

Evo 2による全ゲノム生成と実験検証：Stanford・Arc Institute研究の技術的評価とバイオセキュリティ分析

エグゼクティブサマリ

2026年8月6日、Samuel H. King、Brian L. HieらStanford University／Arc Instituteの研究チームは、ゲノム言語モデルEvo 1／Evo 2を用いて完全なバクテリオファージゲノムを生成し、その一部が**宿主細胞内で感染サイクルを完遂し、増殖可能なファージ**として機能することを実験的に確認した研究を『Science』に発表した。2025年9月のbioRxivプレプリントを経て査読された成果であり、主たる対象は約5.4 kbの一本鎖DNAファージΦX174と、その近縁Microviridaeである。モデルはMicroviridaeに追加学習され、ΦX174を設計テンプレートとして候補を生成、計算フィルタリング後に302候補が合成対象となり、公開資料で確認できる範囲では285設計が実験評価に到達、そのうち16種類、すなわち285に対して5.61%が機能的ファージとして回収された。¹

この研究を「AIが無から生命を作った」と表現するのは不正確である。成功した配列はΦX174近縁の既知の進化空間を足場にしており、モデルによる生成の後に、**計算選別、化学合成、DNAアセンブリ、宿主への導入、増殖選択、配列確認**という強力な現実世界の選択過程が存在する。一方で、「単なるランダム変異」と片づけるのも適切ではない。機能したゲノムは最も近い天然ゲノムから67～392か所の変化を持ち、Evo-Φ36では遠縁ファージG4由来の短いDNA包装タンパク質Jが、他のゲノム変化と共存して機能することがcryo-EMで確認された。²

最も重要な定量結果は、16/285というヒット率そのものだけではない。Evo-Φ69は3回の競合実験すべてで優勢となり、6時間後のcumulative fold changeが16～65倍に達したとプレプリントで報告された。ただし、これは「天然ファージより65倍速く増殖する」という意味ではなく、特定の競合アッセイにおける相対的存在量変化の指標である。

耐性突破実験の系統数は**版により記載が異なる**。2025年プレプリントおよびArc Institute解説では、ΦX174耐性E. coli 3系統（CR1、CR2、CR3）を1～5継代で突破したとされる。一方、Science掲載時のAAAS配信資料では2系統とされ、公開されている個別データではCR1を1継代、CR2を2継代で突破している。本レポートではScience掲載時の情報を基準に2系統と整理し、プレプリント／Arcとの版間差を明記する。Science公開要旨は系統数を明示しておらず、差異の理由は未確認である。ΦX174単独は失敗し、回収された耐性突破ファージは複数のAI設計ゲノムが実験中に組換わったモザイクであった。3, 51

研究のエビデンスは相当に強い。『Science』査読、GenBank、GitHub、公開モデル、EMDBに加え、配列機能、競争適応度、宿主域、構造、耐性突破を複数の測定系で検証している。ただし、独立研究室によるend-to-endのwet-lab再現は2026年8月10日時点で未確認である。原論文内では単純な統計的・進化的ベースラインとの直接比較が不足する一方、Black et al. (2026、プレプリント) の独立再解析は、Evo 2のモデル尤度が単純な生物学的指標以上に生存性を予測し、ランダム変異や継代進化に対して相当な設計効率を示すと報告した。ただし、その効率の多くは天然配列近傍の探索と後段フィルタに由来し、生成物も系統学的には天然ゲノムに近い。約5.6%は事前フィルタ後の値であり、対象ゲノムも非常に小さい。17, 52

バイオセキュリティ上は、**今回作製されたΦX174様ファージそのものの直接的な人健康リスクと、研究が示した「能力」のリスクを分離して評価する必要がある**。実験では非病原性E. coliを用い、16ファージはE. coli Cと近縁のE. coli Wに限定された宿主域を示し、他の6菌株では増殖しなかった。したがって直接リスクは低い。しかし、「完全な機能性ウイルスゲノムを生成モデルで探索できる」という能力は、将来別のウイルス群に転用され得る。Evo 2およびファージ用追加学習モデルが公開されているため、著者らが強調する「ヒト病原ウイルスを訓練データから除外した」というガードレールは重要ではあるものの、**オープンウェイトモ**

デルに対する恒久的な技術的封じ込めとは評価できない。Stanford自身も、改変版が悪意ある主体に利用され得る可能性を認めている。⁵

このため本研究の政策的含意は、「Evo 2を禁止するか公開するか」という二択ではなく、モデル公開、病原体特化ファインチューニング、DNA合成、実験室でのreboot、外部放出という能力スタックの各段階に異なる管理を置くことにある。WHOはデュアルユース研究を研究者だけでなく、研究機関、資金提供者、出版社、企業、政府が共同で管理すべきものとしている。米国では2026年7月に新しいHigh-Risk Life Sciences Research政策が承認され、宿主域・病原性・伝播性などを高める研究をリスクベースで扱う方向が明確化している。また合成核酸については、HHSの2023年ガイダンス（任意の推奨基準）と、連邦資金条件として運用される2024年OSTP Frameworkが重要なチョークポイントとなる。ただし後者は2025年大統領令に基づき改訂・置換が予定されている。6, 53, 34

