアレルギー情報を除くことによって構築したものについて、再度精査を行い更新した。

得られた上記データセット (アレルギーデータと非アレルギーデータ) を用いてアレルギーのみ出現し、非アレルギーには一度も出現しないアミノ酸のモチーフ配列を抽出した。このモチーフデータを利用して、アレルギー性判定・予測システムを検討した。さまざまな数理技術、情報技術を活用することで高精度・高信頼性かつ汎用性のあるアレルギー性判定・予測システムを開発するために、データセットの組み合わせ等を試行錯誤しながら検討を行った。

(5) ゲノム編集生物作製における現象解析と規制の進め方

本課題は、京都大学・木下政人が分担して実施した。ゲノム編集技術により、ミオスタチン遺伝

子を破壊したマダイ、レプチン受容体遺伝子 (*lepr*)、メラノコルチン4型受容体遺伝子 (*mc4r*) を破壊したトラフグについて特性を解析した。ゲノム解析、アレルゲン性、メタボローム解析、ラット短期毒性試験を行って、総合的に解析を実施して、そのデータについて食品安全性の評価法への提言となることを念頭に考察した。

(6) ゲノム編集技術の特性、安全性について

PubMed (NCBI) を用いて、DNA 編集および RNA 編集に関するキーワードを用いて論文検索及び情報取得・解析を行った。RNA 編集技術に関する文献調査から、論文の PMID、Journal、Title、Doi、Abstract、Year、Month、Status、MeSH、Keyword の情報を csv 形式で取得した。さらに、ファイルの統合と重複除去を行った後 (1,753 報)、項目の整理及び技術に関する論文の抽出を行い、重要と思われる論文 18 報を詳細に調べた。RNA 編集技術の食用となりうる生物、特に植物への応用においてその実態と傾向を考察した。

(7) 人材育成 (統計学、バイオインフォマティクス、AI 分野)

分担研究者および協力研究者と共同で行うことで、インフォマティクス関連技術の取得に努めた。

C. 研究の成果

(1) ゲノム編集に関する情報収集解析、ケーススタディーおよびアレルゲン分解性の検討

ゲノム編集技術を中心に研究開発の文献を 2018 年~2020 年について調査した。動物、植物においてゲノム編集技術特に CRISPR/Cas を用いたもの中心であり、技術的には CRISPR/Cas が今後も中核技術でこれに様々な工夫を加えた手法が今後も中心となると考えられた。国内では研究開発が盛んな魚類では、ウナギ、ドジョウなど多様な生物に応用されているが論文報告件数は少ない。植物では、イネでの研究が中国を筆頭に活発で、トマト、コムギ・オオムギ、ダイズが多く研究されている。また、イチゴなど果樹への応用も進んでいる。形質は環境耐性から栄養改変まで多岐にわたる。モデル生物から実際の目的生物での研究が進んでいると考えられた。

アレルゲン分解性については、pH の少しの変化、例えば pH1.2 から pH3 へ変更しただけで完

全に分解していたものが、ほとんど分解されなくなるタンパク (アレルゲン) があることが確認できた。一方、消化酵素であるペプシンの濃度条件ではあまり分解性には影響を与えなかった。今回は、患者血清を用いた検討を十分行えなかったが、反応性が示されれば、これまでアレルゲン性評価のデータとして検討されてきたアレルゲンタンパク分解性試験のあり方を再検討する必要があるかもしれないことが示唆された。

ゲノム網羅的に DNA 二本鎖切断部位を検出する手法とツールの開発検討

イネをモデル生物種とし、論文に報告されている遺伝子を標的として、ゲノム編集を行った場合の潜在的 DNA 二本鎖切断部位 (オフターゲット) 解析手法を検討した。

細胞から抽出精製した DNA を利用する *in vitro* 法である SITE-Seq 法を用いて、信頼性のある潜在的な切断部位候補を抽出した。その候補部位について、実際にイネを形質転換して個体を生育し、カルスから DNA を抽出精製してその候補部位が切断されているかを、NGS を用いたアンプリコンシーケンスで解析した結果、1 つのオフターゲットで切断が確認された。また、このオフターゲット部位は標的部位とは配列類似性が高くなく、殆どの *in silico* 解析法において検索できなかった。以上の結果から、SITE-Seq 法と *in vivo* 検証を組み合わせたオフターゲット検出法は、非常に有用な方法であることが示された。フレームシフトによりアミノ酸配列の変化が生じる。こうしてできる可能性があるタンパクのアレルゲン性を予測するため、アレルゲンデータベースを用いてアレルゲン予測を行った。その結果、今回ゲノム編集を行い、アンプリコンシーケンス解析で明らかになった配列においては、80-Mer Sliding Window FASTA Search および 8-mer FASTA Search のいずれでも既知アレルゲンとホモロジーのある配列は検出されなかった。オフターゲット検索からアレルゲン予測まで含めたスキームを確立することで、一つの安全性評価法として有効であると考えられる。

(2) リスクコミュニケーションに関する研究

リスクコミュニケーションのための資料として、厚労省で提供するパンフレット「新しいバイオテクノロジーで作られた食品について」を作成

した。また、ゲノム編集技術の理解促進のために、パンフレット「ゲノム編集技術応用食品を適切に理解するための6つのポイント」を作成した。

また、一般モニター4,000人、専門家398人を対象としたゲノム編集食品に対するWebアンケートについて、統計処理を含めて詳細に解析した（R2年度分担報告書、添付資料1）。現状では様子見的な態度を示す回答者が多いが、否定的な傾向にあるわけではない。科学リテラシーが高いと専門家信頼・社会受容ともにポジティブな態度が高まる。にリスクとベネフィットに関する情報共有への関心が強いこと、などが明らかになった。

また、海外におけるリスクコミュニケーション事例の包括的調査について、過去5年以内の国内外の遺伝子組換え食品あるいはゲノム編集食品に関するリスクコミュニケーション、サイエンスコミュニケーションの38実施例をWeb検索により抽出した（R2年度分担報告書、添付資料2）。さらに、消費者、事業者、開発者、行政などの多様なステークホルダーが参加する座談会を実施し、そこで出た意見・質問を参考に、頻出する疑問を整理した。

(3) 質量分析インフォマティクスによる化合物同定

本研究の目的はゲノム編集作物等で想定外の質・構造的変化を生じる化合物の迅速な検出と構造の推定である。従来の質量分析のデータ解析では困難だった未知化合物の検出と構造推定を、スペクトル類似性に着目した独自の手法で可能にするために、スペクトルライブラリの構築をまず行った。フラグメントスペクトルライブラリは、公共データベースMassBankに加え、天然化合物を多く含むデータベース群であるGNPSやRespectも追加して、特に植物二次代謝物に対応できるよう整備した。また、独自で500種類の低分子化合物を分析し、ライブラリに追加した。さらに代表的な食品・モデル植物等（ダイズ、トマト、ジャガイモなど）50種類以上から抽出した低分子化合物に関してもスペクトルライブラリに追加した。ゲノム編集ジャガイモに本手法を適用して手法の有用性について検討した。データ解析システムをWebブラウザベースの実行環境で動作させる目的でプログラミングなどの専門知識がない一般的な研究者でも容易にデータ解析とデータ可視化ができるように構築した。

(4) アレルゲンデータベース ADFS のアップデート、および新規タンパクアレルゲン性子測に必要な情報の作製

生化学部で管理運営するアレルゲンデータベース ADFS について、AllergenOnline のアレルゲン情報アップデートに伴って ADFS 登録アレルゲンのアップデート作業を定期的実施した。エビトープ情報は、論文レビューを行った。

次の(5)に用いるアレルゲンおよび非アレルゲン情報の構築について、アレルゲン情報は Comprehensive Protein Allergen Resource (COMPARE) のサイトから最新情報(2,248)に更新、非アレルゲンデータは、UniProt のサイトからどこまでデータを入れ込むかについて試行錯誤しながら、アレルゲン表示対象の食品10品目に相当するタンパクアミノ酸情報(17,372)とした。

また、ADFS サイトの脆弱性に関する対策を行った。

(5) 人工知能を用いたアレルゲン性評価のためのアルゴリズム開発

(4)で構築したデータセットを活用、精度を確認しながら再度データセットの見直しを行い最適化していった。このとき、False Negative を除くためにアレルゲン特異的パターンの定義を整理し、その解釈を明確化して、データ構築を行った。近年のAI研究では、データ駆動型AIシステムの説明性・信頼性が重要であることが要請されるが、本研究で作成したシステムは、統計的信頼性が担保されたパターンのみを用いており、これらの要請を満たしている。最終的に、これまでに論文報告されているシステムの Alledictor、MEME、Allertop と比較を行い、本研究で作成した方法の予測精度がよりよいことが確認できた。

(6) ゲノム編集生物作製における現象解析と規制の進め方

ゲノム編集技術により、ミオスタチン遺伝子を破壊したマダイ、レプチン受容体遺伝子 (*lepr*)、メラノコルチン4型受容体遺伝子 (*mc4r*) を破壊したトラフグについてゲノム解析、メタボローム、アレルゲン性に関する解析データを評価した。外来遺伝子の残存やオフターゲット切断、および、アレルゲン性がないことを確認した。

トラフグについて、フグ毒の体内分布と可食部

(骨格筋)の毒性について検討を行った結果、非ゲノム編集体とゲノム編集体では毒の分布に差はなく、可食部を摂取したラットにおける体重変化や生化学的数値に優位な差は認められなかったことから、明らかな毒性はないものと考えられた。

以上から、ゲノム編集操作したマダイやトラフグについて、食品としての安全性に明確な懸念点は存在しないものと考えられた。これまでに実施した解析で、安全性確認について必要な解析データは得られていると考えられるが、一方で、フグについてはその毒の毒性が極めて強いことから十分な評価が必要と考えられた。

(7) ゲノム編集技術の特性、安全性に関する研究

1) ゲノム編集技術に関する論文調査

ゲノム編集に関する文献 17,499 報 (2010.1.1-2020.10.8) 取得した後、クラスター分析により、ランダムに選んだ 4,000 の文献を 60 のクラスターに分類した。その結果、RNA 編集に関するキーワードを含むクラスターに多くの論文が含まれることが判った。RNA 編集に関するキーワードを含むクラスターの近くに mitochondria、plant、transcriptome、histone、chromatin などのキーワードを含むクラスターが見られた。また、plant(12)は、crop(2)、breeding(2)などと合わせて、ゲノム編集に関わる多くのクラスターに含まれており、広い分野に渡って研究されていることが推察された。

2) RNA 編集技術に関する論文調査

RNA 編集技術に関する文献 (2017.1.1-2020.10.13) 1,753 報を取得した。その中で、RNA 編集技術を用いた研究の数は予想より少なく、多くはゲノムの一塩基編集 (base editing) に関する研究であった。ゲノム編集技術に比べると RNA 編集技術はまだ開発途上であること、応用分野が限られることなどが考えられた。

3) 作物等への応用について

中心的な役割をする分子である ADAR は、それを持たない植物では作動しないので、RNA 編集を作物に利用するためには、外から酵素を導入する必要があり、さらに効果を持続させるために、酵素をゲノムに組み込む必要がある。現時点では

RNA 編集は、作物の品種改良に利用するメリットはなさそうに考えられた。

(8) 人材育成 (統計学、バイオインフォマティクス、AI 分野)

本研究班では、バイオ系 (ウェット) の実験者と情報科学 (ドライ) を専門とする研究者がお互いに連携して、理解して研究を進めていく事が研究開始当初より何よりも重要な点であった。また、近年のビッグデータの取り扱いに関するスキルも分野関係なく求められている現状から、生化学部では、定期的にゲノム解析に関する実習演習を行いドライ解析のスキルの取得に努めた結果、現在では十分な解析が実行できる環境を構築できた。また、その他データ解析に必要なプログラミング等について分担研究者との打ち合わせの中で取得に努めてきた結果、基本的な手法についての理解ができた。

今後はさらにビッグデータの複雑な処理が実行可能な環境と人材育成に努める。

D. その他

内容の詳細については各分担報告書を参照。

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」
総合分担研究報告書

ゲノム編集に関する情報収集解析、アレルゲンタンパク分解性検討、および
人工ヌクレアーゼの特異性を調べる *in vitro* アッセイツールの開発

研究分担者 近藤 一成 （国立医薬品食品衛生研究所）
研究分担者 中村 公亮 （国立医薬品食品衛生研究所）

研究要旨：

本研究では、ゲノム編集技術を用いた動物・植物の開発状況調査、ケーススタディーを実施して今後の開発傾向や課題の抽出を行った。その結果、研究開発はコメやコムギなどの主要作物のほか、リンゴなど果樹でも進んでいる事がわかった。アレルゲン性評価におけるタンパク分解性の反応条件における影響の検討では、人工胃液の pH 変化が分解性に大きく影響する事がわかった。ゲノム編集技術を用いたときの意図しないゲノム上の変化を検出するための手法を検討して、簡便に全ゲノム解析を実施する事なくゲノム全体で起きる DNA 二本鎖切断を検出するツールを開発した。今後一般に公開していく。

研究協力者

中島 治 （国立医薬品食品衛生研究所）
小野 竜一 （国立医薬品食品衛生研究所）
木俣 真弥 （国立医薬品食品衛生研究所）
秋本 智 （国立医薬品食品衛生研究所）
成島 純平 （国立医薬品食品衛生研究所）

A. 研究の目的

ゲノム編集技術や合成生物学を応用した食品や食品素材の開発研究が活発に行われている。ゲノム編集作物では、従来の遺伝子組換え作物のような外来遺伝子を導入することはなく、内在性遺伝子の配列を数塩基欠失により機能欠失させて新たな形質（もち性向上、筋肉量増加、GABA 量増加など）を付与できる。このようなゲノム編集食品の届出制度が令和元年に開始され、安全性に関する必要な情報が確認できれば遺伝子組換え食品とはみなさず安全性審査は不要となる状況となっている。しかし、ゲノム編集技術を応用した食品の安全性の確認に必要な解析手法は定まっていない。ゲノム編集技術特有の問題なども想定されることから、新たな評価手法開発と合わせて研究開発状況や規制制度の調査などを行い検討すべき課題を抽出しておくことは重要である。合成生物学を利用した場合は、大規模な遺伝子改変が行われるため遺伝子組換え食品に該当すると考えられるが、従来の組換える前後の比較による相対的評価、

いわゆる実質的同等性の確認による手法が適応できず、絶対的な評価が必要な場面も出てくると考えられる。そのためには、解析手法はどうあるべきかを検討して必要な評価手法の開発検討を行う事が重要である。このような背景から本分担研究では、以下の3つの検討、(1) データベースを用いて研究開発状況の調査、(2) 人工胃液を用いたタンパク分解性試験、(3) 意図しないゲノム上の変化を解析するためのシステム開発、を実施する。

B. 研究方法

(1) ゲノム編集技術を用いた動植物の研究開発動向調査およびアレルゲン分解性の検討

研究開発動向は、PubMed および有料データベース SciFinder を用いて行った。検索漏れがないようにキーワードを組合せて検索を行った後に重複を除いて整理した。欧州においては、人工胃液による分解性試験を、異なる反応条件で実施することを推奨している。そのため、分解性試験時における反応条件の違いを明らかにしておく必要がある。ここではアレルゲンタンパクのタンパク分解性試験は、人工胃液を作製して異なるペプシン酵素濃度、pH で分解性の変化と患者血清との反応性を電気泳動、ウェスタンブロットを用いて解析した。また、分解していない純度が高いタンパク試料が必要なため、アレルゲンタンパクは精製度が十分なものは市販のものを、精製度が十分でな

いものは食品より抽出精製したものをを用いた。

(2) 人工ヌクレアーゼの特異性を調べる *in vitro* アッセイツールの開発

全ゲノム解析をすることなく、ゲノム全体に対して *in silico* では予測できない意図しない変化である DNA 二本鎖切断を検出するための手法と解析環境の構築を行った。

予測できない意図しない変化である DNA 二本鎖切断部位の予測を、ゲノム DNA に対して CRISPR/Cas9 を用いて切断誘導することで明らかにして (SITE-Seq 法により)、それを実際の生物で検証を行う。すなわち、まずイネ細胞から抽出したゲノム DNA を用いて潜在的な DNA 二本鎖切断部位を予測する。本方法で予測した切断候補部位が、本当に生きた生物でも検出されるかどうかを確認するために、イネをゲノム編集技術のツールである CRISPR/Cas9 で形質転換し、得られたカルスを用いて解析した。ゲノム解析は、解析用サーバーを用いて行った。

(3) 人材育成 (統計学、バイオインフォマティクス、AI 分野)

生化学部内では、定期的な実習トレーニングを行うとともに、分担研究者および協力研究者と共同で行うことで、インフォマティクス関連技術の取得を目指す。

C. 研究結果および考察

(1) ゲノム編集に関する情報収集解析、ケーススタディーおよびアレルゲン分解性の検討

ゲノム編集技術を中心に研究開発の文献を 2018 年~2020 年について調査した結果、動物、植物ともに技術的には CRISPR/Cas を用いたもの中心であった。国内では研究開発が盛んな魚類では、諸外国ではウナギ、ドジョウなど多様な生物に应用されていたが、論文報告件数は少ない。植物では、イネでの研究が最も多く、トマト、コムギ・オオムギ、ダイズが多く研究されている。また、リンゴ、イチゴなど果樹への応用も進んでいることが分かった。形質は環境耐性から栄養改変まで多岐にわたる。また、合成生物学分野では、欧米を中心に研究が進んでいるが具体的に商品化へ進んでいるものは少ないと考えられた。

アレルゲン分解性については、ピーナッツアレ

ルゲン Ara h1 など、pH の少しの変化、例えば pH1.2 から pH3 以上へ変更しただけで、完全に分解していたものがほとんど分解されなくなるタンパク (アレルゲン) があることが確認できた。一方、消化酵素であるペプシンの濃度条件はあまり分解性には影響を与えなかった。今回の検討結果は、これまでアレルゲン性評価のデータとして検討されてきたアレルゲンタンパク分解性試験のあり方を再検討する必要があるかもしれないことを示唆していると考えられた。アレルゲンタンパク分解性試験のある条件で分解性が十分でなかったときには、タンパクアレルゲン性が懸念されるのか、その可能性がないのかを明らかにするための次の新たな試験が必要になってくるが、それに対する検討は諸外国においても行われていない。今後、検討を行っていくことが必要である。

(2) ゲノム網羅的に DNA 二本鎖切断部位を検出する手法とツールの開発検討

イネをモデル生物種として、論文に報告されている遺伝子を標的として、ゲノム編集操作を行った場合の意図しない DNA 二本鎖切断部位 (オフターゲット) 解析手法を検討した。

ゲノム DNA を利用する *in vitro* 法である SITE-Seq 法を用いて信頼性のある潜在的な切断部位候補を抽出し、7つのオフターゲット候補を得た。これらの候補部位 (オフターゲット) の多くは、Web 上で標的配列 (オンターゲット) と類似性がある配列を検索する *in silico* 検索では抽出できないことが分かった。次に、これら SITE-Seq 法を用いて得られた候補部位について、実際にイネを形質転換し、そのカルスからの DNA をアンプリコンシークエンスで解析した結果、1つのオフターゲットで実際に切断が確認された。以上の結果から、SITE-Seq 法と *in vivo* 検証を組合せたオフターゲット検証法は、*in silico* 解析だけではわからない意図しない変化も明らかにできることから、非常に有用な方法であることが示された。

さらに、ゲノム上の変化から、新たなアレルゲンタンパクの生成がないかを調査した。その結果、80-Mer Sliding Window FASTA Search および 8-mer FASTA Search のいずれでも既知アレルゲンとホモロジーのある配列は検出されなかった。

オフターゲット検索からアレルゲン予測まで含めた一連の解析スキームを確立することで、一つの安全性評価法として有効であると考えられる。

この一連の解析作業を、web ベースで行うための解析環境（Galaxy ツール）も独自に開発した。

（3）人材育成（統計学、バイオインフォマティクス、AI 分野）

本研究班では、バイオ系の実験者と情報科学（ドライ）を専門とする研究者がお互いに連携して、理解して研究を進めていく事が研究開始当初より何よりも重要な点であった。また、近年のビッグデータの取り扱いに関するスキルも分野関係なく求められている現状から、生化学部では、定期的にゲノム解析に関する実習演習を行いドライ解析のスキルの取得に努めた結果、現在では十分な解析が実行できる環境を構築できた。また、その他データ解析に必要なプログラミング等について分担研究者との打ち合わせの中で取得に努めてきた結果、基本的な手法についての理解ができた。今後はさらにビッグデータの複雑な処理が実行可能な環境と人材育成に努める。

D. 結論

ゲノム編集技術を用いた動物、植物の研究開発動向調査では、イネ、コムギのほかリンゴなどの果樹への応用も進んでいることが分かった。合成生物学分野では、欧米を中心に研究が進んでいるが商品化へ進んでいるものは少ない。

