

## 生命科学に関するデュアルユースに関する分析

研究分担者 木賀 大介 早稲田大学 理工学術院 教授

### 研究要旨：

合成生物学に関する国際学生コンテストiGEMでのデュアルユース教育について、大会に参加することを通じて調査した。また、昨年に続き、合成生物学を学ぶ学生に、彼らが内容を理解している原著論文の技術を転用した際のデュアルユースのシナリオの検討を依頼し、この内容を分析した。その過程で浮かび上がってきた、タンパク質のアミノ酸配列を生成するAIについては、継続的に注意を向ける必要がある。

### A. 研究目的

本研究課題の主題である生命科学に関するデュアルユース問題について、「先進生命科学技術のデュアルユース問題とバイオセキュリティ」の観点に立ち、この領域で学ぶ学生から見た論文情報のデュアルユースに対する着眼点、および、合成生物学に関する国際学生コンテストiGEMでのデュアルユース教育について、今後に向けてのありべきガバナンスの展望を考察する。

### B. 研究方法

#### 学生から見た論文情報のデュアルユースに関する検討

合成生物学の研究室に所属する学部または修士学生について、彼らが研究室で文献紹介した原著論文について、デュアルユースがあるとしたらどのようなケースが考えられるか、の記載を依頼した。依頼時に、自由意志での参加不参加の判断が可能であること、および、成果報告での各記載と各参加者の対応付けを行わないことを告知した。

紹介論文の一つには、AIを用いた、機能性たんぱく質の配列生成の論文があり、その重要性から、木賀も関連論文を精査した。

#### 合成生物学に関する国際学生コンテストiGEMでのデュアルユース教育についての調査

合成生物学に関する国際学生コンテストiGEMについて、参加学生チームの監督、および、審判委員会のメンバーとして活動し、コンテストにおけるデュアルユース教育について調査した。

### C. 研究結果

#### 学生から見た論文情報のデュアルユースに関する検討

各論文について、一文概要と想定されるデュアル

ユース短文は、学生による記載を木賀が要約したものとなっている。

### 論文1

Horizontal gene transfer enables programmable gene stability in synthetic microbiota.  
*Nature Chemical Biology*. 18. 1245–1252 (2022).  
<https://doi.org/10.1038/s41589-022-01114-3>

#### (一文概要)

環境中の微生物間で、遺伝子が水平伝搬により安定的に共有される条件の解明。

#### (想定されるデュアルユース短文)

環境に悪影響を及ぼす遺伝子を微生物間で拡散させることで、環境中に安定的に維持する。

#### (論文概要)

機能遺伝子の相対的存在量で定義される微生物群集の機能は、群集内の種の組成と強く相関している。淡水や土壌の微生物群集は機能遺伝子レパートリーがその種の組成と有意な相関を示している。しかし、腸内細菌や海水、大型藻類などの多くの生物群集において、異なる種の割合が集団ごと大きく異なるが、機能構造は非常に保存されている。

本研究では移動性遺伝要素（mobile genetic element: MGE）を介した遺伝子の逆流を、機能・組成の分離に寄与するメカニズムとして提唱した。つまり、生物種を超えた遺伝子の動的な移動によって、異なる種の生物で同じ機能遺伝子を共有し、生物種組成が変わっても機能遺伝子の相対的存在量が安定に維持されていると仮説を提唱した。また、微生物群集における遺伝子量の安定性は、遺伝子の水平伝播（HGT）の速度によって制御できることを理論解析と数値シミュレーションによって示した。

(想定されるデュアルユース)

本研究では、細胞間転移可能なプラスミドに介し、伝搬の過程で損失しながらも、共有の速度が十分に早ければ、元から群集内で広く共有されている遺伝子同様、またはそれ以上の安定な機能遺伝子の相対的存在量に達し、安定化を促すことが可能であると証明された。