総合評価：科学的には「ゲノム言語モデルによる完全ゲノム設計を、実体のある感染性・増殖性ファージ系まで接続した」という明確なマイルストーンである。一方、「汎用的なウイルス設計器」「臨床ファージ治療」「AIによる生命創造」と呼べる段階ではない。現在の成果は、高い概念的重要性、良好なデータ公開性、高い計算再解析可能性と未確認の独立wet-lab再現性、低い生成効率、そして将来的には無視できないデュアルユース能力を併せ持つ。

一次ソースと証拠基盤

一次資料一覧

種別	一次ソース	内容と本レポートでの位置づけ	URL
査読論文	King SH et al., “ Generative design of bacteriophages with genome language models ”, Science 393, 2026, DOI 10.1126/science.aec2657	最終的な査読版。完全ファージゲノム生成、16機能ファージ、cryo-EM、耐性突破実験を報告。受理日は2026年6月2日。データ可用性欄でGenBank等への登録も示す。 ⁷	https://doi.org/10.1126/science.aec2657
プレプリント	King SH et al., “ Generative design of novel bacteriophages with genome language models ”, bioRxiv, 2025-09-17, DOI 10.1101/2025.09.12.675911	査読前版。競合実験の定量値など、Scienceの短い公開要旨より詳細な結果を検索可能。正式版ではタイトルから“novel”が外れた。 ⁸	https://www.biorxiv.org/content/10.1101/2025.09.12.675911v1
Stanford プレスリリース	Stanford Report, “ AI designs a novel E. coli killer ”, 2026-08-06	研究者本人の説明。全ゲノムを左から右へ一度に生成、数千候補から絞り込み、約300設計の実験検証、16ファージ、オープンソース方針を説明。 ⁹	https://news.stanford.edu/stories/2026/08/evo-2-ai-tool-e-coli-killer-bacteriophages

種別	一次ソース	内容と本レポートでの位置づけ	URL
研究機関 技術解説	Arc Institute, “ How We Built the First AI-Generated Genomes ”	追加学習データ 14,466 Microviridae、独自 アノテーション、 285設計の実験、宿 主域、67～392変 異、Evo-Φ36、耐性 突破、封じ込めを具 体的に説明。一次的 な研究チーム解説と して極めて有用。 10	https://arcinstitute.org/news/hie-king-first-synthetic-phage
コード・ データ	ArcInstitute/evo2 GitHub, phage_gen	解析コード、ゲノム アノテーション、設 計フィルタ、競合解 析、アセンブリ設 計、生成配列・ viable/nonviable配 列のデータ構成を公 開。 11	https://github.com/ArcInstitute/evo2/tree/main/phage_gen
モデル	evo-design/evo-2-7b-8k-microviridae	Evo 2 7B baseを Microviridaeに追加 学習したモデル。 10,240-token context、12,000 iteration checkpoint として公開されてい る。 12	https://huggingface.co/evo-design/evo-2-7b-8k-microviridae
配列デー タ	NCBI GenBank	Scienceのデータ可用性 欄は、生成ファージの注 釈付きゲノムを PZ509151-PZ509168 と して登録したとする。こ れは18アクセスンで ある。ただし、16種類の 初期機能ファージと、耐 性突破実験等で得られた 派生配列との正確な対応 関係は、各GenBankレ コードの個別確認を要す る。単純に「18種類の機 能ファージ」と解釈して はいけない。 13	https://www.ncbi.nlm.nih.gov/nuccore/ (accession: PZ509151-PZ509168)

種別	一次ソース	内容と本レポートでの位置づけ	URL
構造データ	EMDB EMD-77390 / fitted model 36CQ	Evo-Φ36関連のcryo-EMデータ。EMDBは2.76 Åのsingle-particle mapとfitted model 36CQを掲載。 ¹⁴	https://www.ebi.ac.uk/emdb/EMD-77390
補足資料	Science Supplementary Materials	Science論文ページにはSupplementary Materialsが付随する。補足資料そのものの完全なページ単位照合は、本調査時のScienceサイト側アクセス制限のため 未確認 。したがって、以下で補足資料だけに依存する細かな反応条件等は断定していない。	https://www.science.org/doi/10.1126/science.aec2657
基盤モデル論文	Evo 2 / Nature 2026	Evo 2そのもののアーキテクチャと長文ゲノムモデルの背景。GitHubはEvo 2がOpenGenome2上で自己回帰学習され、最大1 Mb級の文脈を扱うことを説明している。 ¹⁵	https://doi.org/10.1038/s41586-026-10176-5