タンパクアレルゲン分解性の検討は、実際の人の胃内の環境を考慮して（食事 pH は上昇する）異なる複数の反応条件で実施することが必要と考えられた。

ゲノム編集操作による意図しないゲノム上の変化を解析するための解析環境を開発した。本手法により、*in silico* 検索で明らかにできないオフターゲット候補も解析可能になると考えられた。

E. 研究発表・業績

各年度分担報告書を参照

F. 健康危険情報

各年度の分担報告書を参照

G. 知的財産権の出願・登録状況

各年度の分担報告書を参照

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」
総合分担研究報告書

リスクコミュニケーションに関する研究

研究分担者 小泉 望 （大阪府立大学）

研究要旨：

新たなバイオテクノロジーを用いて得られた食品のうち、主としてゲノム編集食品に関する効果的なリスクコミュニケーション手法の確立を目的に Web アンケート調査、国内外の遺伝子組換え食品も含めたコミュニケーション活動の包括的調査、2冊の冊子作成を行った。アンケート調査の結果からは専門家への信頼が鍵因子で、一般回答者の知りたいことと専門家の伝えたいことに強い相関があることから、専門家が伝えたいことを適切な方法で伝えることが重要である。海外では多様なステークホルダーが参画するコミュニケーション活動が行われており、その結果が公開されている。日本の活動でも参考にできる。よりよいコミュニケーションのために作成した平易な冊子、会話形式の Q&A 集をリスクコミュニケーション活動に活用し、その効果を検証することが望まれる。

A. 研究目的

遺伝子組換え食品あるいはゲノム編集技術応用食品（以下、ゲノム編集食品）といった新たなバイオテクノロジーを用いて得られた食品に関して国民、中でも一般の人（専門家でない人）、の疑問や不安が大きいと考えられる。令和元年度にはゲノム編集食品に関する安全性や表示の取扱いルールが決まり、関連する報道も多く見られたが、十分なリスクコミュニケーションが行われ、一般の人の疑問や懸念が解消されたかどうかは定かで無い。また疑問あるいは懸念の内容や専門家と非専門家との間の認識の差も明確で無い。このような状況を考慮し、主にゲノム編集食品を対象としてリスクコミュニケーションを適切に行うために、1) Web アンケート調査、2) 国内外におけるリスクコミュニケーション事例の包括的調査、3) リスクコミュニケーションの場に提供するための冊子2冊の作成を行うことを目的とした。

B. 研究方法

1) Web アンケート調査

一般モニター4,000人、専門家398人を対象としたゲノム編集食品に対する Web アンケート実施した。調査結果について種々の統計解析を行った。

2) 主に海外におけるリスクコミュニケーション

事例の包括的調査

過去5年以内の国内外の遺伝子組換え食品あるいはゲノム編集食品に関するリスクコミュニケーション、サイエンスコミュニケーションの38実施例を Web 検索により抽出した。調査項目は事例名、国名、実施時期、実施名、概要、参考 URL である。さらに、そのうち15事例について詳細な調査を行った。

3) リスクコミュニケーションの場に提供するための冊子の作成

令和元年度はゲノム編集食品および遺伝子組換え食品に関する平易な冊子を作製した。令和2年度はゲノム編集技術で作出された毒を作らないジャガイモ、肉厚のタイ、GABA 高蓄積トマトに関して、消費者、事業者、開発者、行政などの多様なステークホルダーが参加する公開および非公開の座談会を複数回実施した。そこで出た意見、質問を参考に、良く出る疑問を整理した。その疑問に対する分かりやすい答えを用意し、会話形式の冊子の作成に活かした。

C. 研究結果

1) Web アンケート調査の解析

一般回答者4,000人は、登録しているモニターを対象に性別、年齢が均等になるように設定し、専門家398人は学会や研究者コミュニティのメー

リングリストによりボランティアを募った。ゲノム編集を巡る認知、社会受容に関する認識については、一般回答者における認知度は低かったが社会受容に関する否定的な態度も必ずしも高くなかった。専門家コミュニティと一般回答者の間におけるゲノム編集を巡る認識の比較では、規制に関しては一般回答者がゼロリスクを強く求めるのに対して専門家は科学的妥当性を重視した。表示については専門家の半数以上も表示すべきと回答した。ゲノム編集食品が受容されるために重要であることは一般回答者と専門家では異なっており、一般回答者がリスクの有無に強く関心を持っているのに対して、専門家は科学的妥当性、あるいは技術の必要性を挙げた。興味深い結果として、ゲノム編集食品に関して「知りたい事柄」と「伝えたい事柄」に高い共通性が見られた。

一般回答モニターにおけるゲノム編集への意識の因子分析を行い、ゲノム編集食品の社会的受容や専門家の信頼感をめぐる背景について分析したところ、専門家への信頼度が社会受容に大きく貢献することが示された。科学リテラシーが高いことは社会受容につながっているが、一方、Knowledge Difference Score (KDS) が高いことも社会受容にポジティブである傾向も見られた。パス解析を行った結果からは、ゲノム編集食品にネガティブな層は専門家信頼やゲノム編集食品の実食可能性の評価が低くなる傾向にある(図1)。注目すべきは、「農業・食文化価値重視」因子が大きいほど、つまり有機農業や伝統的な食文化などへの関心や評価が高いほど、ゲノム編集食品のリスク関心が高くなる傾向にあることである。しかしながら、「農業・食文化価値重視」は「科学技術肯定」とも正の相関関係があり、「科学技術肯定」は「ゲノム編集食品のベネフィット関心」を引き上げる影響も示された。

2) リスクコミュニケーション事例の包括的調査

英国8事例を筆頭に国内外のコミュニケーション事例38件について調査した。国内は8事例である。国によって状況は異なるものの異なるステークホルダーが参画する対話形式の取り組みが多く、その内容がWebサイト上で公開されている。例えば英国ではEUを離脱することで新しい規制の枠組みを構築できることになるため、王立協会が2017年に市民対話、大規模調査が行われている。ベルギーを中心に行われているCHICというプロ

ジェクトでは、チコリという植物を遺伝子工学の技術を用いて多目的に利用しようという試みを産業、農業、学術、消費者といったステークホルダーが様々な視点から長所、短所について議論することが2019年に行われている。遺伝子組換え食品に否定的な風潮の強い欧州に限らず比較的肯定的に見える米国においても2020年に市民の信頼の獲得と情報提供についての透明性を高めるための取り組みが行われている。日本での取り組みについては情報発信が十分で無いように見受けられる。

3) リスクコミュニケーションの場に提供するための冊子の作成

令和元年度に作成した「新しいバイオテクノロジーで作られた食品について」

<https://www.mhlw.go.jp/content/11130500/000657695.pdf>

に続かたちで「新しいバイオテクノロジーで作られた食品を語る」を作成した。実際に複数回行った座談会や他のプロジェクトでのコミュニケーション活動から抽出したFAQsを整理し、ゲノム編集食品のリスクコミュニケーションに使いやすい冊子を作成した。単なるQ&A集ではなく3人の登場人物(研究者、主婦、高校生)の対話形式を取っており一般の方に身近に感じてもらえるような形式とした。

D. 考察

ゲノム編集食品に関する一般モニター4,000人、ゲノム編集の専門家398人を対象としたWebアンケート調査より、現状ではゲノム編集食品に対する一般モニターの受け止め方は遺伝子組換え食品のようにネガティブではないと思われた。もっとも、ゲノム編集食品に関する知識、関心が低く結果的に「わからない」あるいは「様子見」の姿勢であり、ネガティブな態度として顕在化していないだけかもしれない。また、科学リテラシーはゲノム編集食品の受容と相関はあるが、必ずしも科学リテラシーが上がれば受容度が向上するとは言えないと考えられた。専門家の発言に対する信頼度が受容には重要であると伺えた。専門家が伝えたいことと一般モニターの知りたいことはよく一致しており、専門家はリスクコミュニケーションにおいて伝えたいことを伝えるのが良いが、その伝え方には工夫が必要であると考えられた。

注目すべき点として、従来の農業体系や食文化

価値の重視はゲノム編集食品の受容にネガティブに働く傾向が見て取れた。つまり、ゲノム編集食品を異質なものとしてとらえる価値観がこうした受容の障害につながると考えられる。ゲノム編集技術が品種改良（育種）の一つの手法に過ぎず伝統的な農業体系と競合するものでないことの理解が受容につながる可能性が考えられた。

遺伝子組換え食品の社会受容が困難であることは日本に限らない。特に欧州では日本よりもネガティブに捉えられていることが少なくない。米国でも表示が行われることになるなど、生産が盛んだから社会受容が順調とは必ずしも言えない。このような背景から、世界各国でゲノム編集技術あるいはゲノム編集食品に関して種々のコミュニケーション活動が実施されている。英国は EU を離脱すること、遺伝子組換え食品が非常にネガティブに捉えられてきたことなどから 2017 年に王立協会が中心になり公開討論や意識調査を行い、その結果を公開している。その後も英国では公的機関によるコミュニケーション活動が継続的に行われ、EU では英国に限らず、ベルギー、フランス、ドイツなどでもゲノム編集に関するコミュニケーションが 2018 年から 2020 年にかけて行われている。内容は精査の余地があるが、EU では 2018 年にゲノム編集食品を従来の遺伝子組換え食品と同様に規制するという判決が欧州司法裁判所によってなされてはいるものの、ゲノム編集食品に関するコミュニケーション活動は継続的に行われている。遺伝子組換え食品に比較的寛容な米国あるいはアルゼンチンでもゲノム編集食品に関するコミュニケーション活動が行われている。多くの事例で共通する点は、異なるステークホルダーが参画して多様な視点から議論を行い、その情報を、Web 等を通して多くの人が共有できる仕組みを取っていることである。この点は日本で行われているコミュニケーション活動の多くと異なる。日本では双方向と言いながらやはり講師と参加者の質疑応答に近いものが多く、活動の結果の共有も充分でない。結果として多くの情報はメディアに依存するが、その情報は両極端の両論併記であることが少なくない。こうした状況を考えれば、日本でも多様なステークホルダーの意見を取り入れるコミュニケーション活動を実施し、その内容を多くの国民が容易に知ることができる仕組みが求められる。YouTube の利用なども検討の余地がある。

上述の多様なステークホルダーの意見を取り入

れる大規模コミュニケーション活動の実施にはやはり限界がある。そこで小規模なイベントの実施が求められ、生協などでは勉強会が行われている。そうした際、あるいはその前段階での資料として分かりやすい冊子が必要であると考えた。また多様な意見を冊子に盛り込むことも考えた。結果として令和元年に必要な最小限の情報を載せた平易な冊子を作り、令和 2 年度は 3 名の登場人物：研究者、その従姉弟である主婦、そしてその娘（高校生）を設定し、会話形式でのゲノム編集食品について Q&A 集を作成した。そのコミュニケーション活動への貢献は今後、この冊子が実際にコミュニケーションの場で使用されることによって明らかとなると考えられる。

E. 結論

現状では日本におけるゲノム編集食品への受容状況は明確でないが、遺伝子組換え食品と比べるとネガティブではないと考えられた。ただし、それは認知度が低く、分からないからともいえる。遺伝子組換え食品の場合は危険情報が先行したが、ゲノム編集食品の場合、そのような状況は見られない。従って、今後の取り組みが重要と考えられる。専門家への信頼が重要な要素であり、一般回答者の知りたいことと専門家の伝えたいことに強い相関があることから、専門家が伝えたいことを適切な方法で伝えることが重要である。また海外のリスクコミュニケーション事例からは、異なるステークホルダーの意見収集とその情報発信が重要と考えられ、専門家を巻き込んだそうした場の創出が求められる。海外の事例調査について今後、その効果等についても検証していく必要がある。

様々なリスクコミュニケーションの場においてゲノム編集食品に関する分かりやすい資料が求められる。本研究で作成した会話形式の資料の効果についても今後検証が必要である。

F. 研究発表・業績

1. 論文発表

- 1 Matsuda, M., Iwata, Y., Koizumi, N., & Mishiba, K.-I. (2020). Zeocin-induced DNA double strand breaks affect endoreduplication and cell size in radish cotyledon epidermis. *Cytologia*, **85**, 245-249.
- 2 Mishiba, K.-I., Nishida, K., Inoue, N., Fujiwara, T., Teranishi, S., Iwata, Y., Takeda, S., &

- Koizumi, N. (2020). Genetic engineering of eggplant accumulating β -carotene in fruit. *Plant Cell Reports*, **39**, 1029-1039.
- 3 Mishiba, K.-I., Iwata, Y., Mochizuki, T., Matsumura, A., Nishioka, N., Hirata, R., & Koizumi, N. (2019). Unfolded protein-independent IRE1 activation contributes to multifaceted developmental processes in Arabidopsis. *Life Sci Alliance*, **2**, e201900459.
- 4 Hirata, R., Mishiba, K.-I., Koizumi, N., & Iwata, Y. (2019). Deficiency in the double-stranded RNA binding protein HYPONASTIC LEAVES1 increases sensitivity to the endoplasmic reticulum stress inducer tunicamycin in Arabidopsis. *BMC Res Notes*, **12**, 580.
- 5 Hirohata, A., Sato, I., Kaino, K., Iwata, Y., Koizumi, N., & Mishiba, K.-I. (2019). CRISPR/Cas9-mediated homologous recombination in tobacco. *Plant Cell Reports*, **38**, 463-473.

2. 学会発表

- 1 小泉 望、山口 夕、標葉隆馬 (大阪大学)「日本におけるゲノム編集食品に対する国民の意識：利益、リスク、信頼への関心」、第 62 回日本植物生理学会、令和 3 年 3 月 15 日、島根大学 (オンライン開催)
- 2 小泉 望 (大阪府立大)「ゲノム編集食品に関するパンフレット」、日本植物生理学会、令和 2 年 3 月 19 日、大阪大学 (ポスター発表) ※新型コロナウイルスのため中止、学会としては成立
- 3 小泉 望 (大阪府立大)「ゲノム編集食品の現状と課題」、日本食品微生物学会、令和元年 12 月 12 月 20 日、I-site なんば (招待講演)
- 4 小泉 望 (大阪府立大)「科学技術の社会実装のためのコミュニケーション ゲノム編集食品のルール作りを例に」、日本サイエンスコミュニケーション協会、三鷹ネットワーク大学、令和元年 12 月 7 日

G. 健康危険情報

各年度の分担報告書を参照

H. 知的財産権の出願・登録状況

各年度の分担報告書を参照

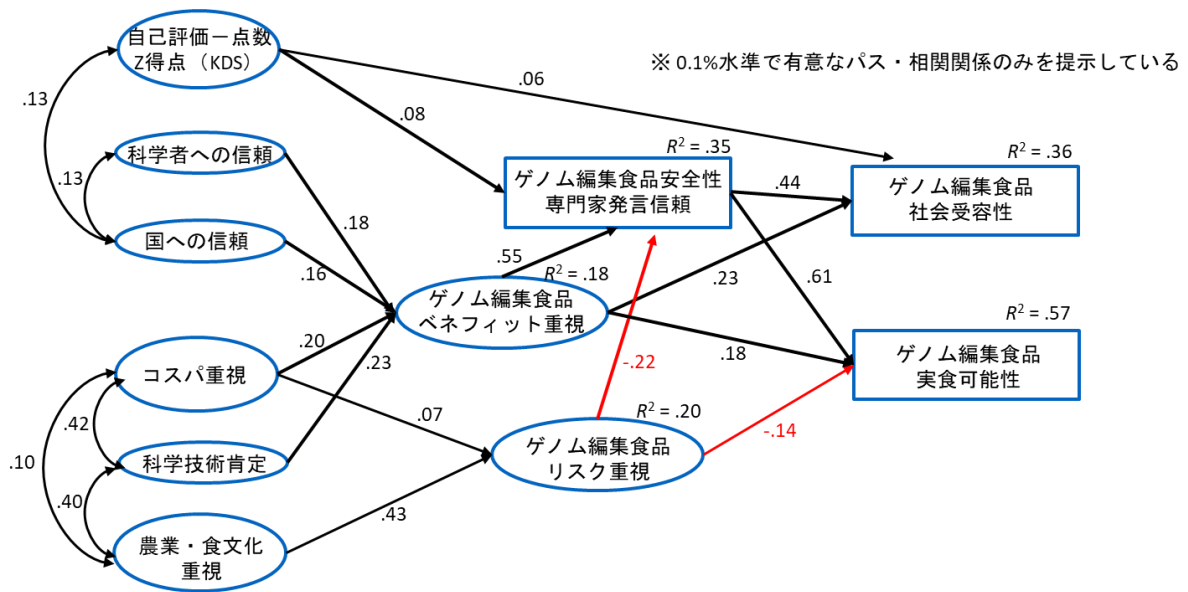


図1 パス解析の結果

メタボロームインフォマティクスによる未知化合物推定

研究分担者 早川英介（沖縄科学技術大学院大学）

研究要旨：

本研究は食品試料の内在性低分子化合物の質的な変動を迅速に検出するデータ解析システムの開発を目的とし、質量スペクトルの類似度をもとにした食品中の未知化合物の検出と構造推定のためのデータ解析システムの開発を行った。スペクトルライブラリ構築、アルゴリズム開発、可視化法の開発を通して、スペクトル類似度計算を核とした迅速な未知化合物の検出と構造推定などが行われるシステムを構築した。

A. 研究目的

本研究はゲノム編集作物のバイオテクノロジーにより作成された食品試料の内在性低分子化合物の質的な変動を迅速に検出するデータ解析システムの開発である。

従来の分析化学ではあらかじめ分析対象の化合物を決めて行う“ターゲット分析”が行われることがほとんどであった。しかし、ゲノム編集食品のように想定外の代謝の変化により未知化合物が懸念されている場合、そのような“ターゲット外”の化合物は従来の化学分析では見過ごされてしまう。本研究では構造的に類似した化合物の質量スペクトル（フラグメントスペクトル）が類似するという特性に着目し、質量スペクトルライブラリやケモインフォマティクスによる化学情報を統合したこれまでにないデータ解析プラットフォームを構築し、食品試料中の未知化合物の迅速な検出と構造推定を可能にすることを目的とした。（図1）

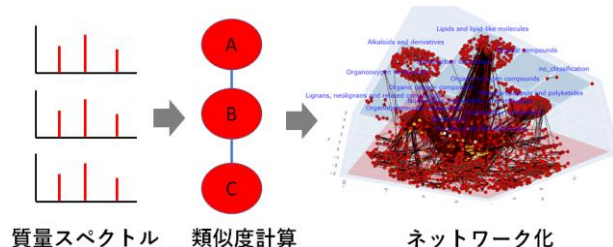


図1

B. 研究方法

本研究では化合物の質量スペクトル（フラグメントスペクトル）の類似度をもとに食品中の未知化合物の検出と構造推定を行うシステムの構築を行った。

データ解析のフローとしては、LC-MS から得られる比較定量情報を伴った質量スペクトルを入力とする。安全と考えられる食品試料（例：非ゲノム編集体）と分析対象の試料（例：ゲノム編集体）の比較定量により、注目している試料で濃度が増加している未知化合物を見出すことができる。

さらに、検出された全質量スペクトルに関して、試料内および質量スペクトルライブラリの全スペクトルに対して総当たりでスペクトル類似度計算を行うことで、構造的に類似している化合物を見出す。この質量スペクトルライブラリに関してはパブリックな質量スペクトルライブラリに加え、in house で作成した化合物標準品と、内在性化合物情報が既知である代表的な食品・植物の抽出物を分析したことで得たライブラリを統合したものを構築した。

ライブラリ内の全化合物に関して、Classyfire により化合物クラスを super class、class、subclass の各階層で付与し、さらに外部毒性情報データベースから抽出した情報を付与した。各化合物の構造とフラグメントスペクトルを解析し、フラグメントピークに相当する部分構造を推定し、データベースに統合した。（図2）

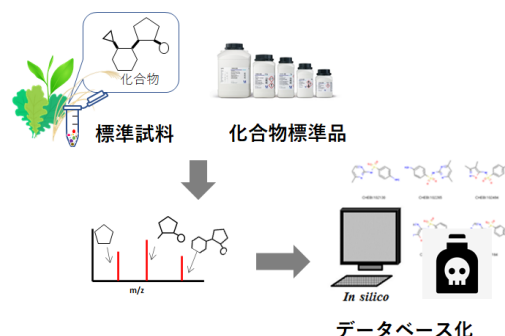


図2

スペクトル類似度計算とこの統合質量スペクトルライブラリを連携することで、試料中の未知化合物の質量スペクトルから化合物クラス、毒性情報、部分構造を得るシステムを構築した。このデータ解析システムの実行環境として、各種パラメータをインタラクティブに反映させることが可能なGUIをPlotly Dashを利用してWebブラウザ上で実行可能なシステムとして実装した。

C. 研究結果

本研究では「B. 研究方法」で述べたデータ解析ワークフローに関し、プログラミング言語：Pythonを用い、ケモインフォマティクスライブラリ：RDKit、データ可視化ライブラリ：Plotly等と連携させることでデータ解析環境の構築を行った。入力データは一般に広く用いられる質量分析データ処理フリーソフト（Mzmine2）の出力形式に対応させた。