よって、この観点から見れば、水平伝搬可能なプラスミドに任意な遺伝子を載せ、宿主細胞を細胞集団に投入し増やせば、結果的に遺伝子情報は細胞集団内で種に関わらず広く保存されるようになり得る。遺伝子の発現制御こそまだ言及されていないが、宿主細胞が除去された後も生物群集の遺伝子プールに特定な遺伝子を残せることが考えられる。

それによって、特定な微生物群集に異常な環境耐性を与え、過度に繁殖させて赤潮のように他の生物の生存環境と資源を奪い生態系に害を成すことや、水平伝搬先の細胞による毒性物質の発現によって、捕食者に害を成して食物連鎖を汚染する可能性も存在すると、私は考える。

## 論文2

Single cell characterization of a synthetic bacterial clock with a hybrid feedback loop containing dCas9-sgRNA.

*ACS Synthetic Biology*. **9**. 3377–3387 (2020).

<https://doi.org/10.1021/acssynbio.0c00438>

(一文概要)

CRISPRiのシステムにより、これまでよりも制御の対応関係に関して柔軟に遺伝子回路を構築

(想定されるデュアルユース短文)

細胞の挙動を制御する遺伝子回路を乗っ取り、細胞のゲノム構造を改変することによる悪用

(論文概要)

遺伝子回路研究が盛んに行われ始めた 2000 年代初頭、Elowitz らによって LacI、TetR、 $\lambda$ -cl の 3 つの遺伝子抑制因子が周期的に抑制し合う負のフィードバック回路が発見された。本論文では、このフィードバック回路に対して 3 つの因子のうちの 1 つを dCas9 タンパク質に変更し、遺伝子回路内に CRISPRi を導入した。

この新たに作成した遺伝子回路を大腸菌内に組み込み、振動を起こすかどうか検証した。

(想定されるデュアルユース)

本論文の新規性は、遺伝子回路内に CRISPRi を導入した点にあると考える。CRISPRi 技術では、dCas9 単体では DNA 切断能はないものの、適切な sgRNA を導入すれば、本来の Cas9 のように DNA の切断能を有する。これは、遺伝子合成技術が安価かつ迅速に行えるようになった昨今では、CRISPRi を

はじめとする CRISPR 技術を用いれば、人為的に任意のゲノム上の編集したい箇所を切断したり結合したりすることは容易であることを意味する。つまり、簡単に DNA 編集を行うことができってしまう。よって、回路を組み込んだ大腸菌のゲノム機能を変えてしまうことも容易であると考えられる。

加えて、本研究では振動周期を CRISPR 技術を用いて特定の配列部を加えたり除去したりすることで容易に変化させることができる。そのため、将来的に日常生活にタイマーなどのタイムキーパーのツールとして応用される可能性が考えられる。

以上の予測を踏まえて、想定される本研究の悪用方法として、第三者によって遺伝子回路に用いる sgRNA が、設計時に想定した sgRNA ではない異なる sgRNA にすり替えられたり、CRISPR 技術を用いて、大腸菌のゲノム情報が書き換えられたりすることが挙げられる。万が一、このような事態が発生した場合、生存環境への高耐性大腸菌や不滅の大腸菌などといった、本来設計時に想定していた範囲を超える挙動を起こす遺伝子回路に書き換えられてしまう可能性がある。もしそのような事態が発生した場合、回路設計者が想定していったタイミングで回路が停止しない、回路の周期が乱れる、本来生産しない物質が生産させてしまう、といったリスクが考えられる。

## 論文3

Cell-free prototyping strategies for enhancing the sustainable production of polyhydroxyalkanoates bioplastics.

*Synthetic Biology*. **3**. ysy016 (2018).