ここで重要なのは、論文だけでなく「失敗配列」も含む解析資産が公開されている点である。GitHubの `phage_gen` には、全生成ファージ、viable、nonviable、天然ΦX174様株、実験室進化株に加え、競合シーケンス解析、変異分類、アノテーション、ゲノム設計フィルタリングのコードが明示的に整理されている。これは、成功例16だけを公表する研究よりも、機械学習側の再解析にかなり有利である。¹¹

ただし、データ公開性と独立再現性は同義ではない。2026年8月10日時点で、別研究グループが公開ウェイトから候補を再生成し、独立にDNAを合成して16/285に近い機能回収率を再現した査読報告は確認できなかった。その意味で本研究は、「計算解析の再現可能性は比較的高いが、end-to-endの**独立wet-lab再現は未確認**」と評価するのが妥当である。

外部評価資料

同号のScienceには、Johns Hopkins Center for Health SecurityのThomas V. InglesbyとMoritz S. HankeによるPerspective “AI-designed viral genomes” が掲載され、「機能するウイルスゲノムを生成できること自体が、緊急性の高いbiosafety／biosecurity問題を生む」と位置づけた。¹⁶

FraserとBondy-DenomyのPREreviewは、構造・機能実験とdesign-build-test系を高く評価する一方、原論文内部では、コンセンサスゲノム、読み枠やGC等を保存したランダム変異、祖先配列再構成、ファミリー統計モデルなどとの直接比較が不足すると指摘した。¹⁷

これに対し、Black et al. (2026, bioRxiv) は公開データを独立に再解析し、Evo 2のモデル尤度が単純な生物学的指標以上に実験的生存性を予測し、ランダム変異や継代進化に比べて相当な設計効率を示すと報告した。⁵²

ただし、その効率のかなりの部分は、既知の天然配列近傍を探索することと、後段の選別フィルタに由来するとされ、生成物も系統学的には天然ゲノムに近かった。したがって「原論文内の直接比較は弱い」という留保は残る一方、公開データに基づく後続再解析はEvo 2の設計上の有用性を一定程度支持している。査読前プレプリントであり、異なるウイルス群への一般化は未確認である。52

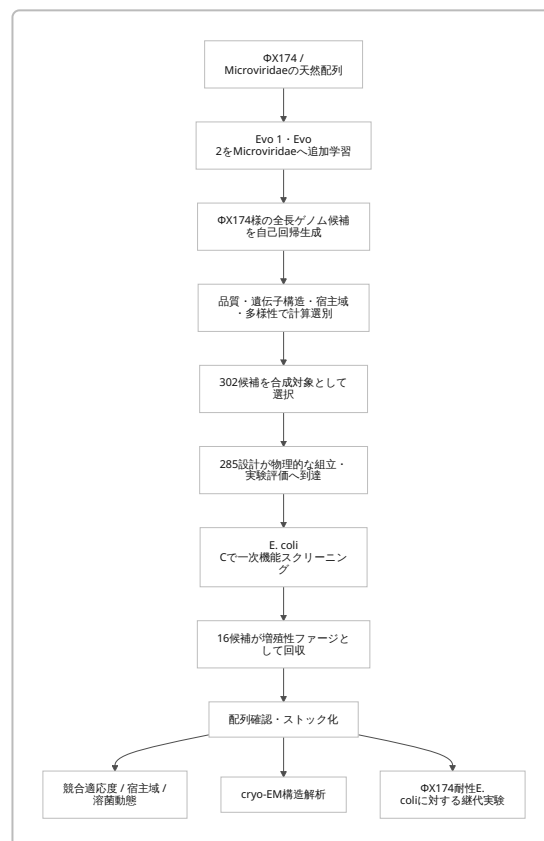
実験設計と検証フロー

生成から機能確認まで

ΦX174は約5.4 kbと小型だが、11遺伝子の多くが重複しており、同じ塩基が複数のタンパク質・調節要素に同時に制約される。そのため、単純な「一遺伝子ずつ最適化」では扱いにくく、全ゲノム設計の試験台として合理的である。Arc Instituteによれば、既存の一般的遺伝子予測ツールでは11遺伝子すべてを正しく拾えず、研究者らはORF探索とファージタンパク質相同性検索を組み合わせた独自アノテーション系を構築した。基本品質フィルタとして、少なくとも7個のΦX174天然タンパク質に対応する予測ヒットを要求した。 18