フラグメントスペクトルライブラリはMassBankに加え、天然化合物を多く含む GNPS、Respectを加えて特に植物二次代謝物に対応できるよう整備した。加えて独自で500種類の低分子化合物を分析しライブラリに追加した。さらに代表的な食品・モデル植物等（ダイズ、トマト、ジャガイモなど）50種類以上から抽出した低分子化合物に関してもスペクトルライブラリに追加した。

この膨大なスペクトルライブラリに対して、「B. 研究方法」で述べた化合物クラス・ピーク部分構造・毒性情報を付与することで、網羅的かつ極めて情報が豊富な独自性の高いスペクトルライブラリが構築された。スペクトル類似度計算とこのライブラリをデータ解析プラットフォーム上で統合することで、スペクトル類似度に基づいた未知化合物の構造解析を迅速に行うシステムが実現された。

さらに、このデータ解析システムをWebブラウザベースの実行環境で動作させることで、プログラミングなどの専門知識がない一般的な研究者でも容易にデータ解析とデータ可視化ができるようになった。

D. 考察

本研究の目的はゲノム編集作物等で想定外の質・構造的変化を生じる化合物の迅速な検出と構造の推定である。従来の質量分析のデータ解析では困難だった未知化合物の検出と構造推定をスペクトル類似性に着目した独自の手法を採用することで実現した。

手法の特徴として、試料中の未知化合物に関して化合物クラスや部分構造、毒性情報など多面的な情報を出力するという点がある。これは単に構造を推定するという点以外にも、変動した化合物“群”の

クラスや共通構造を見出すことであり、分析対象の食品の未知化合物グループを包括的に明らかにすることにもつながるなど、応用範囲は広い。単に試料中の未知化合物を明らかにするだけではなく、変動した代謝パスウェイの推定など、様々な利用法も考えられる。

本解析プラットフォームのもう一つの特徴は、上記のような複雑なデータ解析を容易に行うことが可能という点である。Webブラウザ経由でインタラクティブに出力結果をみながら操作を行うことで、データ解析に不慣れな研究者も操作が可能であり、幅広く普及されることが期待できる。（図3）



図3

E. 結論

本研究により、様々な食品試料の内在性化合物の質的・量的な変動を迅速に解析できるプラットフォームが確立された。この手法はゲノム編集作物のバイオテクノロジーにより作成された食品試料中の未知化合物の分析に極めて有効であり、化合物クラス、部分構造、毒性予測などの豊富な化学情報は、単一の化合物の構造解析にとどまらず、未知化合物群とその代謝パスウェイを包括的に理解することにつながる。

本データ解析システムは一般の研究者でも容易に利用可能なプラットフォームであり、広範囲の食品分析に今後利用されることが期待される。

F. 研究発表・業績

1. 論文発表

該当なし

2. 学会発表

- 1) 日本食品化学学会 第26回総会・学術大会 (2020.5.28)、書面開催「多層質量スペクトルネットワークによる食品中の未知化合物の可視化」
- 2) 日本毒性学会 第47回日本毒性学会学術年会 (2020.6.29)、オンライン開催「An analytical framework using multilayer mass spectral

similarity network to detect unknown compounds for toxicological analysis.」

G. 健康危険情報

各年度の分担研究報告書を参照

H. 知的財産権の出願・登録状況

各年度の分担研究報告書を参照

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
総合分担研究報告書

高精度アレルギー性予測システムの構築に必要な情報の収集に関する研究

研究分担者 爲廣 紀正（国立医薬品食品衛生研究所）

研究要旨：

本研究では、AI を活用した高精度アレルギー性予測システムの構築に向けて学習データの収集と精査を行い、最終的にはアレルギー 2,248 種および非アレルギー 18,906 種を機械学習用データセットとして整備した。また、バイオテクノロジーを用いて得られた食品のリスク管理に関する研究として、弊所で運用・公開しているアレルギーデータベース(ADFS; Allergen Database for Food Safety)のアレルギー情報及びエピトープ情報を追加し、データベースの更新を行った。その結果、3 年間で、アレルギー及びイソアレルギーのアミノ酸配列情報 195、及び、23 種のアレルギーについて総数 66 のエピトープ情報を追加し、アレルギー及びイソアレルギーのアミノ酸配列情報は 2,339 となり、エピトープ既知のアレルギー数は 247 となった。また、システム老朽化が問題視された ADFS の再構築を行い、ADFS 利用に際する脆弱性を改善した。

研究協力者
安達 玲子（国立医薬品食品衛生研究所）

A. 研究目的

現在、様々な遺伝子組換え食品が、生産性の向上や栄養付加を目的として開発されている。組換え食品の分野では、植物だけでなく、動物を宿主とした開発も進んでおり、また最近では、遺伝子組換え植物同士を交配して、付与された機能をスタックすることにより得られるスタック品種も開発されている。しかし、これらのようにバイオテクノロジーを利用して得られた品種について、どのような意図しない形質変化が出現するかを研究している例は少ない。したがって、新たに得られる遺伝子組換え生物について、非意図的な影響等を考慮し、安全性評価の方法等を検討する必要がある。

バイオテクノロジー技術を用いて開発された遺伝子組換え食品のリスクの 1 つの可能性として、アレルギー性増大が考えられる。本研究では、アレルギー性解析法の 1 つとして国立医薬品食品衛生研究所で管理・公開している、アレルギー性予測機能を装備したアレルギー・エピトープ情報データベース(ADFS; Allergen Database for Food Safety)に関して、その情報内容を更新し、充実させることにより、遺伝子組換え食品のリス

ク管理の上で必須であるアレルギー性評価系に関する研究を行う。また、アレルギー予測システムに AI を搭載する事で、予測精度の飛躍的向上を試みる。

B. 研究方法

AI 学習用データセットの準備

米国環境保健科学研究所が組織する HESI により包括的な既知あるいは推定アレルギータンパク配列のレポジトリ“COMprehensive Protein Allergen REsource (COMPARE)”が公開されている。アレルギーの専門家によるピアレビューを経て、情報を公開している点では AllergenOnline と同じであるが、最近改めて情報を整理し 2017 年より公開が開始されたことから、現在最も信頼性の高いデータベースと考えられる。そこで、アレルギータンパク情報として COMPARE の登録タンパク情報を入手した。一方、非アレルギー情報は、わが国で加工食品へのアレルギー表示が義務付けされている特定原材料 7 品目、並びに推奨されている原材料を含む 13 品目について UniProt から査読タンパク情報を収集し、アレルギー性を持つ或いは持つと予想されるタンパクを除いたタンパク情報を整備した。加えて、T 細胞の成熟過程において様々な自己タンパクを抗原提示し、自己タンパクに対する免疫寛容の基盤

を形成する胸腺髄質上皮細胞 (mTEC) の発現タンパクについて、プロテオーム解析ならびに遺伝子発現解析を行った論文から発現タンパク情報を収集・解析し、非アレルギー情報に加えた。

登録アレルギー (アミノ酸配列情報) のアップデート

米国ネブラスカ大学リンカーン校が運営しているアレルギーデータベース (AllergenOnline) における登録アレルギーのアップデート内容を、ADFS に反映させた。

エピトープ情報の追加

平成 30 年度から令和 2 年度の各年度に、前年の 6 月から当該年 5 月までの 1 年間に NCBI PubMed に掲載された論文から、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、アレルギーデータベース (ADFS) のデータに追加した。

C. 研究結果

AI 学習用データセットの準備

アレルギータンパク情報は、令和 2 年 1 月 29 日に COMPARE からリリースされた 2,248 種のタンパク情報を入手した。また、非アレルギータンパク情報は、わが国で加工食品へのアレルギー表示が義務付けされている特定原材料 7 品目 (卵、乳、落花生、そば、小麦、えび、かに) 並びに推奨されている原材料のうち 5 品目 (いくら、さけ、キウイフルーツ、大豆、りんご)、そしてアレルギーの解析が進んでいる 9 品目 (じゃがいも、コメ、人参、トウモロコシ、マスタード、オリーブ、トマト、モモ、牡蠣) について UniProt からタンパク情報を入手・整理し、17,372 の情報を用いた。これらのアレルギー・非アレルギーデータは、学習時により効率的なパターンマイニングを実行できるように分子系統情報を整備した。一方、mTEC は細胞内で発現する組織特異的タンパク等を T 細胞に抗原として提示し、その抗原を認識するよ

うないいわゆる自己攻撃性の T 細胞を排除させる役割を持つ。健康人においては mTEC で発現しているタンパク質に対する抗体は作られない。mTEC 発現タンパク情報の収集・解析により、非アレルギー情報は 1,534 が追加され、計 18,906 となった。

登録アレルギー (アミノ酸配列情報) のアップデート

米国ネブラスカ大学リンカーン校が運営しているアレルギーデータベースである AllergenOnline は、登録アレルギーの全てが国際的なアレルギーの専門家チームによるピアレビューを経ており、登録タンパク質がアレルギーであるというエビデンスの信頼性が非常に高いデータベースである (但しエピトープ情報は含まない)。ADFS における登録アレルギーは平成 20 年度に AllergenOnline の登録アレルギーと統合し、その後も AllergenOnline のアップデートに伴って ADFS 登録アレルギーのアップデートを行っている。平成 30 年から令和 2 年度においても引き続きこのアップデート作業を実施した。

エピトープ情報の追加

エピトープ配列に関しては、平成 30 年度から令和 2 年度の 3 年間で、キーワード検索により抽出された論文は 65 報であった。その中からアレルギー・エピトープ情報が記載されていると思われる 28 報を選択し、ピアレビューを行った。その結果、20 報の論文 (表 1) から 23 種のアレルギーについて、総数 66 種のエピトープ情報を新たに追加した (表 2)。

上記のアレルギー及びエピトープ情報更新作業により、最終的に、ADFS のアレルギー及びイソアレルギーのアミノ酸配列情報は 2,339、エピトープ既知のアレルギー数は 247、構造既知のアレルギー数は 169、糖鎖付加アレルギー数は 127 となった。

ADFS 脆弱性の対応

ADFS は、平成 17 年に公開されて以降、ツールの充実化や情報更新を行っているがミドルウェア等は開発当初のバージョンのまま運用しており、データベースの脆弱性が懸念されていた。そこで、脆弱性対応のため、本研究初年度にミドルウェアのバージョンアップを進めていたが、

ADFS のサイトは不正なスクリプトを挿入することが出来る環境にあり、ADFS のエンドユーザーは不正スクリプトを利用してサイバー攻撃（クロスサイトスクリプティング）を受ける可能性があるため指摘を受けたため公開は中止し、プログラミングを java から php に変更しシステム再構築を行った。また相同検索ツールのプログラム（Blast、FASTA、Pftools）も全て最新のバージョンに更新し OS は最新の RedHat8 にバージョンアップした。加えて、新サーバ機に改修した ADFS プログラムの導入し、Web システムに対して SSL 設定を実施し、インターネット上でのデータ通信を暗号化し、クロスサイトスクリプティングに対応したウェブアプリケーションに改善した。最終的には、外部公開に先んじてペネトレーションテストを行い、検出された不具合について修正等を進めた。

D. 考察

本研究では、AI を活用した高精度アレルギー性予測システムの構築に向けて学習データの収集と精査を行い、アレルギー 2,248 および非アレルギー 18,906 を機械学習用データセットとして整備した。また、学習時に使用する要考慮情報を設定することで、抽出パターンの確度を高めることができた。また、平成 30 年度から令和 2 年度の 3 年間で、アレルギー及びイソアレルゲンのアミノ酸配列情報を 195 種追加、また、23 種のアレルゲンについて総数 66 のエピトープ情報を追加した。本研究により、遺伝子組換え食品のアレルギー性に関する評価・予測系を充実させることができ、現在までに既に開発されている遺伝子組み換え食品、及び多様化するバイオテクノロジー技術により今後作製される新規遺伝子組換え食品のアレルギー性を、より高い精度で評価・予測することが可能となった。

E. 結論

アレルギー予測システムに AI を搭載するために進めた学習データセットの質の向上により、高精度予測を実現可能にした。

平成 29 年 6 月から令和 2 年 5 月までの 3 年間に NCBI PubMed に掲載された論文から、エピトープ配列決定に関する 20 報のピアレビューを行い、23 種のアレルゲンについて、総数 66 のエピトープ情報を ADFS に追加した。また、

AllergenOnline の登録アレルギー（アミノ酸配列情報）に関するアップデートを ADFS に反映し、これらの情報更新により遺伝子組換え食品のアレルギー性評価・予測方法を装備した ADFS のデータベース情報を充実することができた。

F. 研究発表・業績

1. 論文発表

なし

2. 学会発表

- 1) 爲廣紀正、後藤健仁、吉田拓未、花田博幸、佐久間拓斗、安達玲子、竹内一郎、近藤一成. 食の安全性確保に向けたアレルギー性予測機械学習モデルの開発. 第43回日本分子生物学会年会(2020.12 オンライン)
- 2) Norimasa Tamehiro, Reiko Adachi, Kazunari Kondo: Functional food ingredients modulate mast cell signaling. JSA/WAO Joint Congress 2020 (2020.09.17-10.20)

G. 健康危険情報

各年度の分担報告書を参照

H. 知的財産権の出願・登録状況

各年度の分担報告書を参照

表 1 平成 30 年度ピアレビューによりエピソード情報を収集した論文

1.	Fu L, Wang J, Ni S, Wang C, Wang Y. Identification of Allergenic Epitopes and Critical Amino Acids of Major Allergens in Chinese Shrimp (<i>Penaeus chinensis</i>) by Immunoinformatics Coupled with Competitive-Binding Strategy. <i>J Agric Food Chem.</i> 2018 Mar 21;66(11):2944-2953. PMID:29481756
2.	Curin M, Weber M, Hofer G, Apostolovic D, Keller W, Reininger R, Swoboda I, Spitzauer S, Focke-Tejkl M, van Hage M, Valenta R. Clustering of conformational IgE epitopes on the major dog allergen Can f 1. <i>Sci Rep.</i> 2017 Sep 22;7(1):12135. PMID:28939849
3.	Reginald K, Chew FT. Conformational IgE Epitope Mapping of Der p 2 and the Evaluations of Two Candidate Hypoallergens for Immunotherapy. <i>Sci Rep.</i> 2018 Feb 21;8(1):3391. PMID:29467434
4.	Curin M, Garmatiuk T, Resch-Marat Y, Chen KW, Hofer G, Fauland K, Keller W, Hemmer W, Vrtala S, Focke-Tejkl M, Valenta R. Similar localization of conformational IgE epitopes on the house dust mite allergens Der p 5 and Der p 21 despite limited IgE cross-reactivity. <i>Allergy.</i> 2018 Jan 10. PMID: 29319884
5.	Liu GY, Mei XJ, Hu MJ, Yang Y, Liu M, Li MS, Zhang ML, Cao MJ, Liu GM. Analysis of the Allergenic Epitopes of Tropomyosin from Mud Crab Using Phage Display and Site-Directed Mutagenesis. <i>J Agric Food Chem.</i> 2018 Aug 29;66(34):9127-9137. PMID:30107732
6.	He S, Zhao J, Elfalleh W, Jemaà M, Sun H, Sun X, Tang M, He Q, Wu Z, Lang F. In Silico Identification and in Vitro Analysis of B and T-Cell Epitopes of the Black Turtle Bean (<i>Phaseolus Vulgaris</i> L.) Lectin. <i>Cell Physiol Biochem.</i> 2018;49(4):1600-1614. PMID:30223257
7.	Kern K, Havenith H, Delaroque N, Rautenberger P, Lehmann J, Fischer M, Spiegel H, Schillberg S, Ehrentreich-Foerster E, Aurich S, Treudler R, Szardenings M. The immunome of soy bean allergy: Comprehensive identification and characterization of epitopes. <i>Clin Exp Allergy.</i> 2019 Feb;49(2):239-251. PMID:30267550
8.	Lahiani S, Dumez ME, Bouaziz A, Djenouhat K, Khemili S, Bitam I, Gilis D, Galleni M. Immunodominant IgE Epitopes of Der p 5 Allergen. <i>Protein Pept Lett.</i> 2018;25(11):1024-1034. PMID:30430936
9.	Cai ZL, Chen JJ, Zhang Z, Hou YB, He YS, Sun JL, Ji K . Identification of immunodominant IgE binding epitopes of Der p 24, a major allergen of <i>Dermatophagoides pteronyssinus</i> . <i>Clin Transl Allergy.</i> 2019 May 23;9:28. PMID:31139345
10.	Yamamoto K, Ishibashi O, Sugiura K, Ubatani M, Sakaguchi M, Nakatsuji M, Shimamoto S, Noda M, Uchiyama S, Fukutomi Y, Nishimura S, Inui T. Crystal structure of the dog allergen Can f 6 and structure-based implications of its cross-reactivity with the cat allergen Fel d 4. <i>Sci Rep.</i> 2019 Feb 6;9(1):1503. PMID:30728436
11.	Fang L, Li G, Zhang J, Gu R, Cai M, Lu J . Identification and mutational analysis of continuous, immunodominant epitopes of the major oyster allergen Crag 1. <i>Clin Immunol.</i> 2019 Apr;201:20-29 PMID: 29319884
12.	Zenker HE, Ewaz A, Deng Y, Savelkoul HFJ, van Neerven RJJ, De Jong NW, Wichers HJ, Hettinga KA, Teodorowicz M. Differential Effects of Dry vs. Wet Heating of β -Lactoglobulin on Formation of sRAGE Binding Ligands and sIgE Epitope Recognition. <i>Nutrients.</i> 2019 Jun 25;11(6):1432. PMID: 31242665
13.	Eichhorn S, Hörschläger A, Steiner M, Laimer J, Jensen BM, Versteeg SA, Pablos I, Briza P, Jongejan L, Rigby N, Asturias JA, Portolés A, Fernandez-Rivas M, Papadopoulos NG, Mari A, Poulsen LK, Lackner P, van Ree R, Ferreira F, Gadermaier G. Rational Design, Structure-Activity Relationship, and Immunogenicity of Hypoallergenic Pru p 3 Variants. <i>Mol Nutr Food Res.</i> 2019 Sep;63(18):e1900336. PMID: 31207117
14.	Tuppo L, Alessandri C, Giangrieco I, Ciancamerla M, Rafaianni C, Tamburrini M, Ciardiello MA, Mari A. Isolation of cypress gibberellin-regulated protein: Analysis of its structural features and IgE binding competition with homologous allergens. <i>Mol Immunol.</i> 2019 Oct;114:189-195. PMID: 31376732
15.	L'Hocine L, Pitre M, Achouri A. Detection and identification of allergens from Canadian mustard varieties of <i>Sinapis alba</i> and <i>Brassica juncea</i> . <i>Biomolecules.</i> 2019 Sep 14;9(9):489. PMID: 31540036
16.	Cai ZL, Zhang Z, Luo WL, Hou YB, He YS, Chen JJ, Ji K. Identification of immunodominant IgE epitopes of the major house dust mite allergen Der f 24. <i>Int J Mol Med.</i> 2019 Nov;44(5):1888-1898. PMID: 31545417
17.	Glesner J, Kapingidza AB, Godzwon M, Offermann LR, Mueller GA, DeRose EF, Wright P, Richardson CM, Woodfolk JA, Vailes LD, Wunschmann S, London RE, Chapman MD, Ohlin M, Chruszcz M, Pomés A . A Human IgE Antibody Binding Site on Der p 2 for the Design of a Recombinant Allergen for Immunotherapy. <i>J Immunol.</i> 2019 Nov 1;203(9):2545-2556. PMID: 31554696

-
- San Segundo-Acosta P, Oeo-Santos C, Navas A, Jurado A, Villalba M, Barderas R.
Ole e 15 and its human counterpart -PPIA- chimeras reveal an heterogeneous IgE response in olive pollen allergic patients.
Sci Rep. 2019 Oct 21;9(1):15027.
18. PMID: 31636292
-
- Zhang Z, Li XM, Xiao H, Nowak-Wegrzyn A, Zhou P.
IgE-binding epitope mapping of tropomyosin allergen (Exo m 1) from *Exopalaemon modestus*, the freshwater Siberian prawn.
Food Chem. 2020 Mar 30;309:125603.
19. PMID: 31707198
-
- Takashima T, Taku T, Yamanaka T, Fukamizo T, Numata T, Ohnuma T.
Crystal structure and biochemical characterization of CJP38, a β -1,3-glucanase and allergen of *Cryptomeria japonica* pollen.
Mol Immunol. 2019 Dec;116:199-207.
20. PMID: 31731097
-
- Nugraha R, Kamath SD, Johnston E, Karnaneedi S, Ruethers T, Lopata AL.
Conservation Analysis of B-Cell Allergen Epitopes to Predict Clinical Cross-Reactivity Between Shellfish and Inhalant Invertebrate Allergens.
Front Immunol. 2019 Nov 19;10:2676.
21. PMID: 31803189
-
- Múnera M, Martínez D, Labrada A, Caraballo L, Puerta L.
Identification of B Cell Epitopes of Blo t 13 Allergen and Cross-Reactivity with Human Adipocytes and Heart Fatty Acid Binding Proteins.
Int J Mol Sci. 2019 Dec 4;20(24):6107.
22. PMID: 31817065
-
- Bu G, Li T, Zhu T, Xi G.
Identification of the linear immunodominant epitopes in the β subunit of β -conglycinin and preparation of epitope antibodies.
Int J Biol Macromol. 2020 Jul 1;154:724-731.
23. PMID: 32198043
-
- Nesbit JB, Schein CH, Braun BA, Gipson SAY, Cheng H, Hurlburt BK, Maleki SJ.
Epitopes with similar physicochemical properties contribute to cross reactivity between peanut and tree nuts.
Mol Immunol. 2020 May 19;122:223-231.
24. PMID: 32442779
-