<https://doi.org/10.1093/synbio/ysy016>

(一文概要)

生きた細胞内でなく、細胞破碎液を用いた試験管内反応により、生物機能を用いた物質生産条件の探索を効率化

(想定されるデュアルユース短文)

有害な物質の生産システムを迅速に構築

(論文概要)

ポリヒドロキシアルカノエート (PHA) は微生物由来のバイオポリマーであり、石油由来のプラスチックに代わる可能性があるが、PHA の生産コストは高く、より効率的に生産する必要がある。無細胞代謝工学は既にいくつかの生合成経路の最適化に利用されており、これが PHA の生合成経路の最適化にも利用できると考えられた。そこで *in vitro* で PHA 生合成オペロンを試作し、代謝物リサイクル酵素をスクリーニングするための大腸菌無細胞系を何種類か開発した。さらに *in vivo* での PHA の生産を最適化するために産業廃棄物である乳清透過液を加えることで無細胞反応をカスタマイズした。結果、最適

な濃度で乳清を添加することで GFPmut3b の生産量が約 50%向上した。無細胞転写-翻訳反応ではガスクロマトグラフィー質量分析で無細胞の 3-ヒドロキシ酢酸 (3HB) の定量によって PHA 生合成オペロンの活性の違いが明らかになった。最も活性が向上していたオペロンである C104 は in vitro、in vivo の両方で野生型よりも高いレベルの PHA を生産した。

(想定されるデュアルユース)

本研究では in vitro 無細胞翻訳反応系で適切にプロトタイピングすることによって、in vivo でも生産を促進させる、有用な代謝リサイクル酵素のスクリーニングにつながれるというものであった。そこで悪用方法として考え付いたのが in vitro で人の生体内で有害な物質ができるように無細胞翻訳系でプロトタイピングしたのちに、in vivo 戦略に組み合わせて人体に有害な物質を作ることが考えられる。これによって、生物兵器の開発にもつながる。

#### 論文4

Glyceraldehyde-3-phosphate dehydrogenase is required for efficient repair of cytotoxic DNA lesions in *Escherichia coli*.

*The International Journal of Biochemistry & Cell Biology*. **60**. 202–212 (2015).

<http://dx.doi.org/10.1016/j.biocel.2015.01.008>

(一文概要)

バクテリアでは解糖系酵素が DNA 修復過程にも関与

(想定されるデュアルユース短文)

抗生物質を隠れて散布することで人や家畜の健康を乱す

(論文概要)

解糖系酵素のグリセルアルデヒド-3-リン酸デヒドロゲナーゼ (GAPDH) は、ヒトの細胞内で多様な生物学的機能を有する多機能性タンパク質である。一方、細菌においては、分泌タンパク質としての多機能性のみ報告されている。細胞内では、大腸菌由来 GAPDH が DNA 修復産物である 2-phosphoglycolate の代謝に関与する phosphoglycolate phosphatase と相互作用することから、GAPDH が DNA 修復過程に関与していると推定されている。

本研究では、大腸菌の DNA 修復過程と GAPDH の関与について検証された。結果として、(1) GAPDH の欠損または抑制による遺伝毒性薬剤に対する感受性の増加と糸状成長と未修復 DNA 損傷の蓄積、(2) GAPDH 欠損株における自発的 AP サイトと自発的突然変異率の増加、(3) GAPDH と DNA 修復経路のタンパク質 (Endo IV、UDG) や SOS 応答に関するタンパク質 (SSB) との相互作用、が観察された。以上の

ことから、大腸菌の GAPDH が DNA 損傷の修復に必要であることが示唆された。

(想定されるデュアルユース)

本研究では、大腸菌の GAPDH が DNA 修復に関与することが示唆されたが、大腸菌以外の細菌においても同様の機能を持つ可能性がある。GAPDH をターゲットとした細菌感染症治療薬の開発が進められる一方で、治療薬は意図的に細菌の多様性を減らすことにも利用できる。ヒトや家畜の常在細菌といった有用な細菌の多様性の減少による健康被害は直接の原因を特定されずらく、低リスクで他者に損害を与えることができる。

#### 論文5

Orthogonal translation enables heterologous ribosome engineering in *E. coli*.