研究チームはEvoの基盤モデルをそのまま使っただけではない。Arcの一次解説によると、**99% identityで冗長性を減らした14,466件のMicroviridae配列**を用いてsupervised fine-tuningを行い、ΦX174様ゲノムを生成しやすいモデルへ専門化した。したがって、この成果は「汎用Evo 2に自然言語で“ファージを作れ”と頼んだ」という種類の実験ではなく、**ファージ系統データによる専門化+テンプレート誘導+計算選抜**を組み合わせた設計である。 19

Stanfordによれば、ゲノムは途中で人手で遺伝子を差し込むのではなく、短いΦX174由来の開始情報を足場に、モデルが**一回の自己回帰パスで全長候補を生成**した。候補はその後、配列品質、遺伝子構造、宿主特異性、進化的多様性などでフィルタされた。宿主域については、ΦX174でE. coli認識に関わるスパイクタンパク質に近い性質を残すことが設計制約として用いられた。 20



このフローは、Stanford、Arc Institute、Scienceの公開情報を統合した概念図である。302と285の差については、285が実際に評価された設計数であることは一次資料で確認できる一方、残る17件についての個別の製造失敗理由は、今回確認できた一次資料本文だけでは完全には検証できず**未確認**である。 ²¹

ウェットラボ工程

研究チームの一次解説で確認できる実験系は、**合成DNA断片のアセンブリ → competent E. coli Cへの導入 → 宿主増殖阻害の追跡 → 機能候補の配列確認 → ファージストック化**という構成である。ArcはDNAアセンブリにGibson assemblyを使い、96-well形式でOD600を追跡し、感染成立時に2〜3時間程度で宿主濁度が低下することを利用したと説明している。285設計をこのスクリーニング系で評価した。 ¹⁸

ここで「発現」という語には注意が必要である。本研究は個々のウイルスタンパク質を別々に発現させて組み立てたのではなく、**合成した全ゲノムを細菌宿主に導入し、そのゲノムが宿主内で転写・翻訳・DNA複製・カプシド形成・包装まで自律的な感染サイクルを再起動できるか**を評価している。したがって実験の核心は、単一タンパク質の活性確認ではなく、ゲノム全体の相互整合性である。 ²²

デュアルユース上の理由から、本レポートでは、公開リポジトリに存在する**実配列、生成用プロンプト、実行可能な生成コマンド、DNA導入の操作条件、反応組成など、ウイルスゲノムをそのまま再構成するための操作パラメータは再掲しない**。再現性評価に必要な科学的構造は記載するが、実装レシピそのものと科学的分析は分けて扱う。

実験ステップと検証強度

段階	研究で行われたこと	主な選択・コントロール	再現性に関する評価
ファミリー追加学習	14,466 Microviridae配列でモデルを専門化。 ¹⁸	天然Microviridaeの進化的制約を学習	モデルウェイト公開で計算上は比較的高い
全ゲノム生成	ΦX174をテンプレートに、全長を一方方向に生成。 ²³	ΦX174様でありつつWTに近すぎない候補を探索	コード・モデル公開。ただし原生成全件数の一次資料間の完全整合は未確認
アノテーション	独自ORF+相同性パイプライン	≥7の天然ΦX174タンパク質ヒットを基本品質条件。 ¹⁸	GitHubに解析系公開
宿主域設計	スパイクタンパク質を使ってE. coli C tropismを保持	ΦX174由来の宿主認識特性	16/16がC/W、他6菌株陰性。 ¹⁸
DNA製造	選抜候補を化学合成・アセンブリ	計算選抜後のみ物理合成	302候補中285が評価段階へ。個々の失敗理由は未確認
genome reboot	合成ゲノムをE. coli Cへ導入	WT ΦX174など既知ファージとの比較	実験条件の完全な独立再現は未確認
一次アッセイ	96-wellで宿主増殖阻害を追跡	非感染/非機能設計に対するOD挙動	285から16を回収。 ¹⁸
配列検証	16候補をsequence-verifyしストック化	設計配列との整合性	GenBankおよびGitHubが強み。 ²⁴

段階	研究で行われたこと	主な選択・コントロール	再現性に関する評価
適応度	ファージ間競合とシーケンススペース相対量解析	ΦX174などとの競争	GitHubにSNV read assignment / fold-change解析コード。 ¹¹
構造	Evo-Φ36をcryo-EM解析	既知ΦX174構造との比較	EMDB 2.76 Å mapが公開。 ¹⁴
耐性突破	Science掲載時：耐性E. coli 2系統（CR1、CR2）にAI設計ファージ集団を継代。プレプリント／Arcは3系統。	ΦX174単独と比較	CR1：1継代、CR2：2継代で突破。版間差あり。3, 51

封じ込めについてArcは、専用biosafety cabinet、専用廃棄手順、装置を封じ込め区域外へ出さない運用、非病原性E. coliのみの使用を説明している。一方、公開一次資料で具体的なBSL番号、IBC承認番号、リスク評価書全文までは確認できず、これらは未確認である。²⁵