表2 平成30年度新たにADFSに追加したエピトープ情報

	Name	start	end	Sequence	Method	CTYPE	Reference	UniProt acc.No
001	Pen c 1	12	25	KLEKDNAMDRADTL	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	D2KMW0
	Pen c 1	25	40	LEQQNKANRAEKSE	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	D2KMW0
	Pen c 1	47	60	QKRMQQLNLDLQV	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	D2KMW0
	Pen c 1	73	85	EKDALSNAEGEV	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	D2KMW0
	Pen c 1	87	105	ALNRRILQLEEDLERSEER	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	D2KMW0
	Pen c 1	112	128	KLAESQAADSEMRK	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	D2KMW0
	Pen c 1	131	147	ENRSLDEERMDALENQ	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	D2KMW0
	Pen c 1	155	167	AEEADRKYDEVAR	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	D2KMW0
	Pen c 1	187	203	ESKIVELEEELRVVGN	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	D2KMW0
	Pen c 1	221	239	YKEQIKTLTNLKAEEARA	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	D2KMW0
	Pen c 1	243	259	ERSVQKLQKEVDRLDE	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	D2KMW0
	Pen c 1	263	280	EKEKYKSITDELDTQFSE	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	D2KMW0
002	Pen c 2	19	33	AATDCKSLKKYLTK	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	Q4KY22
	Pen c 2	55	72	QSGVENLDSGVGIYAPDA	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	Q4KY22
	Pen c 2	91	107	VGFKQTDKHKHPNKDFGDV	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	Q4KY22
	Pen c 2	123	135	TRVRCGRSMEGYP	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	Q4KY22
	Pen c 2	140	154	LTEDQYKEMESKVSS	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	Q4KY22
	Pen c 2	158	169	SLEGELKGYYP	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	Q4KY22
	Pen c 2	250	266	VNEIEKRIPFSHHDRLG	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	Q4KY22
	Pen c 2	286	299	KLPKLAANRDKLEE	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	Q4KY22
	Pen c 2	320	333	GGIYDISNKRMMGL	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	Q4KY22
	Pen c 2	337	356	QAVKEMQDGILELIKMEKEM	Dot assay/indirect competition ELISA/mast cell degranulation assay	L	PMID 29481756	Q4KY22
003	Der p 2			N10, E25, K77, K96, E102	dot-blot/ELISA/PBMC proliferation assay	C	PMID 29467434	P49278
004	Can f 1				dot-blot/ELISA assay	C	PMID 28939849	O18873
005	Der p 5				dot-blot/ELISA assay	C	PMID 29319884	P14004
006	Der p 21				dot-blot/ELISA assay	C	PMID 29319884	Q2L7C5
007	Scy s 1	44	55	ATQKMQQVEN	Phage display/ Dot blotting/ ELISA	L	PMID 30107732	A7L5V2
	Scy s 1	105	112	RLNTATTK	Phage display/ Dot blotting/ ELISA	L	PMID 30107732	A7L5V2
	Scy s 1	133	140	RSLSDEER	Phage display/ Dot blotting/ ELISA	L	PMID 30107732	A7L5V2
	Scy s 1	143	152	ALENQLKEAR	Phage display/ Dot blotting/ ELISA	L	PMID 30107732	A7L5V2

	Scy s 1	199	206	VVGNNLKS	Phage display/ Dot blotting/ ELISA	L	PMID 30107732	A7L5V2
	Scy s 1	253	264	VDRLEDELVNEK	Phage display/ Dot blotting/ ELISA	L	PMID 30107732	A7L5V2
	Scy s 1			R90 ,E164, Y267	Phage display/ Dot blotting/ ELISA	C	PMID 30107732	A7L5V2
008	Pha v ?	55	66	NVNDNGEPTLSS	ELISA/ lymphocyte proliferation/ cytokine profile analyses	L	PMID 30223257	V5QN77
	Pha v ?	116	125	VGSEPKDKGG	ELISA/ lymphocyte proliferation/ cytokine profile analyses	L	PMID 30223257	V5QN77
	Pha v ?	133	141	NNYKYDSNAHT	ELISA/ lymphocyte proliferation/ cytokine profile analyses	L	PMID 30223257	V5QN77
	Pha v ?	149	160	LYNVHWDPKPRH	ELISA/ lymphocyte proliferation/ cytokine profile analyses	L	PMID 30223257	V5QN77
	Pha v ?	95	103	FNIDVPNNS	ELISA/ lymphocyte proliferation/ cytokine profile analyses	L	PMID 30223257	V5QN77
	Pha v ?	39	47	LQRDATVSS	ELISA/ lymphocyte proliferation/ cytokine profile analyses	L	PMID 30223257	V5QN77
009	Gly m 2	21	27	QVVVQTE	Peptide phage display/peptide microarray	L	PMID 30267550	Q07502
010	Der p 5	90	108	DRLMQRKDLDFEQYNLEM	peptide microarray/ alanine scanning mutagenesis	L	PMID 30430936	P14004
011	Der p 24	1	32	MVHLTKTLRFINNPGRKFFYGLQGYNKYGLY	peptide microarray	L	PMID 31139345	A0A0K2GUJ4
012	Can f 6	28	59	DISKISGDWYSILLASDIKEIEENGSMRVFV	ELISA/ alanine scanning mutagenesis	L	PMID 30728436	H2B3G5
013	Cra g 1	44	49	TSLQKK	ELISA/ amino acid substitution	L	PMID 30807831	B7XC66
	Cra g 1	69	85	TKLEEAKTASEAEQEI	ELISA/ amino acid substitution	L	PMID 30807831	B7XC66
	Cra g 1	99	108	MERSEERLQT	ELISA/ amino acid substitution	L	PMID 30807831	B7XC66
	Cra g 1	134	144	NNASEERTDVL	ELISA/ amino acid substitution	L	PMID 30807831	B7XC66
	Cra g 1	209	224	VQNDQASQREDSYEET	ELISA/ amino acid substitution	L	PMID 30807831	B7XC66
014	Ole e 15	44	55	ATQKMQQVEN	Phage display/ Dot blotting/ ELISA	L	PMID 30107732	A7L5V2
	Ole e 15	105	112	RLNTATTK	Phage display/ Dot blotting/ ELISA	L	PMID 30107732	A7L5V2
015	Exo m 1	133	140	RSLSDEER	Phage display/ Dot blotting/ ELISA	L	PMID 30107732	A7L5V2
	Exo m 1	143	152	ALENQLKEAR	Phage display/ Dot blotting/ ELISA	L	PMID 30107732	A7L5V2
	Exo m 1	199	206	VVGNNLKS	Phage display/ Dot blotting/ ELISA	L	PMID 30107732	A7L5V2
	Exo m 1	253	264	VDRLEDELVNEK	Phage display/ Dot blotting/ ELISA	L	PMID 30107732	A7L5V2
	Exo m 1			R90 ,E164, Y267	Phage display/ Dot blotting/ ELISA	C	PMID 30107732	A7L5V2
016	Blo t 13	55	66	NVNDNGEPTLSS	ELISA/ lymphocyte proliferation/ cytokine profile analyses	L	PMID 30223257	V5QN77
	Blo t 13	116	125	VGSEPKDKGG	ELISA/ lymphocyte proliferation/ cytokine profile analyses	L	PMID 30223257	V5QN77
017	Gly m 5	133	141	NNYKYDSNAHT	ELISA/ lymphocyte proliferation/ cytokine profile analyses	L	PMID 30223257	V5QN77
	Gly m 5	149	160	LYNVHWDPKPRH	ELISA/ lymphocyte proliferation/ cytokine profile analyses	L	PMID 30223257	V5QN77
018	Ara h 1	95	103	FNIDVPNNS	ELISA/ lymphocyte proliferation/ cytokine profile analyses	L	PMID 30223257	V5QN77
	Ara h 1	39	47	LQRDATVSS	ELISA/ lymphocyte proliferation/ cytokine profile analyses	L	PMID 30223257	V5QN77
	Ara h 2	21	27	QVVVQTE	Peptide phage display/peptide microarray	L	PMID 30267550	Q07502
	Ara h 7	90	108	DRLMQRKDLDFEQYNLEM	peptide microarray/ alanine scanning mutagenesis	L	PMID 30430936	P14004
	Ara h 11	1	32	MVHLTKTLRFINNPGRKFFYGLQGYNKYGLY	peptide microarray	L	PMID 31139345	A0A0K2GUJ4
019	Der p 2.0103	28	59	DISKISGDWYSILLASDIKEIEENGSMRVFV	ELISA/ alanine scanning mutagenesis	L	PMID 30728436	H2B3G5
020	Der f 24.0101	44	49	TSLQKK	ELISA/ amino acid substitution	L	PMID 30807831	B7XC66

注) start, end: エピトープ配列の始点及び終点アミノ酸の番号

Ctype: エピトープのタイプ. L: linear, C: conformational

人工知能を用いたアレルギー性評価のためのアルゴリズム開発

研究分担者 竹内 一郎（名古屋工業大学）

研究要旨：

ゲノム編集技術などを用いて人工的に生成した食品のアレルギー性を確認する方法は明らかになっていない。アレルギー性の主要な識別子とされる単一の因子は知られておらず、複数の因子が複雑に関連することでアレルギー性を持つことが示唆されている。また、人工的に生成された食品のアレルギー性を都度実験的に検証するのは様々なコストがかかり、現実的でない。そこで、本研究では、人工知能やデータ科学のアプローチを用い、食品のアレルギー性を高精度、高信頼度で汎用的に判定・予測できるシステムを開発することを目指す。これまでに、様々なアレルギー性を判定・予測のための分析ツールが開発されてきた。残念ながら、これら既存の方法には様々な問題点が存在する。国際連合食糧農業機関/世界保健機関によるガイドラインはアミノ酸配列の類似性に基づく基準であり、精度が低く、大規模データの分析には適していない。また、既知の IgE エピトープに基づく基準、タンパク質構造の物理化学的表現に基づく基準、アミノ酸/ジペプチド組成物に基づく基準など、タンパク質に関する生物科学的・物理化学的な知見に基づく単一、もしくは少数の因子を採用したツールが提案されているが、これらはアレルギー物質の多様性を十分に反映できるものとなっていない。本研究では、以下の3つの課題に取り組む：[課題1] 既存のデータベースを拡張し、アレルギータンパク質と非アレルギータンパク質のデータベースを作成する。[課題2] 課題1で作成したデータベースをもとに、アレルギー特異的なパターン（アミノ酸配列）を統計学的手法により抽出する。[課題3] 課題1で作成したデータベースと課題2で抽出したパターンをもとにアレルギー性判定モデルを人工知能・機械学習手法により作成する。課題1に関して、平成30年度は食品種目のアレルギー性、および、非アレルギータンパク質を含むデータベースを構築したが、令和元年度はこれに非食品アレルギータンパク質を追加し、データベースの大規模化、高精度化を行った。令和2年度は新たなデータを追加してデータベースを更新した。課題2に関して、平成30年度はアレルギータンパク質と非アレルギータンパク質それぞれに特異的なパターンを抽出していたが、令和元年度は特定の種や目に限定しないパターンを抽出できるように手法の改良を行った。令和2年度は更新後のデータベースをもとにアレルギーパターンの再抽出を行った。課題3に関して、平成30年度は汎用的な2クラス分類モデルを用いていたが、令和元年度はデータベースの特徴を考慮した本研究課題に特化した機械学習法を開発した。令和2年度は更新後のデータベースをもとにアレルギー性予測システムを再学習し、既存の方法との比較検証のうえで予測精度の観点で提案するシステムの性能がよいことを確認した。

A. 研究目的

ゲノム編集技術を用いた人工的な農産物の合成が行えるようになり、これまでにない食用タンパク質製品が登場する可能性がある。新たに合成された食用タンパク質は未知の特性を有しており、特定の人が摂取するとアレルギー反応が起こってしまうリスクがある。免疫反応においてタンパク質抗原のアミノ酸配列のうち、抗体が結合する部位をエピトープと呼び、エピトープを認識する抗体を人が持っている場合にアレルギー反応を引き

起こされる。これまでの様々な研究から、いくつかのアレルギータンパク質において共通のエピトープ配列が見出されているが、アレルギー性の単一因子は知られておらず、複数の因子が複雑に関連することでアレルギー性を持つことが示唆されている。既存のアレルギー性判定・予測ツールのうちもっとも基本的なアプローチはアレルギー性を持つタンパク質とのアミノ酸配列の類似性（アミノ酸配列相同性）に基づくものである。しかし、このようなアプローチは偽陽性が高いことが指摘

されており、ゲノム編集技術によって合成される新規タンパク質のアレルギー性判定には十分でない。また、別のアプローチとしては、タンパク質に含まれるアミノ酸の物理化学的な特徴の統計量に基づいてアレルゲン性を判定する試みもなされている。このようなアプローチではアミノ酸の順序や位置関係を適切に考慮できないため、十分な精度ではないことが確認されている。アミノ酸配列パターンを用いたアプローチとして、Alledictorと呼ばれる方法が提案されたが、この方法では一定の長さのアミノ酸配列のみを抽出するものであり、すべてのエピトープを網羅できるようなものではない。このような背景のもと、本課題では人工知能や機械学習のアプローチを用い、食品のアレルゲン性を高精度かつ高信頼度で汎用性のあるアレルゲン性判定・予測が行えるシステムを開発することを旨とする。

本研究では、まず、食品タンパク質の大規模データベースを整備し、アレルゲン特異的な様々な長さのアミノ酸配列を抽出し、これらに基づいてアレルゲン性判定・予測システムを構築する。さまざまな数理技術、情報技術を活用することで、高精度で信頼性が高く汎用性のあるアレルゲン性判定・予測システムを開発することを目的とする。

平成30年度は、アレルゲン性判定・予測システムのプロトタイプを作成し、その高精度化、高信頼度化、汎用化に向けた問題抽出を行った。令和元年度は平成30年度のプロトタイプの問題点を列挙し、それぞれを解決するための新たな数理技術、情報技術の開発を行った。令和2年度は、平成30年度に作成したプロトタイプ、令和元年度のシミュレーションで明らかになった課題を解決し、AIによるアレルゲン性判定・予測システムを完成させた。

【課題1】人工知能や機械学習で判定・予測システムを構築するには訓練データベースが必要である。既存のアレルゲン性判定・予測システムで使われていたデータベースはアレルゲンタンパク質のみを用いたものであった。人工知能や機械学習では正例（positive example）だけでなく、負例（negative example）もあると有効なため、後者をデータベースに追加する必要がある。負例の追加では、アレルゲン性とは無関係のタンパク質データベースを取得し、そこからアレルゲン性のあるものを取り除く作業により行った。

本データベースにおいて注意すべき問題は、アレルゲン性タンパク質数（正例数）と非アレルゲン性タンパク質数（負例数）に偏りがあることである。正例は生物学的な実験によって判定されたものであるため数が少なく、負例は通常のタンパク質データベースから大量に取得できる。一方、通常のタンパク質データベースから大量に取得した負例には誤陰性（False Negative）が多く含まれてしまうため、なんらかの対処が必要である。また、正例と負例の数が食物種目ごとにバラつきがある場合、特定の食物種目に特化したアミノ酸配列がアレルゲン性特異的なアミノ酸配列と誤って発見されてしまうリスクが生じる。

平成30年度には11の食品種目のアレルゲンタンパク質と非アレルゲンタンパク質の訓練データベースを作成した。しかしながら、正例数が十分でないため、令和元年度はさらに非食品タンパク質においてアレルゲン性を持つことがわかっているタンパク質を正例として追加した。令和2年度は、新規にアレルゲン性が判明した食品タンパク質をデータベースに増加するとともに、非アレルゲンタンパク質として登録されていたタンパク質のうち、不確実性が高いものをデータベースから取り除いた。

【課題2】人工知能や機械学習でタンパク質の物性を判定・予測するにはタンパク質の特徴を機械学習が使える数値データとして抽出しなくてはならない。生物情報学で採用されているアプローチとして主に2通りのものがある。1つ目のアプローチは、タンパク質を構成するアミノ酸の物理化学的な特徴（疎水性、分子量など）を求め、その平均、分散、相関などを特徴として抽出することである。2つ目のアプローチは、アミノ酸の部分配列のうち、特定の物性を有するタンパク質に特化して頻出する部分配列を特徴として抽出することである。アプローチ1ではアミノ酸の順序や位置を考慮できないため、本研究ではアプローチ2を採用する。

また、一般に、機械学習における特徴抽出は、教師なし学習と教師あり学習の2つのアプローチが存在する。本研究においては、前者はアレルゲン性タンパク質の情報のみから特徴抽出を行うことに相当し、既存のアレルゲン性判定・予測システムの多くではこのアプローチを採用されている。

本研究では、より判定・予測に有用な特徴を抽

出するため、教師あり特徴抽出のアプローチを採用した。平成30年度には食品タンパク質のみを扱っていたため、我々のグループが別の目的で既に確立した方法をそのまま適用することができた。令和元年度は非食品のアレルゲン性タンパク質を正例として追加したが、そのため対処が必要となった。これは、既存の教師あり特徴抽出法を用いると、特定の非食品タンパク質に特化したアミノ酸部分配列がアレルゲン性特異的アミノ酸配列として誤って抽出されてしまうためである。令和2年度は、課題1で更新されたデータベースをもとにパターンの再抽出を行った。さらに抽出したパターンと既知のエピトープとの一致性を検証した。

【課題3】 正例と負例を含む訓練データベースを用いて、正負が未知の事例を判定・予測する問題は教師あり学習 (supervised learning) と呼ばれている。アレルゲン性タンパク質を正例、非アレルゲン性タンパク質を負例とみなせば、典型的な教師あり学習問題と解釈できるが、いくつか本課題特有の課題を解決する必要がある。まず、本課題の1つ目の特徴は訓練データベースに含まれるタンパク質が独立同一分布 (i.i.d.; independently, identically distributed) に従わない点である。この場合、通常の教師あり学習で多用されるクロスバリデーションなどのリサンプリング法をそのまま利用することができず様々な工夫が必要となる。また、正例数と負例数に偏りが生じてしまう点も本課題の特徴であり、注意深く対処する必要がある。本研究で用いるデータベースにおいて、食品タンパク質に関しては正例が負例に比べて極端に少なくなってしまうしており、非食品タンパク質に関しては正例のみが存在する状況になっている。

また、アレルゲン性の原因となるエピトープはさまざまな長さであることが知られているため、さまざまな長さのアミノ酸部分系列特徴を抽出できるような工夫が必要である。さらに、アレルゲン性の判定は統計的信頼性が担保されたものである必要があるため、抽出された特徴的信頼性定量化を行う必要がある。加えて、特定の食品種目に特化したものでなく、一般的な特徴を抽出するための工夫が必要である。

平成30年度では、訓練データベースが非IIDである点と食品タンパク質における正例と負例の偏

りを考慮したモデル作成法を構築した。令和元年度では、さらに非食品タンパク質を訓練データベースに追加した際の対処法を検討した。令和2年度は、課題2で更新されたパターンをもとに予測モデルの再学習を行った。既存のアレルゲン性予測システムとの比較を行い、提案法の予測精度が高いことを確認した。

B. 研究方法

課題1の訓練データベースの構築においては、アレルゲン性を持つ食品タンパク質の正例としてCOMPAREデータベースのものを利用した。同じくアレルゲン性のない食品タンパク質の負例としてUniProtデータベースより取得した。UniProtデータベースは汎用的なタンパク質データベースであるため、アレルゲン性を持つものも含まれている。そのため、既存のエピトープを含むもの、アレルゲンに関連するキーワードが付記されているものなどを削除した。またプロトタイプとして作成したアレルゲン性判定・予測システムにおいて偽陽性であったタンパク質に関して個別にデータベースを精査し、アレルゲン性を持つ可能性があるものは削除するなどの措置をとった。後述のように、課題2、3においては食品種目の情報を活用するため、食品種目分類の精査を行い、あいまい性のあるタンパク質はデータベースから削除するプロセスを行った。その他にもプロトタイプシステムや諸々のタンパク質データベースを活用することで訓練データベースの大規模化と高精度化を実現した。上述のように、本データベースに含まれる事例 (タンパク質) は独立同一分布 (IID) に従わないので、食品種目ごとにデータ分割を行うLeave-Food-Outクロスバリデーションと呼ぶ方法に基づいてデータ分析を実施した。