*Nature Communications*. **12**. 599 (2021).

<https://doi.org/10.1038/s41467-020-20759-z>

(一文概要)

バクテリアにタンパク質合成システムを 1 経路追加して持たせる。

(想定されるデュアルユース短文)

元の経路は生存に、新たな経路で有害たんぱく質を自在に生産

(論文概要)

本論文では、他の微生物由来の rRNA や r タンパク質で構成された異種リボソームを大腸菌内で発現させ、その活性評価方法を新たに開発した。また、この評価基準を用いて異種リボソームの活性向上への取り組みを行い、異種リボソームと大腸菌のリボソーム間で生じるサブユニット間の交換についても検討を行った。

本実験では、異種 16SrRNA が人工の mRNA のみに結合し、大腸菌の 16SrRNA は大腸菌の mRNA に結合してそれぞれ翻訳する、直交翻訳とよばれる方法を用いた方法で活性評価を行った。直交翻訳を用いることで、以前のような大腸菌の増殖能を測定する方法と比べて、異種リボソームが大腸菌の生存に関わる要素を排除し、翻訳能力そのものを測定することが可能となった。また、異種リボソームの rRNA をコードした遺伝子の RNA プロセシング部位を大腸菌の配列に置換し、rRNA 周囲に結合する r タンパク質の一部を大腸菌のものから異種リボソーム提供生物のものに変えることで、異種リボソームの活性を向上させた。また、大腸菌 16SrRNA との配列一致度が 92.9~97.0% である生物種由来の異種リボソームでは、大腸菌のサブユニットではなく、その種のサブユニット同士での結合が起こりやすいことが判明した。

(想定されるデュアルユース)

本論文での重要な点は、直交翻訳を用いることで、大腸菌の生存に関係なく異種リボソームの活性を測定することが出来る点であると考えている。この技術を応用することで、対象生物の生存能力を保ったまま、異種リボソームで効率的かつ持続的に有害物質を生産することが可能となるだろう。例えば、特定の生物に有害な物質を生産する菌を海洋中に放出し、その生物の数を減少させるといった使い方が考えられる。将来的には、我々のような多細胞生物の中でも、異種リボソームとmRNAを導入することで、生存に関係無く体内で有害な物質を生産することができると考えられる。

## 論文6

Large language models generate functional protein sequences across diverse families.

*Nature Biotechnology*. (2023).

<https://doi.org/10.1038/s41587-022-01618-2>

### (一文概要)

公開データベースの野生型タンパク質の配列を学習し、酵素機能を持ちつつ、既知酵素とは大きく配列が異なるタンパク質を生成

### (想定されるデュアルユース短文)

有害遺伝子の受託合成を防ぐためのスクリーニングシステムをすり抜けるタンパク質配列を生成

### (論文概要)

Transformerベースの自然言語生成モデルをタンパク質のアミノ酸配列の生成タスクにfine tuneし、タンパク質ファミリーにユニークに存在しているPfamIDを入力としてタンパク質の配列を出力するProgenというモデルを開発した論文。Progenのトレーニングデータには2億8000万のタンパク質配列とPfamIDの対応データセットが使用されており、それらの配列が巨大な表現空間に写像される。生成されたタンパク質は文脈の前後関係から尤もらしいアミノ酸配列かつPfamIDで指定されたタンパク質ファミリーと類似の機能を持った新しいタンパク質である。本論文では5種のlysozymeサブグループを生成し、先行の評価指標でランク付した上位配列の活性をin vivo実験で検証し、活性を持つ新しいタンパク質配列が生成されていることを確認した。

### (想定されるデュアルユース)

一般に遺伝子合成受託サービスにおいては、毒素が含まれていないことを誓約書として確認したり、タンパク質データベースを参照して、毒素タンパク質、有害、危険なタンパク質や配列をスクリーニングする仕組みが存在していることもある。

