

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」

総括研究報告書（令和2年度）

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	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	吉場 聡子	（国立医薬品食品衛生研究所）

研究要旨：

ゲノム編集食品の届出制度が平成 31 年度（令和元年）10 月に開始となり食品衛生法上の取扱いも明確化された。今後、制度実施にあたり、安全性確認を科学的なエビデンスをもって行うための手法開発整備も重要な課題である。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。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を行った。具体的には、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。

その結果、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。今後は検体数を増やした検討も必要である。本検討により、ゲノム編集食品の事前相談・届出制度にある必要項目に沿った十分なデータを取得してその安全性を確認した。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。ここでは、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例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DNA二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ギー分解性の検討

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卵白オボムコイドについて検討するために、タンパクの生成を行った。

## 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

オフターゲットを、次世代シーケンサー (NGS) を用いて予測する手法 SITE-Seq 法を用いて検討した。抽出 DNA を用いて潜在的オフターゲットをゲノムワイドに予測する SITE-Seq 法の、*in vivo* での検証を行った。予測された候補オフターゲット部位について、イネカルスを用いて検証した。シーケンス解析は、SITE-Seq 法の予測部位に対してデータ量を厚く、信頼性を確保しつつ低頻度の変異を検出するためにアンプリコンシーケンスを行うことで、*in vivo* での検証を行った。

### (2) リスクコミュニケーションに関する研究

一般モニターと専門家を対象にした web アンケートを前年度実施してそれを解析した。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。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

化合物の質量スペクトル (フラグメントスペクトル) 類似度をもとにして、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拡張を行った。スペクトル類似度計算および部分構造推定に用いるフラグメントスペクトルライブラリは、公共データベース及び実際に標準品、植物抽出体、食品試料を測定して得られたスペクトルをライブラリーへ追加して作製した。毒性情報に関しては網羅的な毒性物質データベースである T3DB を XML 解析することでスペクトルライブラリ内の化合物に毒性および毒性のカテゴリ情報の付与を行った。データ解析の実行環境として、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。測定試料として、食用及び有毒植物 (菌類含む)、ゲノム編集ジャガイモなどの試料を用いて検討した。

### (4) アレルギーデータベース ADFS のアップデート、および新規タンパクアレルギー性予測に必要な情報の作製

アレルギー情報について、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питープ情報について、2019 年 6 月から 2020 年 5 月までの 1 年間に NCBI PubMed に収載された論文を査読して必要なエпитープ情報を含む場合はその情報を整理し、アレルギーデータベース (ADFS) のデータに追加した。

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。

### (5) 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

これまでに、食品タンパク質の大規模データベースを整備し、アレルギー特異的な様々な長さのアミノ酸配列を抽出、これらに基づいてアレルギー性判定・予測システムを検討してきた。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ことを目的としている。令和 2 年度は、平成 30 年度に作製したプロトタイプ、令和元年度のシミュレーションで明らかになった以下の 3 つの課題を解決し、AI によるアレルギー性判定・予測システムを完成させた。

#### (課題 1 の解決)

アレルギー性タンパク質数 (正例数) と非アレルギー性タンパク質数 (負例数) に偏りがあることについて、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、令和 2 年度は、新規にアレルギー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、非アレルギー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 の解決)

アレルギー性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。この時、2 つのアプローチがある。一つは、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り、もう一つは、タンパク質を構成するアミノ酸の疎水性や分子量などの平均、分散、相関を求めて特徴として抽出することであるが、本研究ではアミノ酸の順序も重要なことから、これが考慮できる前者の方法を用いている。教師データとして、アレルゲンデータの他に、非アレルゲンデータを用い、アレルゲン性の判定・予測により有用な特徴を抽出した。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の解決)

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。そして、非食品タンパク質データに関しては正例のみが存在する。一方で、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。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する。

#### (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

作出したゲノム編集トラフグについて特性を解析して、食品安全性の評価法への提言となることを念頭に検討した。今年度は、ゲノム編集トラフグの毒分布の変化、可食部の毒性という点を中心に検討した。