平成30年度では、データベースが食品タンパク質のみから構成されていたが、令和元年度には非食品タンパク質も追加した。なお、非食品タンパク質でアレルゲン性のないものを網羅的に収集するのは困難であることが判明したため、本研究では、非食品タンパク質に関しては、アレルゲン性を有する正例のみを扱うこととした。令和2年度は、本研究開始後に新たにアレルゲンとして登録されたアレルゲン性食品タンパク質をリストアップし、これをデータベースに追加した。一般に、タンパク質が非アレルゲン性であることを完全に確定するのは困難であるため、非アレルゲン性タ

ンパク質としてデータベースに登録されているもののうち、False Negative が存在する可能性が排除できない。令和2年度は、非アレルギー性タンパク質を精査し、不確実性の高いものを削除してデータベースの信頼性を向上させた。

課題2の特徴抽出においては、本研究に特化したさまざまな工夫を行った。まず、異なる長さのアミノ酸部分配列を抽出できるようにするため、我々が開発したデータマイニング分野の技術を利用した。系列データから特定の性質を持つ部分系列を抽出する技術は系列マイニングと呼ばれ、さまざまな方法が提案されている。系列マイニングでは、系列を木構造と呼ばれるデータ構造で表現し、枝刈りと呼ばれる手順を導入することにより、膨大な部分系列から、特定の性質を満たすものを探索することができる。本研究の基本的な方針は、アレルギー性タンパク質に高頻度で含まれ、非アレルギー性タンパク質には低頻度でしか含まれない（あるいはまったく含まれない）ような部分配列を探索することである。頻度の違いを定量化する指標には様々なものがあるが、本研究ではフィッシャーの正確検定 (Fisher Exact Test) に基づく指標を利用した。例えば、20種類のアミノ酸において長さ10までのアミノ酸の種類は10の20乗となり、その頻度を数えたデータテーブルを作るとは実質的に不可能である。

我々は、系列マイニングにおける木構造の枝刈りをフィッシャーの正確検定と統合する方法を開発した。詳細は割愛するが、この方法では、統計的に有意となり得ない部分配列を木構造の枝刈りによって排除できるため、膨大な数の候補から予測に最適な部分配列を選択することができる。また、アレルギー性予測モデルの信頼性を高めるため、統計的な有意性を持つ部分配列のみを用いることが望ましい。ある部分配列の出現頻度がアレルギー性タンパク質と非アレルギー性タンパク質で異なるかどうかの統計的検定を行う場合、フィッシャーの正確検定の p 値を利用することができる。しかしながら、膨大な部分系列の候補のなかから特に頻度の違いの大きなものを抽出してきた場合、選択バイアスが生じてしまい、所望の誤検出率を制御できなくなる。この選択バイアスの問題は多重検定問題 (multiple hypothesis testing) と呼ばれており、その補正を行うためにはフィッシャーの正確検定によって得られた p 値を適切に補

正しなくてはならない。もっともよく使われている多重検定補正にボンフェローニ補正 (Bonferroni correction) と呼ばれるものがあるが、選択における候補数が多い場合、補正が保守的になってしまう問題点が指摘されている。本研究ではこの問題に対処するため、Westfall Young法と呼ばれるランダム化に基づく方法を採用した。これらの方法法の開発と本データベースへの適用は主に平成30年度に行ったが、令和元年度もアルゴリズムの改良や新たなデータへの適用などを行った。

令和元年度は、主に、非食品タンパク質においてはアレルギー性を持つ正例のみがデータベースに含まれる点を考慮して特徴抽出を行った。この点を特に考慮せずに通常の機械学習アルゴリズムを適用すると、アレルギー特異的でなく、非食品タンパク特異的なパターンが誤って検出されてしまう。この問題を回避するため、アレルギー特異的なパターンとして、条件1) 食品タンパク質に含まれるか、条件2) 非食品タンパク質のうち複数の目に含まれる、のどちらかの条件を満たすもののみを抽出することとした。

令和2年度には、諸々のタンパク質データベースを活用し、非食品タンパク質でアレルギー性を持たないものをデータベースに加えることができないか検討を進めた。また、令和元年度に引き続き、パターン同定アルゴリズムの高速化を行った。さらに同定されたパターンと既存のエピトープとの一致度の検証を行った。

課題3のアレルギー性判定・予測システムの構築は上述の Leave-Food-Out クロスバリデーションを利用した教師あり学習によって行った。アミノ酸部分配列パターンを特徴として抽出したため、テスト対象のタンパク質がパターンを含むか否かをバイナリ表現した線形分類器をベース手法として採用した。パターン数が多いと解釈性が低く過学習のリスクがあるため、スパース正則化や二次正則化 (Ridge Regression) を導入した。平成30年度は主にこのプロトタイプモデルに基づく考察を行った。令和元年度は、さらに、パターンが完全に含まれる (exact match) だけでなく、パターンが部分的に類似している場合 (non-exact match) も考慮できるような工夫を導入した。20種類のアミノ酸の物理化学的な特徴に基づいてアミノ酸種間の類似度を定義し、タンパク質にパターン

が含まれる程度を連続量として定量化した。加えて、令和元年度は、抽出されたパターンの生物学的な考察として、既存のエピトープとの一致度の確認や、結合性の確認なども行った。令和2年度は、アレルゲン性予測システムの性能を定量評価するため、既存の方法である Alledictor 法、MEME 法、Allertop 法と比較し、本研究で構築した方法の予測精度が高いことを確認した。

C. 研究結果および考察

データ駆動型アプローチによってアレルゲン要因を同定する際、アミノ酸配列が、アレルゲン性特異的なものであるのか、種（カテゴリ）特異的なものであるのかを判断するには注意が必要である。本研究では、データベースに含まれるカテゴリを、(a) ペアカテゴリ（アレルゲン性タンパク質と非アレルゲン性タンパク質の両者がデータベースに含まれているようなカテゴリ）、(b) アレルゲンオンリーカテゴリ（アレルゲン性タンパク質のみがデータベースに含まれているようなカテゴリ）、(c) 非アレルゲンオンリーカテゴリ（アレルゲン性タンパク質のみがデータベースに含まれているようなカテゴリ）に区別して処理することでこの問題を解決した。本研究におけるアレルギー特異的パターンの定義として、次の3つの要件、（要件1）ペアカテゴリに含まれる、もしくは、複数のアレルゲンオンリーカテゴリに含まれる、（要件2）非アレルゲンタンパク質には一切含まれない、（要件3）統計的に有意である、を考え、このような性質を持つパターンを膨大な候補の中から漏れなく効率的に同定するための方法を開発した。実際、この方法によって同定されたアレルゲン特異的パターンは既知のエピトープと一致しているものが多く、アレルゲンの要因となるものを正しく発見できていることが示唆される。また、アレルゲン特異的パターンを用いてアレルゲン性判定・予測システムを構築したことで、解釈性と信頼性の高い予測システムの構築が可能となった。

D. 結論と今後の展望

本研究で作成したアレルゲン性判定・予測システムは、統計的に信頼性の高い情報のみを用いて構成されているため、予測の説明・解釈が容易であり、信頼性が高いという利点を持つ。本研究で作成したアレルゲン性判定・予測システムは既存

のものより予測精度が高いことを確認した。今後は、データベースの更新とそれに伴う判定・予測システムの更新のプロトコル化し、さらなる改善を目指す。

E. 研究発表・業績

1) 論文発表

1. Karasuyama M., Inoue K., Nakamura R., Kandori H., Takeuchi I. Understanding colour tuning rules and predicting absorption wavelengths of microbial rhodopsins by data-driven machine-learning approach. *Scientific Reports*, **8**, paper no.15580, 2018.
2. Yoshida T., Takeuchi I., Karasuyama M. Safe Triplet Screening for Distance Metric Learning. *Neural Computation*, **31**(12), 2432-2491, 2019.
3. Umezu Y., Takeuchi I. Selective inference via marginal screening for high dimensional classification. *Japanese Journal of Statistics and Data Science*, **2**(2), 559-589, 2019.
4. Inatsu Y., Karasuyama M., Inoue K., Takeuchi I. Active learning for level set estimation under input uncertainty and its extensions. *Neural Computation*. **32**, 2486-2531. 2020

2) 学会発表

1. Yoshida T., Takeuchi I., Karasuyama M. Safe Triplet Screening for Distance Metric Learning. Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD2018), 2018.
2. Ndiaye E, Takeuchi I. Computing Full Conformal Prediction Set with Approximate Homotopy. Advances in Neural Information Processing Systems (NeurIPS2019), 2019.
3. Ndiaye E, Le T., Fercoq O., Salmon J., Takeuchi I. Safe Grid Search with Optimal Complexity. International Conference on Machine Learning (ICML2019), 2019.
4. Tanizaki K., Hashimoto N., Inatsu Y., Hontani

H., Takeuchi I. Computing Valid P-values for Image Segmentation by Selective Inference. IEEE/CVF Conference on Computer Vision and Pattern Recognition 2020 (CVPR2020), 2020.

5. Duy V.N.L., Toda H., Sugiyama R., Takeuchi I. Computing Valid p-value for Optimal Changeoint by Selective Inference using Dynamic Programming. The 34th Conference on Neural Information Processing Systems (NeurIPS2020), 2020.

F. 健康危険情報

各年度の分担報告書を参照

G. 知的財産権の出願・登録状況

各年度の分担報告書を参照

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」
総合分担研究報告書

ゲノム編集生物作製における現象解析と規制の進め方

研究分担者 木下 政人（京都大学）

研究要旨：

我々は、これまでにゲノム編集技術（CRISPR/Cas9）を用いて、増肉形質を示すミオスタチン遺伝子破壊マダイ、高成長形質を示す食欲関連遺伝子破壊トラフグシステムの作製に成功している。本研究では、これらのゲノム編集養殖魚を用いて、非ゲノム編集個体との性状の相違を検討した。1）アレルギー性の検討：検討した結果、編集様式（欠失塩基数やその位置）により、既知のアレルゲン物質との相同性が異なることが明らかとなったが、既存のアレルゲンと相同性の高い配列は確認されなかった。2）メタボロミクス解析：食欲関連遺伝子破壊トラフグを用いて、筋肉に存在する代謝産物の網羅的比較を行った。その結果、エネルギー代謝に関わる物質にわずかに違いが見られたが、非編集魚と有意に異なる点は見つからなかった。3）ゲノム解析：食欲関連遺伝子破壊トラフグでのゲノム解析を行なった結果、オフターゲット変異はされず、また、ゲノム編集ツールの残存も確認されなかったことから、ゲノム配列に非編集魚と有意に異なる点は見つからなかった。4）フグ毒蓄積様式の検討：食欲関連遺伝子破壊トラフグを用いて、フグ毒（テトロドトキシン：TTX）の体内分布様式を検討した結果、ゲノム編集魚および非編集魚いずれにおいても筋肉、皮、精巣には蓄積されず、肝臓と卵巣に蓄積された。この知見は従来のトラフグのフグ毒含有部位と同一であった。5）ラットを用いた投与試験：レプチン受容体遺伝子機能欠失トラフグの骨格筋ホモジネートをラットに2週間経口投与した結果、外見的所見、血液学的検査、血液生化学的検査の結果では異常は認められず、ゲノム編集トラフグの可食部である骨格筋に毒性は見られなかった。

以上の結果から、ゲノム編集魚は非編集魚と同等の安全性が示された。また、アレルギー性の検討のところで観察されたように、同一遺伝子を編集したものであっても、各編集タイプ（欠失塩基数やその位置）において特性が異なることから、安全性は作出された系統を個別に評価することが必要であると考えられた。

A. 研究目的

近年急速に発展してきたゲノム編集技術は、「生物種を選ばず、ゲノム上の狙った配列を改変できる」という特性から、農林水産物の育種に応用され始めている。この技術では、「短期間で狙った形質を持つ品種」を作製することが可能であるため、本技術は品種改良方法・育種方法として、今後定着していくものと考えられる。

しかしながら、本技術は歴史が浅いため、この技術で作製された食品に対して、消費者が安全性に不安を持っているのが現状である。加えて、安全性を確認する機関においても安全性への具体的な評価基準の策定に至っていない。

水産物は、これまでの「とる漁業が中心」の時代から、「作る・育てる漁業」へと変わってきたものの、品種の作製や育種が、作物や畜産物に比べて

大幅に遅れている。一方、世界的な人口増加と健康食志向の高まりから養殖業が発展してきており、消費者ニーズに合った水産物の作出が望まれるようになってきた。このような背景から、短期間で優良形質を固定化できるゲノム編集技術の導入が水産物育種に活用され始めた。今後、多様な水産物においてゲノム編集技術による新品種・新食品が作製されていくものと思われる。

そこで本研究では、自身で作出したゲノム編集マダイおよびトラフグを用い、それらの特性を検討し、得られた結果をもとに今後の食品安全性評価法策定に提言を与えるのを目的とした。

B. 研究方法

1）アレルギー性の検討：ゲノム編集によりミオスタチン遺伝子（*mstn*）を破壊したマダイ3系統

およびトラフグ3系統、レプチン受容体遺伝子 (*lepr*) を破壊したトラフグ2系統、メラノコルチン4型受容体遺伝子 (*mc4r*) を破壊したトラフグ1系統について、予想される全アミノ酸配列、新生ペプチドとその直上10アミノ酸部分、および、塩基欠失部位を挟んだ両側の終止コドン内で予想されるペプチドを用い、web上のアレルゲン検索サイト(後述)により、それらのアレルゲン性を検討した。平成31(令和元)年度は通常の読み枠(フレーム)に加えて、異なる読み枠および逆鎖でのアミノ酸配列においても評価した。アレルゲンとする基準は、①全長のアミノ酸配列において E.value < 1 の相同性を示す、②80アミノ酸のウィンドウサイズで35%以上の相同性を示す、③8アミノ酸配列が完全に一致する、ものとした。「FAOでの6アミノ酸の相同性評価」は過大評価をしているとの指摘があるため、Allergen Online の評価基準である「8アミノ酸の相同性」により評価した。また、E.value < 1.0 となったものについては、より低値での解析を行った。

アレルゲン検索サイト

Allergen Online (University of Nebraska-Lincoln)

Allergen Database for Food Safety (ADFS) (国立医薬品食品衛生研究所)

2) メタボロミクス解析：レプチン受容体遺伝子破壊トラフグ(4塩基欠失系統)の第2世代と非ゲノム編集トラフグ(野生型)(それぞれ19ヶ月齢の個体を3個体ずつ)の背部骨格筋を採取し、キャピラリー電気泳動飛行時間型質量分析装置(Capillary Electrophoresis Time-of-flight Mass Spectrometry: CE-TOFMS)を用いて、低分子代謝産物を網羅的に解析した(ヒューマン・メタボローム・テクノロジーズ株式会社に委託)。また、ゲノム編集によりミオスタチン遺伝子を破壊したマダイ系統(14塩基欠失)およびレプチン受容体遺伝子を破壊したトラフグ系統(4塩基欠失)の背部骨格筋を採取し、凍結乾燥後、粉末化した。この粉末50mgにメタノール/クロロホルム/水(2.5/1/1)を加えた後、水溶性画分を回収後、乾留・TMS誘導化を行い、GC-MS(GSMS-QP2010 Ultra)にて解析を行った。分析により得られた水溶性一次代謝物を網羅的に検出し、多変量解析ソフトSIMAにより解析した。

3) ゲノム解析：レプチン受容体破壊トラフグ第3世代の4個体、および、野生型4個体(雌雄2個体ずつ)それぞれから、ゲノムDNAを抽出して全ゲノム配列を解読し、Integrative Genomics Viewer (IGV)による解析を行った。トラフグゲノム配列データベースから、レプチン受容体中のターゲット配列(CCACTGTGTGCTGTTCCATCT:太字がPAM配列)と欠失および挿入を含め mismatches が2塩基以内の配列をオフターゲット候補とし、同領域におけるゲノム編集魚と野生型魚における塩基配列の変化を検討した。加えてPAM配列近傍のシード配列が完全に一致している領域については、PCRで増幅後、塩基配列解析を行い、変異の有無を検討した。

4) フグ毒蓄積様式の検討：ゲノム編集によりレプチン受容体遺伝子、メラノコルチン4型受容体遺伝子を破壊したトラフグおよび非編集トラフグに市販のテトロドトキシン(和光TTX206-11071)を0.4mg/Kgとなるように単回経口投与した。投与24時間後に即殺、解剖し、皮、筋肉、肝臓、生殖巣を摘出して使用するまで冷凍保存した。解凍後、細切したのち1gに0.1%酢酸4mLを添加し100°Cで10分間加熱後、冷却、遠心分離し、上清をフィルターろ過してTTXの抽出を行なった。抽出液をHPLC-FLDに供することでTTXを蛍光検出器(励起波長381nm、検出波長505nm)で定量した。

5) ラットを用いた投与試験：レプチン受容体を破壊したトラフグ4個体、および、非ゲノム編集トラフグ4個体(雌雄2個体ずつ)それぞれから、背部普通筋を採取し、1.5倍量のミリQ水を用いてホモジナイズした後、目開き0.42mmのメッシュを通し、筋肉ホモジネートとした。これをラット(Slc:SD系統、SPFグレード、雄、6週齢)に各トラフグホモジネートにつき2匹に、ゾンデを用いて胃内に1日1回注入した。これを14日間行なったのち、血液を採取し、血液学的検査および血液学的検査を行なった。また、試験前中後での各個体の体重、試験期間後の肝臓および腎臓の重量測定を行った。試験期間を通じて外見の観察を行った。なお、ラットへの投与試験は株式会社ケー・エー・シーに委託し実施した。

C. 結果および考察

1) アレルゲン性の検討：マダイのミオスタチンタンパク質には、野生型においても、全長アミノ酸配列でアレルゲン性を示す配列（グリアジンと相同性を示す）が存在したが、それ以外にはアレルゲン性を示唆する配列は検出されなかった。また、新生アミノ酸とその上流10アミノ酸配列中の8アミノ酸配列、および、別フレームではアレルゲン性が予測される配列は検出されなかった。野生型トラフグとゲノム編集トラフグのミオスタチン遺伝子から産生されるアミノ酸配列を図1に示す。ミオスタチン遺伝子ゲノム編集トラフグの全長を解析した結果、8塩基欠失(-8b)系統においてアオカビのタンパク質と相同性を示す配列が検出されたが、その他の系統ではアレルゲン性を示す配列は検出されなかった(図2)。通常とは異なるフレームあるいは、相補鎖について検討した結果、各系統において、アレルゲン性を示す配列が検出されたが、いずれも E.value < 0.05 ではアレルゲン性は示されなかった。

野生型トラフグとゲノム編集トラフグのレプチン受容体遺伝子から産生されるアミノ酸配列を検討した結果、通常フレームで全長アミノ酸を検討した場合、2塩基および4塩基欠失のいずれの系統においてもイネの α アミラーゼと相同性を示す配列が検出された。しかしながら、事実上アレルゲンとして有効であると考えられる E.value < 0.05 で再評価した場合には相同性は検出されなかった。センス鎖の1塩基シフトしたフレームでは、いずれのゲノム編集系統でもオレオシンの相同性が示されたが、E.value < 0.05 ではアレルゲン性を示さなかった。他のフレームおよび相補鎖のすべてのフレームではアレルゲン性は検出されなかった。

野生型トラフグとゲノム編集トラフグのメラノコルチン4型受容体遺伝子から産生されるアミノ酸配列を評価した結果、センス鎖の通常フレームではアレルゲン性は示されなかった。一方、通常ではない読枠ではアレルゲン性を示すものも有った。中でもネツタイシマカの唾液腺中アレルゲンと E.value = 0.0078 でアレルゲン性を示すものがあった。

2) メタボロミクス解析：レプチン受容体遺伝子破壊(4塩基欠失)トラフグおよび野生型トラフグの背部骨格筋に含まれる低分子代謝産物の網羅

的解析を CE-TOFMS を用いて行った。その主成分分析の結果を図3に、パスウェイ解析の結果を図4に示す。主成分分析では、第一主成分(PC1)により、レプチン受容体遺伝子破壊個体と野生型個体が区別された。この第1主成分の負荷因子は、正の寄与因子として寄与率順に、Betonicine (trans-4-hydroxy-L-proline betaine)、セリン、グリシン、コハク酸があり、また、負の寄与因子として寄与率順に、5-ヒドロキシリジン、N6-メチルリジン、リジン、アミノアセトンが示された。パスウェイマップ上での、これらの因子の分布は系統だっておらず、特定のパスウェイがゲノム編集操作(レプチン受容体遺伝子破壊)によって大きく影響を受けたことはないと判断された。