本論文において提案されたProgenを用いると、天然タンパク質と類似の機能を持つが、天然タンパク質と異なるアミノ酸で表現されているためそのス

クリーニングに引っかからないようなタンパク質を生成することが可能であり、これが実現すれば、遺伝子合成の段階でのバイオテロ防止が無効化される。

## 合成生物学に関する国際学生コンテストiGEMでのデュアルユース教育についての調査

2005年に始まったこのコンテストには、近年では毎年、世界中から300チーム、3000-5000人の学生が参加し、合成生物学分野において、有力な人材供給源となっている。「遺伝子工学のロボコン」と宣伝されることもあるこのコンテストでは、生物実験と他の領域との組合せが重要になる。すなわち、数理モデルやシミュレーションを取り入れた研究デザイン、産業化を見据えた活動、社会との倫理・安全面での関わり、これらが審査システムで重視されている。ゆえに、研究室に配属前が大半である学部学生は、これら多面的な活動を意識することになる。

現在では生物産業の基盤となりつつある合成生物学であるが、2000年代初頭のこの分野の立ち上げには、MITに長年勤めていた情報科学者Tom Knightが深くかかわったがゆえに、当時の生物系の研究者には稀な思想「標準化、オープンソース、数値シミュレーション」が根底に流れている。また、Tom Knightの傍にMITで新興技術のリスクガバナンスを専門としていたKenneth Oyeがいたことも、合成生物学分野が設立当初より、社会とのかかわりを重視することに繋がった。

iGEMには、審判委員会・Engineering委員会以外にも、起業委員会やsafety and security委員会をはじめ、社会との関りを担当する諸委員会がある。これら委員会の構成は、NPO法人iGEMの本部のメンバーが少数と、外部に本務を持つ専門家集団から構成される。この構成によって、iGEMは、新興技術と社会との関係性や、新興技術の学生への教育自体を研究する場ともなっている。

iGEM大会中に開かれる各種の展示ブースやワークショップのなかに、FBIによる連年の啓発活動も含まれる。過去においては、メインステージで、FBIのメンバーが「私たちは君たちを勧誘しに来た」と述べていたこともあった。我が国の政府が、情報科学のセキュリティのためにホワイトハッカーを登用するのと同じ発想と思われる。また、2022年のワークショップでは、2週間前にbioRxivに公開された、新型コロナウイルスで感染性の高いオミクロン株のスパイクタンパク質を毒性の高い従来株に組み込んだ仕事の紹介が、ボストン地域を管轄するFBIの担当者から語られることもあった。

数百の学生チームの中には、合成生物学の最新技術を応用しよう、という着想が生まれることも多い。特定の遺伝子を標的生物種のゲノム中に拡散するジーンドライブについて、CRISPR-Cas9系を用いた2015年の論文を受けて最初にiGEMに登場した2016年のプロジェクトへのiGEM本部の迅速な対

応を、安全委員会のメンバーが論文化している。当時、ジーンドライブに特化した公的な規制の枠組みは存在せず、iGEMでも将来的な規則策定のみが想定されていた段階で、秋の学生プロジェクト審査中にジーンドライブの試みに気づいたという。すると、iGEMの大会には関連研究者が多く参加しているために、すぐに打合せが始まり、その後の数か月間のミーティングをうけ、翌年春のiGEM活動開始前には、各種専門家の意見を反映した、大会としての事前審査ポリシーを策定することができた、と述べている。

#### D. 考察

##### 学生から見た論文情報のデュアルユースに関する検討

各種の合成生物学研究について、生産の効率化や細胞機能の制御といった合成生物学の得意とする方策が、有害物質の生産にも直結する、ということが、合成生物学について1年程度学んだ学生にとっては容易に着想できることが示された。