定量結果と統計的評価

「300から16」の意味

一般報道では「約300のAI設計から16を選抜」と簡略化されるが、実験統計としては工程を分ける必要がある。公開情報を合わせると、**302が物理合成に進む候補として選ばれ、285が実験評価に到達し、16が機能的ファージとして回収された**という構造である。Arc一次資料は285の試験と16のsequence-verified candidatesを明記し、302という上流数はScience関連資料・二次解説でも一貫して報告されている。²¹

したがって、

$$\text{実験評価後の機能ヒット率} = \frac{16}{285} = 5.61\%$$

である。302を分母にしたend-to-endの選抜候補ベースでは、

$$\frac{16}{302} = 5.30\%$$

となる。また、302から285まで到達した比率は94.37%である。これらは公表された整数から本レポートで計算した値であり、論文そのものの表記とは区別すべきである。²¹

16/285を二項比率とみなし、探索的にWilson法で95%信頼区間を計算すると**約3.48～8.92%**となる。ただし、これは「Evoが任意のプロンプトから5.6%の確率で機能ウイルスを出す」という推定ではない。285候補は大量の生成物から**事前の計算フィルタと人間の選択を通過した非ランダム標本**だからである。したがって、5.6%をモデル本体の無条件成功確率として解釈するのは統計的に誤りである。²²

原始生成数については、StanfordとArcの一次解説は「thousands」と表現する一方、一部メディアは約70万候補と報じている。今回参照できたScience補足資料の完全本文でその総数を直接照合できていないため、**「約70万」は未確認情報として扱う**。したがって16/700,000のような数字を本研究の“真の成功率”として示すこともしない。²⁶

機能ファージの多様性

16機能ゲノムは、最も近い既知天然ゲノムに対して**67~392変異**を持っていた。Evo-Φ2147は近縁phage NC51に対して93.0% ANI (average nucleotide identity) で、Arcは一部のウイルス分類閾値では別種相当になり得ると説明している。また、16中13のゲノムには、既知の天然配列集合で見つからない変異が含まれた。¹⁸

ただし、「自然界に存在しないウイルス」と断言する場合には注意が必要である。データベースに未登録であることと、地球上の自然集団に実在しないことは同値ではない。厳密には、**既知の天然配列にはない組合せ・変異を含む人工設計ファージ**と表現するのが適切である。これはデータベース観測範囲からの論理的帰結であり、自然界全体の非存在を証明したものではない。¹⁸

適応度の評価

プレプリントの公開検索本文では、Evo-Φ69が**3回のcompetitionすべてで他ファージを上回り、6時間後のcumulative fold changeが16~65×**だったと報告されている。GitHubには競合実験のシーケンスをQCし、SNVベースで各ファージにreadを割り当て、fold changeを計算する解析コードが存在する。⁸

ここで「16~65×」は、生物学的fitnessの万能スカラーでも、burst sizeが65倍という意味でもない。競合環境での相対的頻度変化を積算したアッセイ指標であり、培地、宿主状態、初期混合、時間窓などに依存する。Science Media CentreでSimon Clarkeは、16ファージのうち**3 variantsが親ΦX174を上回った**と研究結果を要約している。²⁷

したがって研究が支持するのは、「AI設計の一部が**特定のin vitro条件で天然テンプレート以上の競争性能を持ち得る**」までである。「臨牀的に強力」「自然界より一般にfitnessが高い」とする外挿は支持されない。

構造的検証

Evo-Φ36は、遠縁のG4ファージに由来するDNA包装タンパク質Jに対応する構成を持つ。G4側のJは25アミノ酸で、ΦX174の38アミノ酸より短い。Arcは、過去の単純な合理的置換ではうまくいかなかった組合せが、Evo設計のゲノム文脈では機能したと説明している。cryo-EMではこの短いJタンパク質が異なる向きでカプシド内に配置されることが確認され、EMDBには2.76 Åのマップが登録されている。²

この結果は本論文で特に強い部分である。単に「細菌のODが下がった」だけでなく、**AIが生成した複数の変化が物理的なウイルス粒子の組立互換性につながったことを原子レベルに近い構造データで検証している**ためである。もっとも、これが「Evoが補償変異の因果機構を理解していた」ことを証明するわけではない。モデルが統計的制約を利用して成功配列を探索できることと、生物物理学的因果を明示的に理解していることは別問題である。Jordi García Ojalvoも、この研究は「なぜ一部のゲノムが生き残り、他が失敗するのか」という機構理解そのものは提供しないと指摘している。²⁷

宿主域と耐性突破

16の機能ファージはいずれもE. coli Cと近縁E. coli Wで機能し、**その他6つの試験菌株では増殖しなかった**。これはテンプレート／スパイクベースの宿主域制約が、少なくとも試験パネルでは機能したことを示す。ただし試験した宿主範囲は8菌株程度に限られ、環境中の全細菌に対する宿主域を証明したものではない。¹⁸