毒分布については、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および非編集トラフグへのテトロドトキシン投与により、皮、筋肉、肝臓、生殖巣への毒の分布を、HPLCを用いて定量測定した。可食部の毒性については、背部普通筋を採取し、これをラットへ投与、短期(14日間)毒性試験を行った。

#### (7) 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。

#### (8) 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る。

### 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

ゲノム編集技術を用いた研究開発の文献2020年について調査した結果、動物ではゲノム編集技術を用いたもの191件、植物ではゲノム編集技術を用いたもの310件、きのこは1件であった。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件数は少ない。植物では、イネでの研究が活発で、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、リンゴ、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。技術的には、CRISPR/Casがほとんどである。

アレルゲン分解性については、検討に用いるため卵白オボムコイドの精製を行った。

#### 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

アンプリコンシーケンスの各リードをCRISPResso2で解析した結果、本研究で標的として用いたOsFH15-gRNAは高い切断活性を有していることが示唆された。一方でオフターゲット部位については、1つのオフターゲットのみ、5塩基欠失が確認された。今回、細胞から抽出したDNAを利用する*in vitro*法であるSITE-Seq法を用いて潜在的な切断部位を予測し、その一部は*in vivo*においても実際に確認できた。このオフターゲット部位は標的部位(オンターゲット)とは配列類似

性が高くなく殆どの *in silico* 検索法で検索できなかった。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証法は、非常に有用な方法であることが示された。

さらに標的遺伝子 OsFH15 はエキソン上に設計されているため、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率することで、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

前年度実施した一般モニター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(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とともに、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。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。このうち15実施例について詳細な調査を行った(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

本研究では、昨年度までに構築した基本的な

データ解析システムを拡充することで未知化合物に関する豊富な情報を迅速に出力するシステムへと発展させた。食品から抽出した未知化合物を含む試料のLC-MSデータを本解析プラットフォームで解析した場合、未知化合物であっても「化合物クラス・構造的に類似した化合物・部分構造・構造類似性に基づく毒性」の情報が自動的に得られるようになった。未知化合物のLC-MSデータから豊富な情報を見出すデータ解析系はこれまでに存在せず、本データ解析プラットフォームは独自性の高いものと言える。一般に高度な質量分析データ解析アルゴリズムはインストールや操作が難解なものが多く、非インフォマティクスの一般の分析化学の研究者に実際に利用されることが少ないというケースが多々あった。本研究のデータ解析プラットフォームはWebブラウザ上でGUIによりインタラクティブに操作が可能となるように設計しており、より幅広い分野の研究者に使用されることが期待される

### (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予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、登録アレルゲンは平成20年度に AllergenOnline の内容と統一し、その後も AllergenOnline の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を行っている。令和2年度もアップデート作業を実施した。エピトープ情報は、論文レビューを独自に行い、アレルゲン及びイソアレルゲンのアミノ酸配列情報は2,339、エピトープ既知のアレルゲン数は247、構造既知のアレルゲン数は169、糖鎖付加アレルゲン数は127となった。

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

最新のアレルゲン情報をデータセットに加えるとともに、アレルゲンおよび非アレルゲンデータセットの精査を行った。特に、タンパク質に完全にアレルゲン性がないことを確定させ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登録されているもののなかには一定の False Negative が含まれている可能性が否めない。課題 2 に対して、令和 2 年度は、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。課題 3 に対しては、既存のアプローチである 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

フグ毒の体内分布の検討：フグ毒テトロドトキシン TTX の蓄積は、ゲノム編集、非編集ふぐのいずれにおいても肝臓および雌の生殖巣（卵巣）飲みに見られた。フグ毒蓄積には、サキントキシン・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が、ゲノム編集の標的ではなくその代謝には影響を与えないと考えられた。

#### (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 1 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る、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。

#### D. 結論

各分担研究者の課題について、計画 3 年目において概ね予定通りに実施され、成果を得ることができた。詳細については、分担報告書を参照。

#### E. 研究発表・業績

各分担報告書に記載

#### F. 健康危険情報

該当なし

#### G. 知的財産権の出願・登録状況

該当なし