また、レプチン受容体遺伝子破壊トラフグを GC-MS (GSMS-QP2010 Ultra) にて分析を行った。レプチン受容体遺伝子破壊トラフグの結果を図5に示す。Q2値は0.5以下であったため、主成分分析の結果の信頼性は高くないが、レプチン受容体ゲノム編集トラフグと野生型トラフグ間で成分の差異があることが示唆された。そのため、最も差異を作り出す条件で解析を行った結果、グリシンが増加し、リジンが減少していることが推察された。

野生型マダイとミオスタチン遺伝子破壊マダイの GC-MS (GSMS-QP2010 Ultra) 解析によるメタボロミクスを比較した結果を図6に示す。主成分分析(図6-b)の信頼性を示す Q2値が0.5以下であり(図6-a)、主成分分析解析の信頼性は低いと判断される。つまり、比較した個体間で有意な差がないと判断された。

3) ゲノム解析：レプチン受容体遺伝子破壊トラフグと非編集トラフグの全ゲノム配列解読は、各個体でゲノムサイズの30倍以上のデータを取得した。塩基配列データ検索の結果、ターゲット配列と2塩基以下のミスマッチ(欠失・挿入を含む)を有する配列は61箇所存在した。図7に全塩基配列解析の結果をIGVで可視化した一例を示す。図7内の拡大図に示すように、他個体と比較し塩基配列に違い(欠失)が存在すれば白抜きで示される。図7ではターゲット配列領域を表示しているため、その中の同じ4塩基がゲノム編集魚でのみ欠失している。ゲノム編集魚に特異的にみられる欠失領域は、標的領域のみであった。つまり、オフターゲット影響は観察されなかった。また、異なる

る手法でのオフターゲットの検証として、候補領域を PCR により増幅後、電気泳動によるバンドシフトの観察と塩基配列解析を行った(図8)。その結果、いずれの領域にも、変異は観察されなかった。

図9に Cas9 RNA、図1に guideRNA の残存性を示す。いずれにおいてもゲノム編集処理時に用いた RNA 配列は確認されなかった。一方、プラスミドのバックボーン配列(RNA に転写しておらず、ゲノム編集処理には用いていない配列)の一部に、ゲノム編集魚および野生型魚の両方で観察された。これは、トラフグゲノム内に一般的に侵入した細菌などの断片であると思われる。

4) フグ毒蓄積様式の検討：ゲノム編集によりレプチン受容体遺伝子を破壊したトラフグ(ゲノム編集トラフグ)および非編集トラフグの各組織における TTX の蓄積量を、投与した TTX 量に対する比率(%)で示す(表1)。TTX を投与しない場合は、ゲノム編集魚および非編集魚のいずれにおいても検討した全ての組織で TTX の蓄積は認められなかった。このことからトラフグ自身が TTX を合成しないことが改めて示された。ゲノム編集魚において、皮、筋肉、雄の生殖巣(精巣)では、TTX 量は、検出限界(0.2 $\mu\text{g/g}$ 組織)以下であった。一方、個体間に大きなばらつきはあるが肝臓および雌の生殖巣(卵巣)では、顕著な TTX の蓄積が見られた。この結果は、非編集魚でも同様であった。トラフグのフグ毒蓄積には、サキシトキシン・テトロドトキシン結合タンパク質(Puffer-fish Saxitoxin Tetrodotoxin Binding Protein: PSTBP)が関与していることが報告されている。今回のゲノム編集のターゲット遺伝子は、食欲に関係する遺伝子であり、PSTBP の代謝に関与するものではない。このことから、TTX 蓄積に関与しない遺伝子をゲノム編集することでは、トラフグ体内での TTX 蓄積部位は変化せず、非編集魚と同様であることが示された。

5) ラットを用いた投与試験：試験期間中のゲノム編集トラフグホモジネート、または、非編集トラフグホモジネートを投与したラット全てにおいて一般状態、および投与試験終了後の解剖時の肝臓および腎臓を肉眼的観察に異常所見はみられなかった。

体重の増加においても、ゲノム編集トラフグホ

モジネート、または、非編集トラフグホモジネートを投与群ともに、実験開始時から 55%~69% の増加を示し、両者に差は見られなかった。

血液学的検査結果を表2に示す。赤血球数においてゲノム編集トラフグホモジネート投与群では $7.15 \times 10^6 \mu\text{L}$ 、非編集トラフグホモジネート投与群では $6.87 \times 10^6 \mu\text{L}$ となり統計的な有意差が示されたが、ヘモグロビン濃度やヘマトクリット値など関連するパラメーターに優位な差が見られなかったこと、他の試験での対照群の数値とゲノム編集トラフグホモジネート投与群の方が近いこと、などから事実上、赤血球数には差異がないと判断された。白血球数、血小板数などその他の指標では優位差は示されなかった。

血液生化学的検査において、ゲノム編集トラフグホモジネート投与群のアスパラギン酸アミノ基転移酵素(ASAT)は非編集トラフグホモジネート投与群と比較して統計学的に有意な高値を示した。しかし、ASAT の変動幅は正常の範囲内であり、その他の肝機能の指標となる酵素には変化が見られないことから、ASAT の上昇は肝機能障害を示唆する結果ではないと判断した。血液生化学的検査のその他の項目については、有意差は見られなかった。

以上の結果から、ゲノム編集トラフグ筋肉ホモジネートに毒性はないと判断した。

D. 結論

- ・ アレルゲン性の検討項目で、同じ遺伝子を編集しても、その編集の様式(欠失延期数や部位)により、新規に産生される可能性があるペプチドは異なり、かつ、そのアレルゲン性についても異なることが示された。このことより、ゲノム編集食品の安全性を評価する際、作出された系統ごとにその安全性を評価する必要があることが示された。
- ・ アレルゲン性を評価する際、基準値(E.value や評価する連続塩基数など)の設定により、結果が異なる。そのため真に有効な基準値の検討と設定が求められる。
- ・ メタボロミクス解析の結果、レプチン受容体遺伝子編集トラフグにおいて特定の代謝経路の亢進や減退が見られず、個体としての代謝経路が大きく変化することは観察されなかった。また、フグ毒(TTX)蓄積様式にも変化が見られなかった。これらのことから、編集を施した遺伝子

と直接的な因果関係のない代謝経路に及ぼすゲノム編集の影響は非常に低いと考えられた。

- ・今回用いたレプチン受容体遺伝子ゲノム編集トラフグにおいてラットへの経口投与で毒性が見られなかったこと、フグ毒投与試験でフグ毒を投与しなかったゲノム編集トラフグにフグ毒が検出されなかったこと、また、これまで分担者が作出したミオスタチン遺伝子ゲノム編集マダイでも食味試験において検査者に異常は見られなかったことなどから、ゲノム編集操作によって有害な物質が産生されるものではないことが示された。
- ・全ゲノム解析の結果、ゲノム編集ツールの残存やオフターゲットは検出されなかった。またPCR法によってもオフターゲットは観察されなかった。この様にオフターゲットの発生頻度は大変低いと考えられる。
- ・動物のゲノム編集で一般的であるRNAやタンパク質を用いたゲノム編集方法では、それらの残存せいの可能性は非常に低いと考えられる。

E. 研究発表・業績

1. 論文発表

なし

2. 学会発表

なし

3. 講演会・説明会等の社会貢献

1) 市民向け説明会

- ・平成30年度科学技術コミュニケーション推進事業 未来共創イノベーション活動支援「共に考えるゲノム編集の未来」、2018年9月18日、奈良市立一条高校、合計約1,200名、奈良市立一条高校とJST、プレゼンター
- ・バイオカフェ in 愛知県図書館、2018年10月21日、愛知県立図書館、合計約30名、愛知県図書館、プレゼンター
- ・河内長野市民大学 くらまろ塾本部講座、2018年10月31日、河内長野市立市民交流センター、合計約32名、河内長野市、プレゼンター
- ・平成30年度 広島バイオフィォーラム、2018年11月15日、サテライトキャンパスひろしま、合計約60名、広島バイオテクノロジー推進協議会、講演者
- ・元村有希子のサイエンスカフェ、2018年11月

26日、毎日メディアカフェ、合計約30名、毎日新聞社、講演者

- ・シンポジウム「ゲノム編集と食の安全・安心」2019年2月10日、日比谷コンベンションホール、合計約200名、たねと食とひと@フォーラム、パネリスト
- ・千里ライフサイエンスフォーラム、2019年2月14日、千里ライフサイエンスセンタービル会議室、合計約100名、千里ライフサイエンス振興財団、講師
- ・ゲノム編集生物と社会について考える、2019年7月6日、日本学術会議講堂、約150名、日本学術会議農学委員会・食料科学委員会合同遺伝子組換え作物分科会、講演者
- ・農林水産省アウトリーチ事業、2019年11月13日、名古屋大学農学部、約70名、農林水産省、講師
- ・農林水産省アウトリーチ事業、2019年11月22日、立命館高校、約20名、農林水産省、講師
- ・農林水産省アウトリーチ事業、2019年11月22日、立命大学、約20名、農林水産省、講師
- ・ステークホルダー会議、2019年12月5日、京都テルサ、約50名、京都生活協同組合、講師
- ・養殖業へのゲノム編集技術の可能性、2020年8月21日、アイビーホール青山会館、参加者約50名、一般財団法人日本水産油脂協会、講演者
- ・食べる？食べない？ゲノム編集マダイ・サイエンスアゴラ 2020、2020年11月22日、Online、参加者数約100名、国立研究開発法人科学技術振興機構、講演者

2) 業界関係者向け説明会

- ・水産ジャーナリストの会、2018年10月24日、三會堂会館、合計約20名、水産ジャーナリストの会、講演者
- ・食生活ジャーナリストの会、2018年11月17日、東京ウィメンズプラザ、合計約80名、食生活ジャーナリストの会、講演者
- ・知的財産セミナー、2019年6月29日、京都産業会館、約30名、日本弁理士会関西会京都地区会、招待講演者
- ・ワークショップ：「ゲノム編集食品の安全・安心」、2019年7月17日、大阪大学東京オフィス、約20名、一般社団法人ゲノム編集学会、プレゼンター、話題提供者

- ・ アクションプラン 2019、2019年7月31日、大阪帝国ホテル、約300名、三菱食品、招待講演者
 - ・ 第63回 滋賀県学校保健・安全研究大会、2019年10月10日、近江八幡市 男女共同参画センター、約60名、滋賀県教育委員会
 - ・ トランスフォーマティブ化学生命融合研究大学院プログラムセミナー、2019年10月15日、名古屋大学農学部、約70名、名古屋大学、招待講演者
 - ・ 養殖情報交換会、2019年12月10日、シンフォニアテクノロジー響ホール伊勢、約50名、国立研究開発法人水産研究・教育機構 増養殖研究所、招待講演者
 - ・ JBA 発行と代謝講演会シンポジウム、2020年2月17日、東京大学・中島董一郎記念ホール、80名、バイオインダストリー協会、招待講演
 - ・ ワークショップ「ゲノム編集食品に関する多様な意見をどう取り上げるか?」、2020年2月25日、メルパルク京都、20名、ゲノム編集の未来を考える会 (JST 来共創イノベーション活動支援)、招待講演
 - ・ ゲノム編集マダイ・スタークホルダー会議、2020年12月3日、東海コープ事業連合 (本部)、参加者約50名、くらしとバイオ 21・東海コープ事業連合、講演者
- ・ 第31回日本生命倫理学会年次大会 シンポジウム、2019年11月29日、東北大学

5) 総説など

- ・ 木下政人 「肉厚マダイ」の事例から考える養殖業へのゲノム編集技術の活用、月刊養殖ビジネス (緑書房)、1月号 pp61-64、2020
- ・ 木下政人 ゲノム編集技術を使った肉厚マダイの作出と品種改良期間の短縮、JATAFF ジャーナル (農林水産・食品産業技術振興協会)、No.2 (2月号) pp8-12、2020
- ・ 木下政人、岸本謙太、吉浦康寿、家戸敬太郎「養殖魚でのゲノム編集—肉厚マダイ・トラフグの作出を例に」、ゲノム編集食品—農林水産分野への応用と持続的社会的実現 (株式会社エヌ・ティー・エス)、2021年1月刊行

F. 健康危険情報

各年度の分担報告書を参照

G. 知的財産権の出願・登録状況

各年度の分担報告書を参照

3) 行政関係者向け説明会

- ・ 農林水産省消費安全局、水産庁との意見交換会 (非公開)、2018年12月17日、水産庁会議室、合計約13名、水産庁、ゲノム編集魚類の現状の説明者
- ・ ゲノム編集水産物に関する検討会 (非公開)、2020年2月27日、農林水産省、20名、農林水産省消費安全局、有識者

4) 学会 (招待講演)

- ・ 第4回ゲノム編集学会、2019年6月3日、タワーホール船堀
- ・ 第31回日本比較免疫学会シンポジウム、2019年9月4日、九州大学
- ・ Marine Biotechnology Conference 2019、2019年9月12日、静岡県清水市清水文化会館
- ・ 第17回食品安全フォーラム (日本薬学会)、2019年11月29日、日本薬学会長井記念ホール

WT :	1	MQLSPSMLHF	SLMISLSLV	LSGQETHQQP	PVGSPEDETEQ	CVTCDVRRQHI
mstn^{-8b/-8b} :	1	MQLSPSMLHF	SLMISLSLV	LSGQETHQQP	PVGSPEDETEQ	CVTCDVRRQHI
mstn^{-41/-41} :	1	MQLSPSMLHF	SLMISLSLV	LSGQETHQQP	PVGSPEDETEQ	CVTCDVRRQHI
mstn^{-8a/-8a} :	1	MQLSPSMLHF	SLMISLSLV	LSGQETHQQP	PVGSPEDETEQ	CVTCDVRRQHI
WT :	51	KTMRINAIKS	QILSKLRMKE	APNISRDTVK	QLLPKAPPLQ	QLLDQYDVLG
mstn^{-8b/-8b} :	51	KTMRINAIKS	QILSKLRMKE	APNISRDTVK	QLLPKAPPLQ	QLLDQYDVLG
mstn^{-41/-41} :	51	KTMRINAIKS	QILSKLRMKE	APNISRDTVK	QLLPKAPPLQ	QLLDQYDVLG
mstn^{-8a/-8a} :	51	KTMRINAIKS	QILSKLRMKE	APNISRE EAAP	AQSA AAAAAAAA AP	RPVRRAGR*
WT :	101	DDNRDVVTEE	DDEHAITETI	MM	MATEPASVVQV	• • •
mstn^{-8b/-8b} :	101	DDNRDVVTEE	DDEH GDDHDD	GH*		
mstn^{-41/-41} :	101	DDNRD HDDGH*				

376

図 1. ミオスタチン遺伝子破壊トラングおよび野生型トラングのミオスタチン遺伝子から産生されるアミノ酸配列。

色字は新生ペプチドを示す。WT:野生型、-41: 41 塩基欠失、-8a および -8b: 8 塩基欠失。

配列全長の比較

	-8a (Full)	-8b (Full)	-41 (Full)	WT
Full FASTA (E.value < 1)	なし	ペルオキシソーム 膜タンパク (アオカビ)	なし	グルテニン
FAO/WHO (>35% in 80 aa)	なし	なし	なし	なし

新生アミノ酸配列とその上流10アミノ酸をクエリとした比較

	-8a (10 + new aa)	-8b (10 + new aa)	-41 (10 + new aa)
FAO/WHO (8 aa exact match)	なし	なし	なし