耐性実験の数値は版により異なる。2025年プレプリント／Arc解説ではΦX174耐性E. coli 3系統 (waa operon変異) を1~5継代で突破したとされる一方、Science掲載時のAAAS配信資料では2系統とされる。公開されている個別データに基づき、本レポートではCR1を1継代、CR2を2継代で突破した結果を主要結果として整理する。ΦX174単独は成功せず、回収ファージは2~3種類のAI設計由来要素を組み合わせたモザイクゲノムであった。3, 51

これは臨床的には興味深いが、解釈のポイントは「AIが耐性菌専用薬を直接設計した」ではない。むしろ、AIが進化に投入できる遺伝的多様性を増やし、その後の組換え・選択が耐性突破解を探索したという結果である。

定量結果の要約

指標	結果	解釈上の注意
合成候補	302	既に計算フィルタ後の候補。生のモデル出力ではない。 ²⁸
実験評価到達	285	Arc一次資料で明記。 ¹⁸
機能ファージ	16	growth inhibition後、sequence verification、stock化。 ¹⁸
16 / 285	5.61%	本レポート計算。事前選抜標本のhit rate
16 / 285 Wilson 95% CI	3.48–8.92%	本レポートによる探索的統計。論文記載値ではない
最近傍天然ゲノム との差	67–392変異	「完全de novo」の証拠ではなく、既知進化空間内外の組合せ探索を示す。 ¹⁸
Evo-Φ2147	93.0% ANI vs NC51	高い新規性だが、依然として近縁ファージ空間にある。 ¹⁸
Evo-Φ69競争	16–65× cumulative fold change / 6 h	相対頻度指標であり、65倍の一般的増殖速度を意味しない。 ²⁹
宿主域	C、Wで陽性、他6株陰性	試験パネル外の宿主域は未確認。 ¹⁸
resistant E. coli	2系統 (CR1、CR2)	waa operon変異。Science掲載時資料を基準。プレプリント／Arcは3系統。
耐性突破	1継代 (CR1)、2継代 (CR2)	モザイク化・追加進化を伴う。版間差に注意。3, 51
cryo-EM	2.76 Å	EMD-77390。 ¹⁴

一部二次解説は、16機能ゲノムのうち9が生成時の配列のまま機能し、7がreboot後に追加変異を獲得したと報じている。しかし、この内訳を今回アクセスできたScience一次本文・Arc一次解説から直接照合できなかったため、9/7の内訳は本レポートでは「未確認」とする。²⁸

同様に、主要図の各エラーバー、全 biological replicate 数、個別比較に使われた有意差検定、multiple-testing補正の詳細は、Science Supplementary Materialsの完全本文をページ単位で確認できなかったため未確認である。したがって、競合アッセイの強い効果量は評価する一方、細かなp値については本レポートでは断定しない。

バイオセーフティとデュアルユース評価

直接的危険性と能力リスクを分離する

今回の実験材料そのものについて、研究チームは非病原性のE. coli C等を用い、専用封じ込め設備・廃棄手順を採用したと説明している。16生成ファージの宿主域も試験範囲ではE. coli C/Wに制約されていた。したがって、今回得られた16ファージの直接的なヒト感染リスクは低いと評価できる。¹⁰

ただし、Arcが「人健康へのリスクはない」と強く表現している点は、著者機関によるリスク評価である。8株程度の宿主域試験はヒト感染性がないことを支持するが、全環境宿主、長期進化、水平伝播、複雑な微生物生態系での影響まで実証したものではない。このため、本レポートでは「ゼロリスク」ではなく**直接リスクは低い**が、**環境リスクには残余不確実性があると表現する。** 18

より重要なのはcapability riskである。研究は、ゲノム言語モデルが「個別タンパク質」ではなく、**宿主細胞内で感染サイクルを完遂して増殖可能なウイルス全ゲノム**の候補探索器として機能し得ることを示した。同号のInglesbyとHankelは、この能力が成立したこと自体をbiosafety/biosecurity上の転換点として扱っている。16

「病原体を訓練データから除く」ことの限界

Arc Institutelは、Evoの訓練データからヒトウイルス配列を意図的に除外したため、ヒトウイルス生成を防止できると説明している。これは**研究チームのガードレール設計としては有意義**である。 25

しかし、この主張を「Evo 2という技術はヒト病原体設計に転用できない」という意味に拡張することはできない。ファージ専用Evo 2自体が公開ウェイトとして配布されており、Stanfordも「modified versions could be used by bad actors」と認めている。したがって、訓練データ除外は**現行モデルの初期状態を安全側にする措置であって、公開ウェイトの将来的再学習・別用途化まで強制するアクセス制御ではない**、というのが技術的に妥当な評価である。 30