また、Chat-GPTのような、生成モデルに基づいた対話型AIが、タンパク質のアミノ酸配列も生成可能であることについて、学生は柔軟に適応し、その悪用についても想像している。言語や画像の生成AIが話題になると、半年以内には合成生物学におけるツールに関しての論文が発表されており、実際は言語や画像の生成と、生体分子配列の生成の研究は、同時並行的に進んでいたと考えられる。これらの研究には、大量のデータからの学習のために、億円単位の料金が想定されたため、その試みは検知しやすいかもしれない。一方、ひとたびその学習データが公開されると、カスタマイズ自体は比較的安価に、場合によっては、数十万円のコンピュータでも可能になるといわれている。さらに、遺伝子回路のデザインについては数年、微生物ゲノムゲノム全体のデザインについても10年程度で達成されることを見据え、安全性担保のためにどのような対策が可能かを、継続的に議論する必要がある。

##### 合成生物学に関する国際学生コンテストiGEMでのデュアルユース教育についての調査

iGEMの会場で、最新のプレプリントに記された懸念されうる研究について、FBI担当者を含め欧米の関連研究者達の生の声を聴くことができたように、新興技術に関する研究者が集う場としての意義もiGEMにはある。

また、学生のプロジェクトが、ジーンドライブの安全性を高めることを目的としたものであったことは、iGEMが社会とのかかわりの意識の啓発を学生に対して続けてきたことの結果であったかもしれない。

##### タンパク質の電子情報からの機能推定を行う人工知能に関する調査 (追加)

配列がわかっているものの機能が未知なタンパク質について、その機能を推定するためには、立体構造情報が極めて重要となる。昨年度の本研究で、タンパク質のアミノ酸配列から立体構造予測ツールについて調査した。この段階でも、生物学者一般に、確度の高い予測立体構造情報の入手を可能にした点で、大きな意義があったが、本年度は、多数の天然型タンパク質の配列を公開しているデータベースを学習し、機能を持ちつつ、天然タンパク質との類似性が低い配列を生成することも可能なツールについての論文公表が相次いでいる。

Madani, A., Krause, B., Greene, E.R. et al. Large language models generate functional protein sequences across diverse families.

*Nature Biotechnology*. (2023).

<https://doi.org/10.1038/s41587-022-01618-2>

Yeh, A.H.W., Norn, C., Kipnis, Y. et al. De novo design of luciferases using deep learning.

*Nature*. **614**, 774–780 (2023).

<https://doi.org/10.1038/s41586-023-05696-3>

Watson, J.L., Juergens, D., Bennett, N.R. et al. Broadly applicable and accurate protein design by integrating structure prediction networks and diffusion generative models.

*bioRxiv*. **12**, 09, 519842 (2022).

<https://doi.org/10.1101/2022.12.09.519842>

これら新規な生成ツールと、デュアルユースの既存の議論対象が組み合わせられることで、昨年の想定と同様に、どちらの方向の使用にも、大きな進展が予期される。すなわち、環境微生物を含めたゲノム情報の公開と入手、および、必要とされる塩基配列について、電子情報から遺伝子の合成を外注して物質としてのDNAを入手することである。この意味で、遺伝子合成に対する適切な監視を、より確実に行う必要がある。

#### E. 結論

合成生物学を学ぶ学生にとって、積極的にデュアルユースを検討する機会は稀であるが、ひとたび検討するならば、いくつものシナリオを想定する能力があることが分かった。タンパク質について、配列からの立体構造予測技術のみならず、新規配列生成AIの進展は著しく、デュアルユースの観点から、この技術と、配列情報の公開および遺伝子合成の受託合成のスクリーニングについての議論を引き続き行っていく必要がある。

F. 健康危険情報 なし

#### G. 研究発表

1. 論文発表 なし (執筆中、2023年度公表予定)
2. 学会発表

「合成生物学の国際学生コンテストiGEM等における教育事例」木賀大介. STS学会 2022年11月27日.

- |           |    |
|-----------|----|
| 1. 特許取得   | なし |
| 2. 実用新案登録 | なし |
| 3. その他    | なし |

H. 知的財産権の出願・登録状況