図2. ミオスタチン遺伝子破壊トラフグのアレルゲン性の検討

センス鎖の通常フレームでの結果を示す。

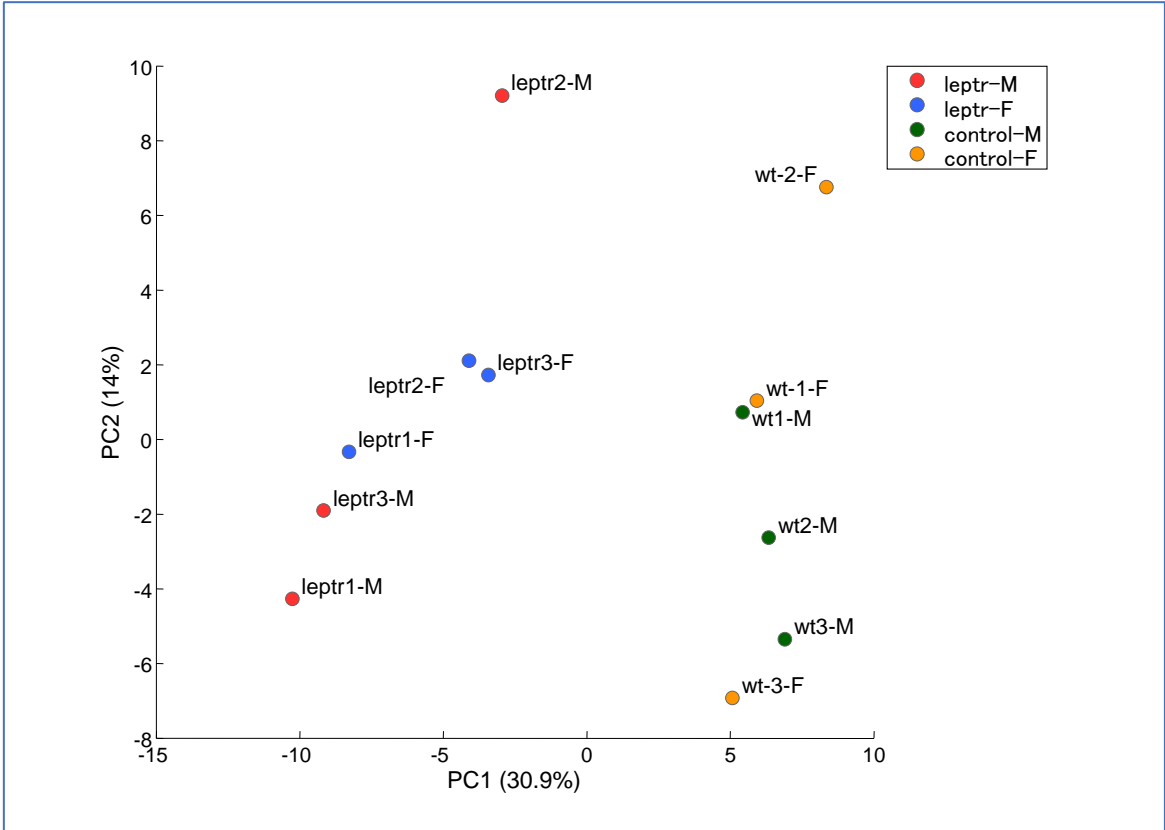


図3. レプチン受容体遺伝子破壊トラフグと野生型トラフグのメタボローム解析（主成分分析）

control: 野生型トラフグ、leptr: レプチン受容体遺伝子破壊トラフグ、
M: 雄、F: 雌

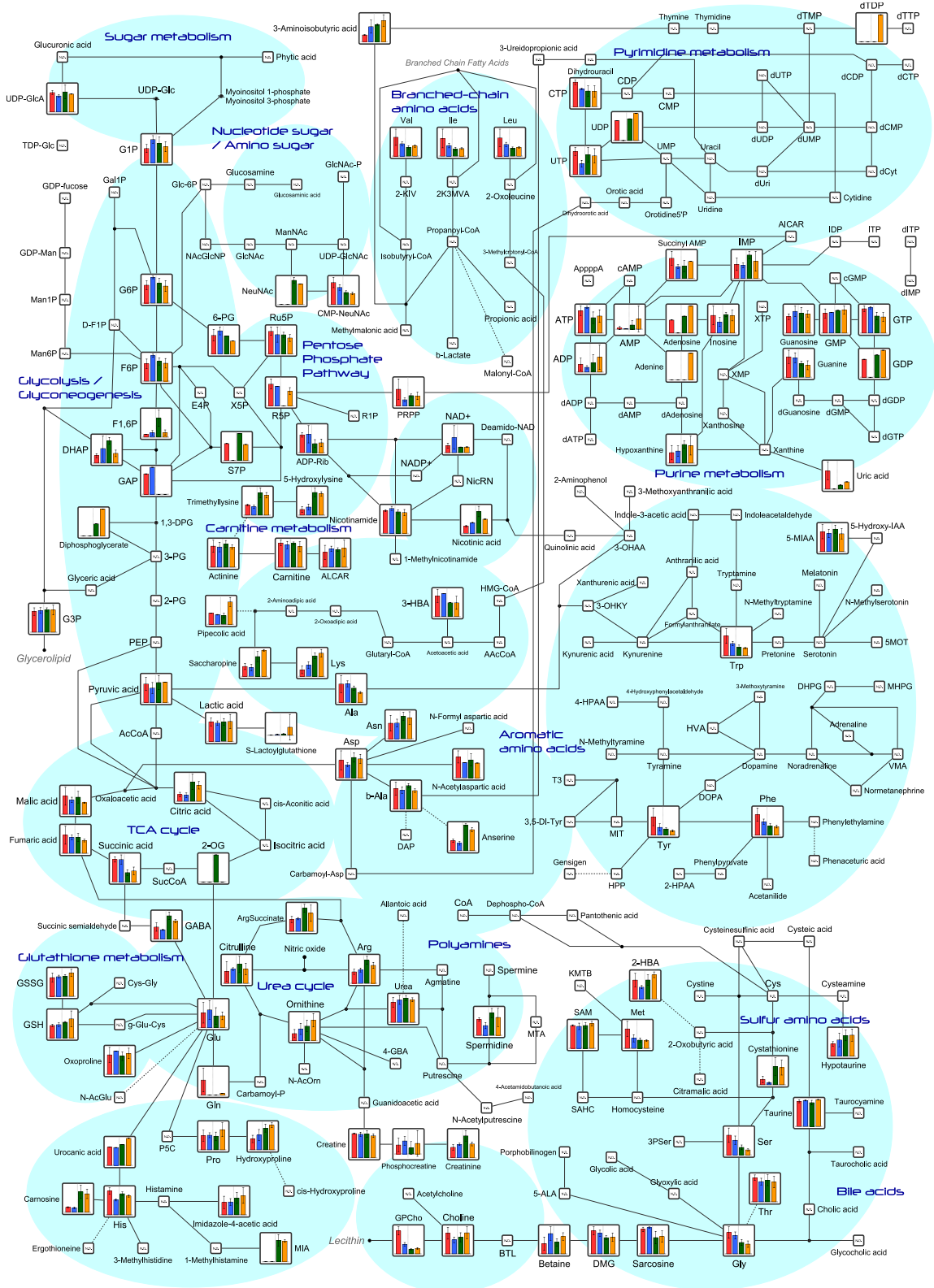
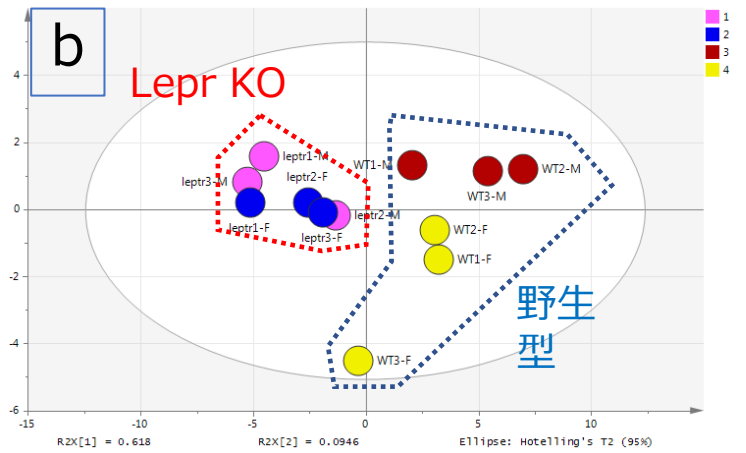
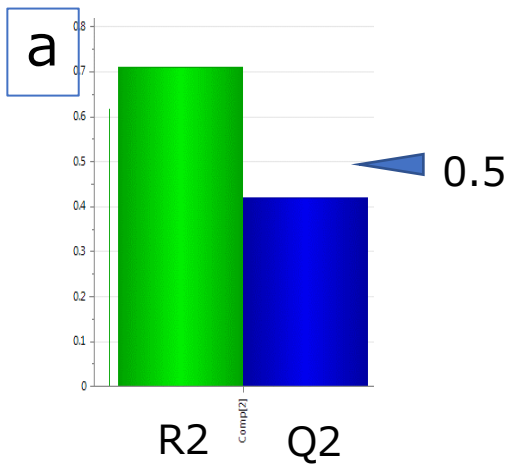


図4. レプチン受容体遺伝子破壊トラフグと野生型トラフグのメタボローム解析 (各パスウェイでの変動概略図)

赤 : lepr KO 雄、青 : lepr KO 雌、緑 : 野生型雄、オレンジ : 野生型雌



a: データ解析の有意性を示す。Q2値が0.5以上の場合、データ解析結果が信頼される。今回のQ2値は0.5以下であった。b: 主成分分析の結果。c: 主成分分析に寄与する因子の解析。lepr KO トラフグ、WT: 野生型トラフグ

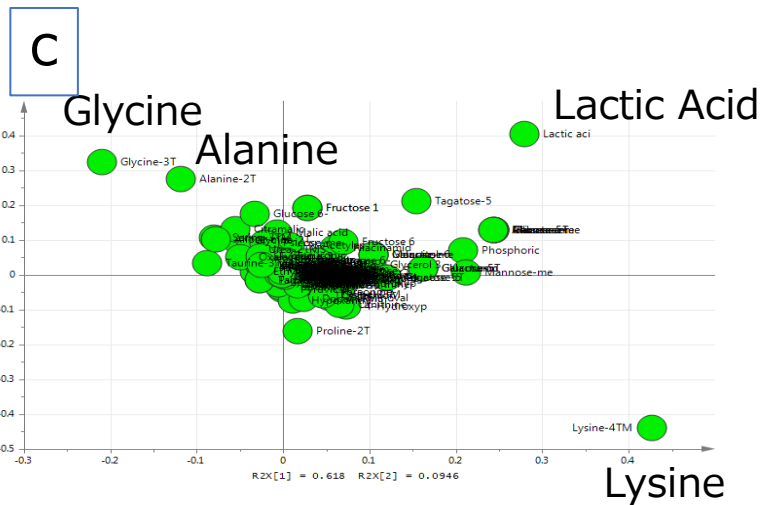


図5. レプチン受容体遺伝子破壊トラフグ系統メタボロミックス解析

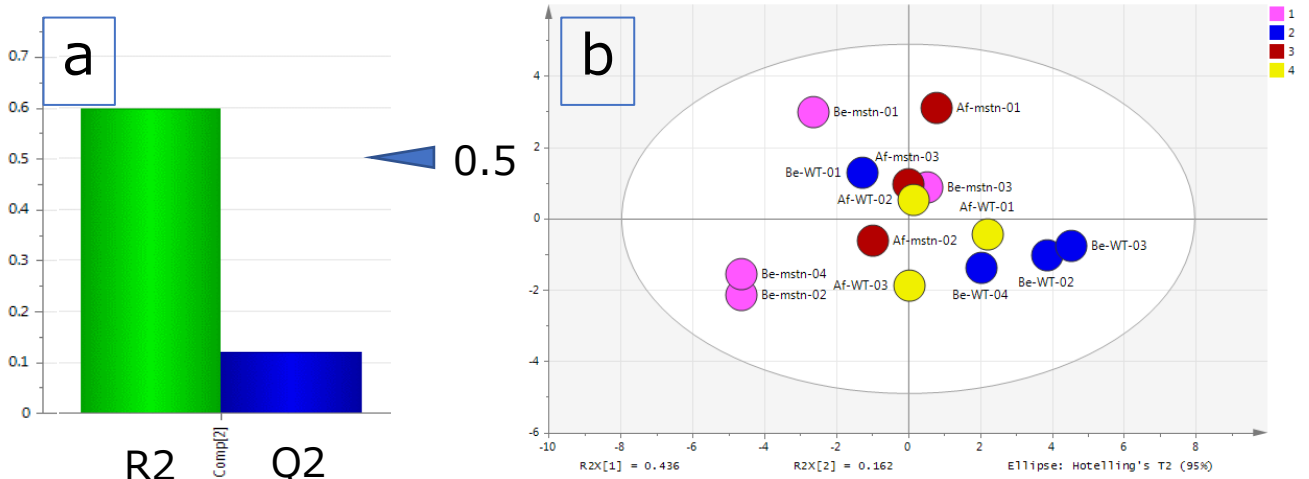


図6. ミオスタチン遺伝子破壊マダイ系統メタボロミックス解析

a: データ解析の有意性を示す。Q2値が0.5以上の場合、データ解析結果が信頼される。今回のQ2値は0.5以下でありbに示される解析の信頼性は低いと判断される。つまり、比較した個体間で有意な差がない。b: 主成分分析の結果。

mstn : mstn KO マダイ WT : 野生型マダイ

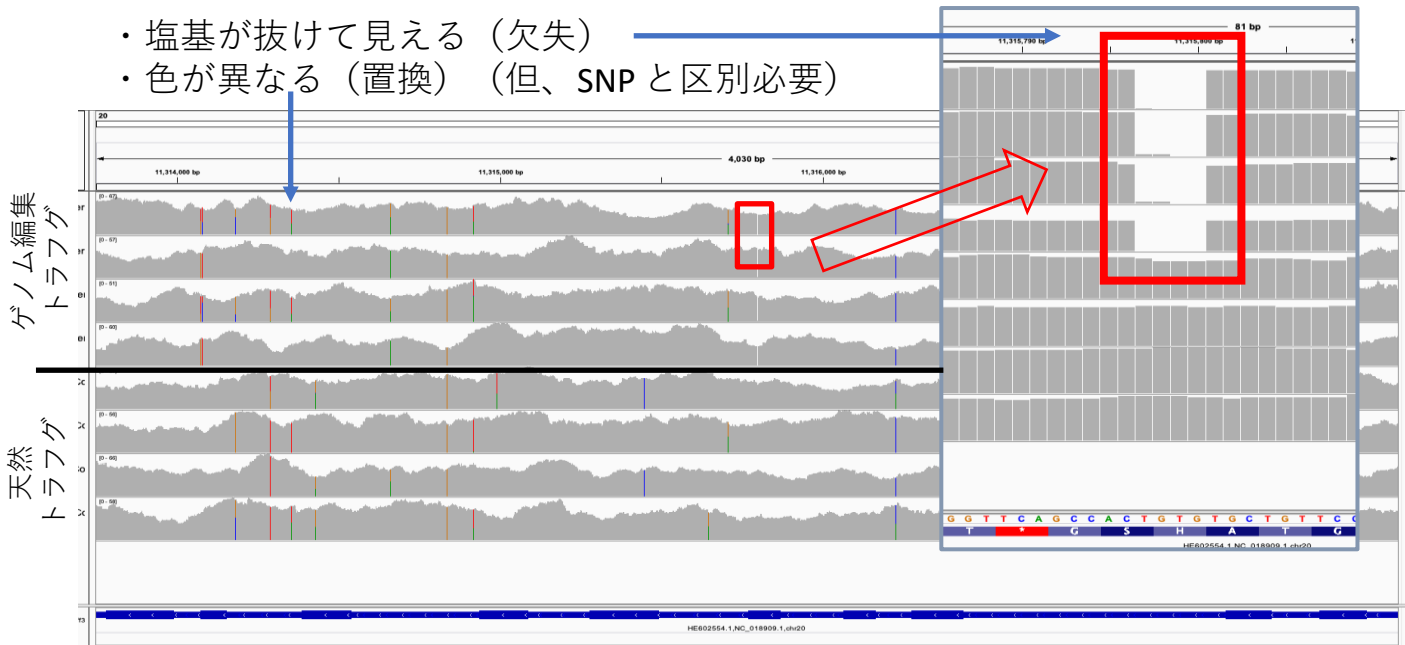


図7. レプチン受容体遺伝子破壊および野生型トラフグの全塩基配列解析結果のIGV像

ターゲット入れる付近の結果を示す。縦軸はリード数を示す。横軸は塩基配列を示す。拡大図に示されるように、塩基欠失箇所は白抜きで示される。

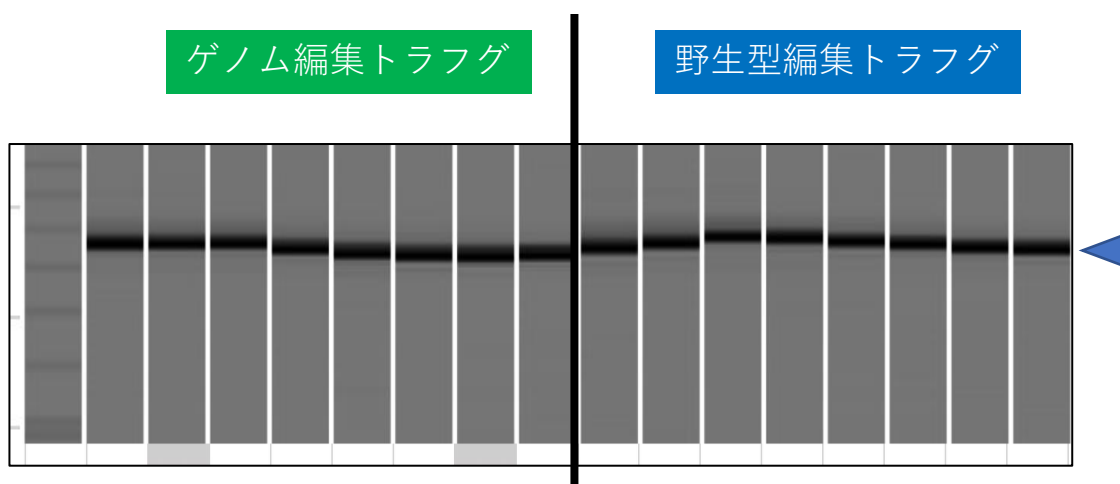


図 8. レプチン受容体遺伝子破壊および野生型トラフグのオフターゲット領域のPCR産物

一例として、オフターゲット候補配列:OT#2 を示す。ゲノム編集および野生型個体それぞれ 8 個体の各 PCR 産物を自動電気泳動装置により解析した。続いて、これらの塩基配列をサンガーシーケンス法により解読した。その結果、いずれの領域にも変異は観察されなかった。

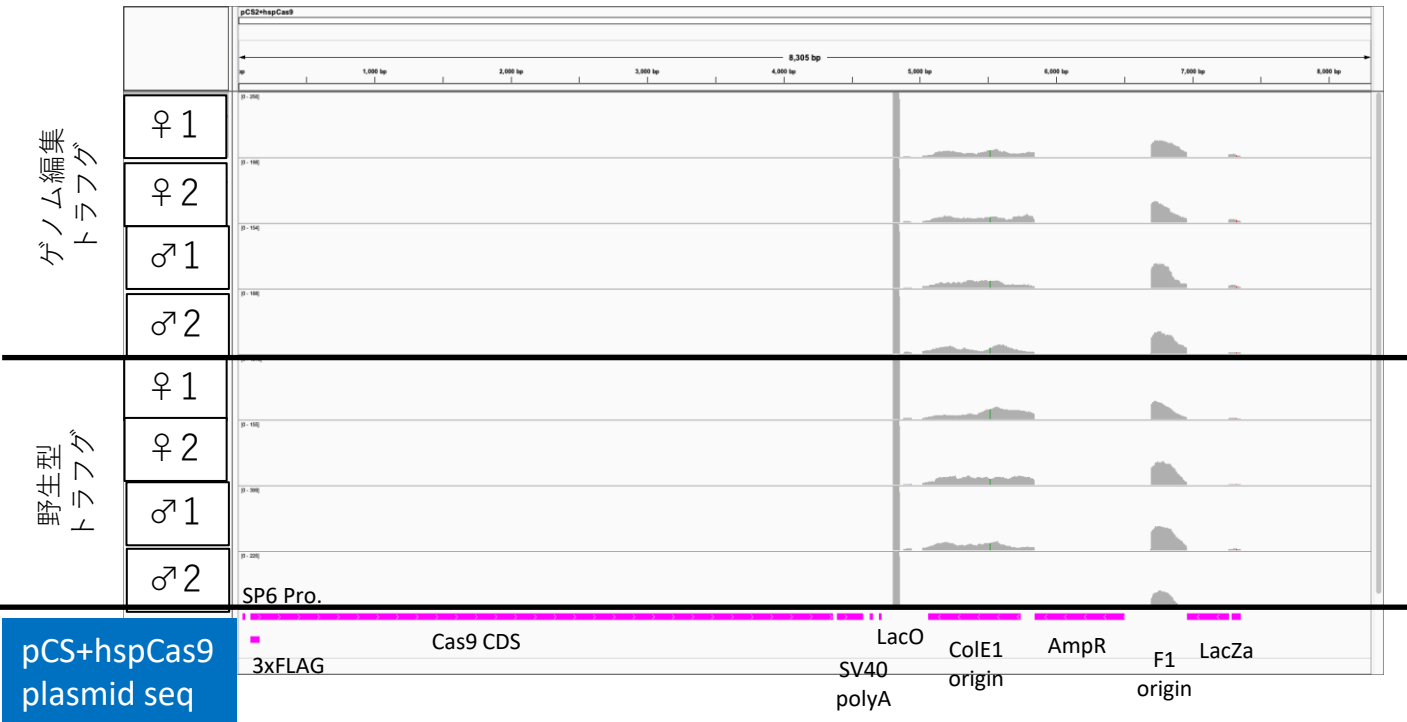


図9. レプチン受容体遺伝子破壊トラフグシステムのゲノム編集ツール残存性の評価 (pCS+hspCas9)

各個体の全塩基配列データをクエリーとしてCas9 RNA 合成時に用いた pCS+hspCas9にマッピングした。ゲノム編集個体および野生型個体ともに、プラスミドバックボーンの配列にマップされるものはあるが、Cas9 領域にはマップされなかった。

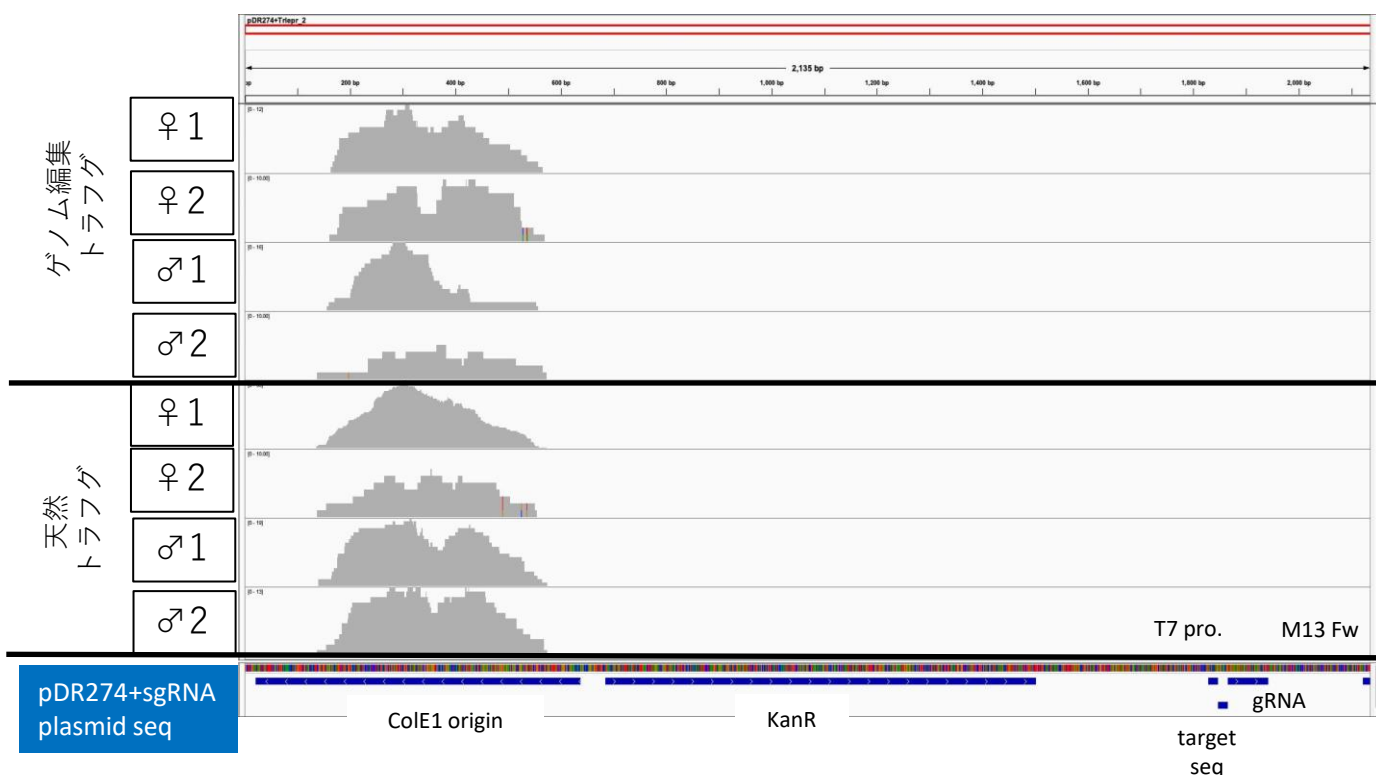


図10. レプチン受容体遺伝子破壊トラフグシステムのゲノム編集ツール残存性の評価 (pDR274+sgRNA)

各個体の全塩基配列データをクエリとしてCas9 RNA 合成時に用いた pDR274+sgRNAにマッピングした。ゲノム編集個体および野生型個体ともに、プラスミドバックボーンの配列にマップされるものはあるが、標的配列を含む guide RNA 領域にはマップされなかった。