この点で、安全性は三層に分けて考えるべきである。

モデル層では、危険なウイルス群の学習除外、危険能力の評価、モデルウェイトへのアクセス政策が問題になる。**合成層**では、デジタル配列を物理DNA/RNAへ変換する際の注文者・配列スクリーニングが重要になる。**実験層**では、宿主、封じ込め、実験目的、外部放出の有無についてIBC等の制度的レビューが必要になる。Evo研究は、モデルだけを規制しても、あるいはDNA合成だけを規制しても十分ではないことを示している。これは本研究と政策文書を合わせた分析上の推論である。 31

合成可能性とチョークポイント

今回の研究が示したのは、約5.4 kb級のファージゲノムについては、デジタル候補から化学合成・アセンブリを経て実体化できるということである。Stanford自身、DNA合成は依然として高価で、全ゲノムを細菌へ移す工程も容易ではないと説明している。したがって2026年現在、**AI出力が即座にウイルスになるわけではなく、物理合成と実験能力が大きな摩擦点**である。 32

米HHSの2023 Synthetic Nucleic Acid Screening Guidancelは、従来の規制対象病原体だけでなく病原性・毒性に寄与するSequences of Concernを広く対象とし、50 nt未満の短い配列も含むスクリーニング、一本鎖・二本鎖DNA/RNA、ペンチトップ合成装置、注文者確認まで含めるよう勧告した。**ただし、このガイダンスは新たな規制を設けるものではなく、遵守は任意のベースライン基準である。** 33, 53

2024年OSTP Frameworkは、連邦資金を受ける生命科学研究で適合プロバイダーからの調達を条件とした。もっとも、2025年のExecutive Order 14292は同Frameworkの改訂・置換を命じ、2026年8月10日時点のHHS公開ページも新しい枠組みの公表後に更新するとしている。したがって、2024 Frameworkを現在の最終確定制度として引用するのは不正確である。 34

本研究は核酸合成スクリーニングの回避を実証していない。最も離れたEvo-Φ2147でも最近傍天然ファージに93.0% ANIであり、Black et al.の再解析も生成物が系統学的に天然ゲノムへ近いと評価した。したがって、将来、既知配列から大きく逸脱する生成配列が得られた場合に文字列一致だけでは不十分となり得るという点は、本研究からの直接証明ではなく政策的推論である。中長期には推定機能・毒性・宿主域を含むscreeningが重要となる。18, 52

リスク評価マトリクス

リスク	現時点の蓋然性	潜在影響	総合評価	根拠・主要な管理策
今回の16ファージによる直接的ヒト感染	低	低～中	低	非病原性E. coli対象、試験宿主域がC/Wに限定。 18
意図しない細菌宿主への感染・環境影響	低～中／不確実	中	低～中	6追加株で陰性だが、全環境微生物群を網羅していない。環境放出前の広範host-range/ecology試験が必要。 18
病原性ウイルス群への方法論転用	現在は中程度の技術障壁	高～極高	中～高の能力リスク	オープンモデルの再専門化可能性。研究者自身もmodified versionsの悪用可能性を認める。 30
将来の高乖離生成配列に対する合成スクリーニング上の課題	中長期に中	高	中～高	本研究は回避を実証しておらず、生成配列も天然ゲノム近傍。将来仮説として、機能ベースscreeningと顧客確認を強化。52, 53
産業・農業用細菌を狙う悪意あるファージ設計	低～中	中～高	中	人病原体以外にも経済・農業への標的化が理論上あり得る。WHOは人・動物・農業・環境を含むOne Health型管理を推奨。 37
自動化design-build-testによる能力スケールアップ	現在は低～中、上昇傾向	高	中～高の将来リスク	現在は合成・wet-labがボトルネック。モデル性能向上と合成費低下で摩擦が減る可能性を著者も指摘。 25
誤った「AI設計＝安全」判断による事故	中	中～高	中	モデル制約と実際の表現型は一致しない場合がある。実験的host-range/phenotype検証を維持すべき。

悪用シナリオについては、政策上は「ヒト・動物・植物病原体に同じwhole-genome design paradigmを移す」「産業上重要な細菌を狙うファージを探索する」「将来、既知配列から大きく逸脱する生成配列が得られた場合に既存スクリーニングが十分か」といった能力クラスを想定すべきである。**本研究はスクリーニング回避を実証しておらず、具体的な回避法を示すものでもない。**また、実行可能な配列設計法、実験条件、screening回避法を公開すること自体がデュアルユースリスクを増すため、政策分析では脅威モデルと実装レシピを分離する必要がある。