表1. ゲノム編集トラフグおよび非編集トラフグのTTX投与後の組織別蓄積率

サンプル ID	雌雄	TTX投与	体長 (mm)	体重 (g)	TTX蓄積量 (投量に対する割合 %)				
					皮	筋肉	肝臓	生殖巣	
ゲノム編集 トラフグ	1	♀	無	354	2040	0	0	0	0
	2	♀	無	355	1909	0	0	0	0
	3	♀	無	428	2933	0	0	0	0
	4	♂	無	425	2679	0	0	0	0
	5	♀	無	397	2953	0	0	0	0
	6	♀	有	352	2092	0	0	29.7	6.1
	7	♀	有	358	2065	0	0	30.7	1.0
	8	♀	有	390	2389	0	0	24.3	8.0
	9	♀	有	410	2696	0	0	17.3	16.4
	10	♂	有	387	2378	0	0	28.5	0
	11	♂	有	380	2381	0	0	18.7	0
非編集 トラフグ	12	♂	無	360	1590	0	0	0	0
	13	♀	無	360	1560	0	0	0	0
	14	♂	有	375	1640	0	0	60.7	0
	15	♂	有	365	1710	0	0	47.0	0
	16	♂	有	380	1600	0	0	51.2	0
	17	♂	有	380	1590	0	0	48.4	0
	18	♂	有	350	1600	0	0	61.9	0
	19	♀	有	360	1560	0	0	31.0	20.3
	20	♀	有	360	1510	0	0	36.7	10.9
	21	♀	有	395	1490	0	0	41.4	6.9
	22	♀	有	395	1640	0	0	35.6	17.6
	23	♀	有	365	1600	0	0	34.7	11.5

TTX 量が検出限界 (0.2 μg/g組織) 以下の場合は、蓄積量を“0”とした。

表2. トラフグ筋肉ホモジネート投与ラットにおける血液学的検査

試験群	動物番号	RBC ($\times 10^6/\mu\text{L}$)	Hb (g/dL)	Ht (%)	MCV (fL)	MCH (pg)	MCHC (g/dL)	WBC ($\times 10^3/\mu\text{L}$)	Platelet ($\times 10^3/\mu\text{L}$)	Reticulocyte (%)
非編集トラフグ 筋肉ホモジネート 投与群	101	6.52	12.7	40.5	62.0	19.4	31.3	4.11	1102	5.15
	102	6.93	12.9	41.7	60.3	18.7	31.0	5.21	1196	5.22
	103	7.08	13.5	43.7	61.8	19.0	30.8	8.12	1125	4.96
	104	6.85	13.2	42.3	61.8	19.2	31.1	6.07	955	5.18
	105	6.79	12.9	42.4	62.4	18.9	30.3	7.10	1017	5.14
	106	7.07	13.0	42.7	60.4	18.4	30.5	4.63	1057	4.73
	107	6.75	12.8	40.7	60.2	19.0	31.5	6.55	1103	5.79
	108	6.97	13.3	42.4	60.8	19.0	31.3	8.17	1283	5.05
Mean	6.87	13.0	42.1	61.2	19.0	31.0	6.25	1105	5.15	353.8
SD	0.19	0.3	1.1	0.9	0.3	0.4	1.53	102	0.3	17.5
ゲノム編集トラフグ 筋肉ホモジネート 投与群	201	6.98	12.8	42.0	60.2	18.4	30.6	5.92	1080	5.48
	202	7.02	13.0	42.3	60.3	18.5	30.7	4.52	981	4.65
	203	7.27	13.7	45.1	62.0	18.9	30.4	7.43	1269	5.11
	204	7.16	13.1	42.6	59.5	18.3	30.7	5.50	1088	4.79
	205	7.03	13.0	41.6	59.1	18.5	31.4	5.91	1122	4.52
	206	7.15	13.4	43.1	60.3	18.8	31.1	4.36	959	4.95
	207	7.52	14.4	45.7	60.8	19.2	31.6	6.76	1110	3.96
	208	7.07	13.5	42.5	60.1	19.1	31.8	4.98	1093	5.41
Mean	7.15**	13.4	43.1	60.3	18.7	31.0	5.67	1088	4.86	346.9
SD	0.18	0.5	1.5	0.9	0.3	0.5	1.07	95	0.50	31.4

RBC: 赤血球数、Hb: ヘモグロビン濃度、Ht: ヘマトクリット値、MCV: 平均赤血球容積、MCH: 平均赤血球色素量、MCHC: 平均赤血球色素濃度、WBC: は血球数、Platelet: 血小板数、Reticulocyte: 網状赤血球数、*: Significant difference (p<0.01) 各項目の平均値について Studentのt検定を用いた。

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」
総合分担研究報告書

ゲノム編集の特性、安全性に関する調査研究

研究分担者 吉場 聡子 （国立医薬品食品衛生研究所）

研究要旨:

本研究では、ゲノム編集技術の発展に伴って発展している RNA 編集技術について、その技術の特徴や有効性、研究開発状況の調査を行い、現状の把握を行なった。まず、文献データベースを用いて、ゲノム編集全般に関する文献調査を行い、研究分野についてその傾向の分析を行った。さらに、その中から RNA 編集に関係するキーワードを抽出し、RNA 編集技術を利用した研究の文献収集を行って、その情報をもとに、技術の概要及び応用研究、課題等についてまとめた。本調査より、RNA 編集技術は現在、ヒトを対象とする研究が主であり、疾患の原因となる遺伝子が明らかになっている中枢神経系を中心とする神経疾患やがんの治療研究が始まっている一方で、植物における RNA 編集の応用研究は見当たらず、現時点では食品分野への応用は可能性が低いと考えられた。

A. 研究目的

CRISPR/Cas9 をはじめとするゲノム編集技術は、現在医療から農業・畜産などさまざまな分野で研究が行われており、社会的にも実装が進んでいる。ゲノム編集は、比較的容易にゲノムの恒久的な改変が行われることから、実用においては、特にオフターゲット（意図しない DNA の改変）リスクを低減すること、適切に安全性評価を行うことが重要な課題となっている。

一方で、ゲノム編集技術の発展に伴って、さまざまな技術や新たな分野が派生的、相乗的に発展してきている。その中で RNA 編集技術は、ゲノム DNA を改変することなく、RNA 配列を編集することで遺伝情報を書き換えるため、リスクの高いゲノム編集に比べてより安全に遺伝情報を操作できると期待されている。

本研究では、まず近年のゲノム編集技術に関する研究の動向を調査し、その傾向を把握する。さらに RNA 編集技術について文献調査を行い、その技術の特徴、有効性、技術的な問題点などを把握し、現在医療分野での研究開発が著しいこの技術の、食品分野への応用の可能性を探ることを目的とする。

B. 研究方法

文献データベースとして、PubMed を用いて、以下の通りキーワードによる検索及び情報取得

を行った。

ゲノム編集に関する文献調査

E-utilities (Entrez API)を利用して、Python を用いて論文の PMID, Journal, Title, Doi, Abstract, Year, Month, Status, MeSH, Keyword の情報を csv 形式で取得した。検索ワードは「genome + editing」、期間は 2010/01/01 - 2020/10/08 とした。さらに取得した文献情報の中から、Title, Abstract, MeSH に含まれる単語を抽出して、クラスター分析 (Mini Batch K-平均化法) を行った。文献の収集及びデータ分析にあたっては、参考文献 1、2 を参考とし、Python の実行は全て JupyterLab 上で行った。

RNA 編集技術に関する文献調査

ゲノム編集に関する文献のクラスター分析で得られたキーワードを参考として、RNA 編集に関する 5 種類の検索用キーワード「ADAR or C-to-U or A-to-I or adenosine-to-inosine or cytidine-to-uridine, mRNA + editing, messenger + RNA + editing, site-directed + RNA + editing, Cas13 + RNA + editing」、期間 2017/01/01 - 2020/10/13 でそれぞれ PMID, Journal, Title, Doi, Abstract, Year, Month, Status, MeSH, Keyword の情報を csv 形式で取得した。さらに、ファイルの統合と重複除去を行った後、項目の整理及び技術に関する論

文の抽出を行い、重要と思われる論文をエクセルファイルにまとめて「RNA 編集技術に関する文献」とした。

C. 研究結果と考察

1. ゲノム編集に関する文献調査

研究方法に記載の方法で、ゲノム編集に関する文献を 17,499 報 (2010.1.1-2020.10.8) 取得した後、クラスター分析により、ランダムに選んだ 4,000 の文献を 60 のクラスターに分類した。その結果、deaminase, adar, adenosine などの RNA 編集に関係するキーワードを含むクラスターに多くの論文が含まれることがわかった。またその結果を tSNE により可視化した。類似性の高いキーワードを含むクラスターはより近傍に配置されるが、RNA 編集に関係するキーワードを含むクラスターの近くに mitochondria, plant, transcriptome, histone, chromatin などのキーワードを含むクラスターが見られた。また、plant は、crop, breeding などと合わせて、ゲノム編集に関わる多くのクラスターに含まれており、広い分野に渡って研究されていることが示唆された。

2. RNA 編集技術に関する文献調査

研究方法に記載の方法で、RNA 編集技術に関する文献 (2017.1.1-2020.10.13) を取得した。キーワードにより回収した文献数は多かった (1,753 報) もの、その中で、RNA 編集技術を用いた研究の数は予想より少なく、多くはゲノムの一塩基編集 (base editing) に関する研究であった。ゲノム編集技術に比べると RNA 編集技術はまだ開発途上であること、応用分野が限られることなどが考えられ、また、現時点では医療分野での研究が主であり、食品分野への応用はほぼ見られなかった。

3. RNA 編集技術について

RNA 編集技術は、一度編集を行うと恒久的にその遺伝情報が細胞に変化が受け継がれるゲノム編集と異なり、DNA 配列には変化を加えず、転写産物である RNA に対して塩基の書換えを行う技術である。RNA 編集は、細胞が持つ進化的に保存されたシステムであり、1987 年に脱アミノ化酵素 adenosine deaminase (ADAR) が RNA 編集活性を持つことが報告されてから、そのメカニズムや利用について研究が行われてきたが、近年のゲノム

編集技術の発展で改めて注目されることになった。

現在 ADAR を用いた RNA 編集システムは、内在の ADAR を利用することで、核酸オリゴの導入のみで編集が可能な RESTORE、LEAPER などのシステムと、ADAR の酵素活性を含むデアミナーゼドメインに Cas13 などの RNA 認識モジュールを連結した、REPAIR、RESCUE などのシステムの 2 種類に大別される。特に、*in vivo* での治療においては、外来遺伝子の導入が不要になる前者のシステムは、安全性の面からも大きなメリットがあり、臨床への応用に向けた研究も行われている。

D. 結論

RNA 編集は、ゲノム DNA に恒久的な変化を加えることなく、効果も一過性のため、ゲノム編集に比べてリスクが低いというメリットがある。一方で、編集効率が低いこと、編集の多様性が不十分であること、またオフターゲットの予測や検証方法などの情報が少ないことなど、テクノロジーとしてはまだゲノム編集ほど確立されていない分野である。現時点では RNA 編集技術は、医療分野での利用が主であり、特に点変異を病因とする中枢神経系疾患の治療で研究が進んでいる。一方で、編集効果の持続が重要になる、作物への利用はあまりアドバンテージがなく、応用研究が見られなかった。

参考文献：

- 1) <https://github.com/tatsuya-takahashi/PubMed-API-Script>
- 2) <https://nbviewer.jupyter.org/github/tatsuya-takahashi/PubMed-API-Script/blob/master/PubMed.ipynb>

E. 研究発表・業績

なし

F. 健康危険情報

なし

G. 知的財産権の出願・登録状況

なし

研究成果の刊行に関する一覧表
(平成30年度～令和2年度)

書籍

著者氏名	タイトル名	書籍名	出版社名	出版地	出版年	ページ
木下政人	「肉厚マダイ」の事例から考える養殖業へのゲノム編集技術の活用	月刊養殖ビジネス	緑書房	東京都	2020	1月号 61-64
木下政人	ゲノム編集技術を使った肉厚マダイの作出と品種改良期間の短縮	JATAFFジャーナル	農林水産・食品産業技術振興協会	東京都	2020	8巻 2号 8-12

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Takabatake, R., Kagiya, Y., Minegishi, Y., Futo, S., Soga, K., Nakamura, K., Kondo, K., Mano, J., Kitta, K.	Rapid screening detection of genetically modified crops by loop-mediated isothermal amplification with a lateral flow dipstick.	<i>Journal of Agricultural and Food Chemistry</i>	66	7839-7845	2018
Soga, K., Nakamura, K., Kishine, M., Takashima, Y., Miyahara, T., Kimata, S., Mano, J., Takabatake, R., Ozeki, Y., Kitta, K., Kondo, K.	Studies on the detection of maize genomic DNAs in cornflakes using real-time PCR. (リアルタイムPCRを用いたコーンフレーク中のトウモロコシゲノムDNA検出について)	<i>Bulletin of National Institute of Health Sciences</i>	136	31-39	2018

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Nakanishi, K., Fujii, U., Nakamura, K., Ohtsuki, T., Kimata, S., Soga, K., Kishine, M., Mano, J., Takabatake, R., Kitta, K., Kawakami, H., Akiyama, H., Ikeda, M., Kondo, K.	Effect of sodium carboxymethyl cellulose in processed rice foods on detection of genetically modified rice- derived DNA.	<i>Japanese Journal of Food Chemistry and Safety</i>	25	77-85	2018
Nakamura, K., Ishigaki, T., Kobayashi, T., Kimata, S., Fujii, U., Soga, K., Kishine, M., Takabatake, R., Mano, J., Kitta, K., Kawakami, H., Nishimaki-Mogami, T., Kondo, K.	Identification of chickpea (<i>Cicer arietinum</i>) in foods using a novel real-time polymerase chain reaction detection method.	<i>Journal of Food Composition and Analysis</i>	71	8-16	2018
Kishine, M., Noguchi, A., Mano, J., Takabatake, R., Nakamura, K., Kondo, K., Kitta, K.	Detection of DNA in highly processed foods. (高度加工食品からの原材料 農産物 DNA 検出の検討)	<i>Food Hygiene and Safety Science</i>	59	151- 156	2018
Tsugawa, H., Nakabayashi, R., Mori, T., Yamada, Y., Mikiko Takahashi, M., Rai, A., Sugiyama, R., Yamamoto, H., Nakaya, T., Yamazaki, M., Kooke, R., Bac- Molenaar, J., Oztolan- Erol, N., Keurentjes, J., Arita, M., Saito, K.	A cheminformatics approach to characterize metabolomes in stable isotope-labeled organisms.	<i>Nature Methods</i>	16	295- 298	2019

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Burla, B., Arita, M., Arita, M., Bendt, A.K., Cazenave-Gassiot, A., Dennis, E.A., Ekroos, K., Han, X., Ikeda, K., Liebisch, G., Lin, M.K., Loh, T.P., Meikle, P.J., Orešič, M., Quehenberger, O., Shevchenko, A., Torta, F., Wakelam, M.J.O., Wheelock, C.E., Wenk, M.R.	MS-based lipidomics of human blood plasma: a community- initiated position paper to develop accepted guidelines.	<i>Journal of Lipid Research</i>	59 (10)	2001- 2017	2018
Sakuma, T., Nishi, K., Kishimoto, K., Nakagawa, K., Karasuyama, M., Umezu, Y., Kajioka, S., Yamazaki, S.J., Kimura, K.D., Matsumoto, S., Yoda, K., Fukutomi, M., Shidara, H., Ogawa, H., Takeuchi, I.	Efficient learning algorithm for sparse subsequence pattern-based classification and applications to comparative animal trajectory data analysis.	<i>Advanced Robotics</i>	33	134- 152	2019
Karasuyama M., Inoue K., Nakamura R., Kandori H., Takeuchi I.	Understanding colour tuning rules and predicting absorption wavelengths of microbial rhodopsins by data-driven machine-learning approach.	<i>Scientific Reports</i>	8	#15580	2018
Kanamori K., Toyoura K., Honda J., Hattori K., Seko A., Karasuyama M., Shitara K., Shiga M., Kuwabara A., Takeuchi I.	Exploring a potential energy surface by machine learning for characterizing atomic transport.	<i>Physical Review B</i>	97	#12512 4	2018

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Kawaguchi, N., Tomita, C., Naradate, R., Higami, T., Nakamura, K., Date, K., Aikawa, K., Ogawa, H.	A novel protocol for the preparation of active recombinant human pancreatic lipase from <i>Escherichia coli</i> .	<i>Journal of Biochemistry</i>	164 (6)	407- 414	2018
Soga, K., Nakamura, K., Ishigaki, T., Kimata, S., Ohmori, K., Kishine, M., Mano, J., Takabatake, R., Kitta, K., Nagoya, H., Kondo, K.	Data representing applicability of developed growth hormone 1 (GH1) gene detection method for detecting Atlantic salmon (<i>Salmo salar</i>) at high specificity to processed salmon commodities.	<i>Data in Brief</i>	27	#10469 5	2019
Narushima, J., Kimata, S., Soga, K., Sugano, Y., Minegishi, Y., Kishine, M., Takabatake, R., Mano, J., Kitta, K., Kanamaru, S., Shirakawa, N., Kondo, K., Nakamura, K.	Rapid DNA template preparation directly from a rice sample without purification for loop-mediated isothermal amplification (LAMP) of rice genes.	<i>Bioscience, Biotechnology, and Biochemistry</i>	84	670- 677	2020
Soga, K., Nakamura, K., Ishigaki, T., Kimata, S., Ohmori, K., Kishine, M., Mano, J., Takabatake, R., Kitta, K., Nagoya, H., Kondo, K.	Development of a novel method for specific detection of genetically modified Atlantic salmon, AquAdvantage, using real-time polymerase chain reaction.	<i>Food Chemistry</i>	305	#12542 6	2020

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Mishiba, K.-I., Iwata, Y., Mochizuki, T., Matsumura, A., Nishioka, N., Hirata, R., & Koizumi, N.	Unfolded protein-independent IRE1 activation contributes to multifaceted developmental processes in Arabidopsis.	<i>Life Science Alliance</i>	2(5)	#20190 0459	2019
Hirata, R., Mishiba, K.-I., Koizumi, N., & Iwata, Y.	Deficiency in the double-stranded RNA binding protein HYPONASTIC LEAVES1 increases sensitivity to the endoplasmic reticulum stress inducer tunicamycin in Arabidopsis.	<i>BMC Research Notes</i>	12	#580	2019
Hirohata, A., Sato, I., Kaino, K., Iwata, Y., Koizumi, N., & Mishiba, K.-I.	CRISPR/Cas9-mediated homologous recombination in tobacco.	<i>Plant Cell Reports</i>	38	463- 473	2019
Yoshida T., Takeuchi I., Karasuyama M.	Safe Triplet Screening for Distance Metric Learning.	<i>Neural Computation</i>	31 (12)	2432- 2491	2019
Umezu Y., Takeuchi I.	Selective inference via marginal screening for high dimensional classification.	<i>Japanese Journal of Statistics and Data Science</i>	2	559- 589	2019

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Soga, K., Kimata, S., Narushima, J., Sato, S., Sato, E., Mano, J., Takabatake, R., Kitta, K., Kawakami, H., Akiyama, H., Kondo, K., Nakamura, K.	Development and Testing of an Individual Kernel Detection System for Genetically Modified Soybean Events in Non-identity-preserved Soybean Samples.	<i>Biological and Pharmaceutical Bulletin</i>	43	1259- 1266	2020
Takabatake, R., Onishi, M., Minegishi, Y., Futo, S., Soga, K., Nakamura, K., Kondo, K., Mano, J., Kitta, K.	Development of a Novel Detection Method Targeting an Ultrashort 25 bp Sequence Found in Agrobacterium-Mediated Transformed GM Plants.	<i>Journal of Agricultural and Food Chemistry</i>	68	15327- 15334	2020
Takabatake, R., Onishi, M., Mano, J., Kishine, M., Soga, K., Nakamura, K., Kondo, K., Kitta, K.	Evaluation of Conversion Factors for Genetically Modified Maize Quantification. (遺伝子組換えトウモロ コシ定量のための内標比 の算出)	<i>Shokuhin Eiseigaku Zasshi</i>	61	235- 238	2020
Matsuda, M., Iwata, Y., Koizumi, N., & Mishiba, K.-I.	Zeocin-induced DNA double strand breaks affect endoreduplication and cell size in radish cotyledon epidermis	<i>Cytologia</i>	85	245- 249	2020
Mishiba, K.-I., Nishida, K., Inoue, N., Fujiwara, T., Teranishi, S., Iwata, Y., Takeda, S., & Koizumi, N.	Genetic engineering of eggplant accumulating β -carotene in fruit	<i>Plant Cell Reports</i>	39	1029- 1039	2020

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Inatsu Y., Karasuyama M., Inoue K., Takeuchi I.	Active learning for level set estimation under input uncertainty and its extensions	<i>Neural Computation</i>	32	2486- 2531	2020