倫理・政策と専門家評価

現行ガバナンスとの整合

米国では、本研究発表直前の2026年7月20日に**United States Government Policy for Stopping High-Risk Life Sciences Research**が承認され、7月28日に公表された。HHSの政策史によれば、2024年のDURC/PEPP統合政策は「developed, but not implemented」であり、2025年Executive Orderを経て2026年政策に置き換えられた。したがって2026年8月現在の評価では、この新政策を基準にすべきである。 38

新しいHHS/USGの説明は、重大な社会的結果をもたらし得るdangerous gain-of-functionの例として、**harmful consequencesの増強、治療・予防への抵抗性、安定性・伝播性・散布能力、host range/tropismの変更**などを列挙している。³⁹

本研究はΦX174様の非ヒト病原ファージを用いており、直ちにこの高リスクカテゴリーへ該当すると断定できる根拠はない。一方、研究内容にはファージのfitness、多様性、host range、抵抗性突破といった、**政策が一般論として注視する表現型と概念上重なる要素**が含まれる。特定の研究が政策上正式にDGOFに該当するかについて、米当局による本論文固有の公式判断は**未確認**である。したがって「この研究は米国で禁止対象」「完全に規制外」のどちらも現時点では断定できない。⁴⁰

また、この実験の大半は2025年のプレプリント以前に実施されているため、2026年7月に成立した政策を**過去の実験に遡及的に適用して研究の適法性を判断することは適切でない**。現在の政策はむしろ、同種研究を今後どう扱うべきかというprospective governanceの基準として重要である。⁴¹

WHOの2022 **Global guidance framework for the responsible use of the life sciences** は法的拘束力のある国際条約ではないが、デュアルユースリスクを研究者個人の判断だけに任せず、政府、研究機関、研究者、資金提供者、出版社、民間企業、セキュリティ関係者を含む「shared responsibility」として扱う。この枠組みは、AI、synthetic biology、genome editing等を含む新技術のforesightにも適用される。³⁷

専門家コメント

Thomas Inglesby / Moritz Hanke, Johns Hopkins Center for Health Security

同号のScience Perspectiveは、機能的なウイルス全ゲノムを生成できることが実証された点を、緊急性のあるbiosafety/biosecurity課題として扱っている。これは「今回のΦX174が危険」というより、**技術クラスが新しい能力境界を越えたことへの懸念と読むのが適切**である。¹⁶

Simon Clarke, University of Reading

データ品質を高く評価し、16のAI設計ファージがE. coliへ感染し、3 variantsが親株を上回ったこと、耐性突破を示したことを重要視した。その一方で、よりfitnessの高い非天然ウイルスを作り得ることが示されたため、倫理・規制・安全性の問題は重大であり、今回の研究者が設けたguardrailsを他の研究者が必ず採用する保証はないと指摘した。²⁷

Patrick Cai, University of Manchester

「予測」から「完全な機能ゲノム生成」へ進んだsynthetic genomics上のマイルストーンと評価した。ただし現在は小型ファージであり、より大きく複雑なゲノムへ拡張する際にも**計算予測を厳密な実験検証と結び付けることが必要**だとしている。²⁷

Jordi García Ojalvo, Pompeu Fabra University

成功効率が低いことを重要な制約として挙げ、数千候補から機能したのが16だけである点を、ある意味ではゲノムモデルの「hallucination」と表現した。他方、各生成ゲノムをwet-labで一つずつ試さなければならず、現状の低効率は自動的な危険ゲノム量産への実務的障壁にもなると評価している。また、この手法は「なぜ機能するのか」の機構理解を自動的には与えないとも指摘した。⁴²

Marc Güell, Pompeu Fabra University

実験設計を非常に高い水準と評価し、単一タンパク質設計から、10以上のタンパク質が協調する数kbの合成ファージ全体へ進んだ点を、generative biological designの重要な転換として位置づけた。⁴³

Fraser / Bondy-Denomy, PREREview

wet-labの構造・機能検証自体は説得力があるものの、「この成果に大規模ゲノム言語モデルがどれほど必要だったか」の比較が不足していると指摘した。これは非常に重要な批判である。成功した配列が天然近縁空

間に比較的近い場合、コンセンサス+制約付き変異、祖先配列、より単純な統計モデルでも同程度のhit rateが得られる可能性を排除できない。 17

規制当局による本研究固有のコメントについては、2026年8月10日時点で、HHS、NIH、WHO等がKing et al.の論文そのものについて正式なリスク判定を出したことは未確認である。したがって本レポートでは、規制当局については論文への推測コメントではなく、現在有効・公表済みの政策文書のみを根拠として扱った。

推奨される公開・アクセス政策

研究成果の公開はall-or-nothingで扱うべきではない。WHOのshared-responsibility原則、米国のhigh-risk research政策、核酸合成screeningの考え方を合わせると、**成果の社会的検証可能性を保ちながら、実行能力を不必要に拡散しない段階的公開**が適切である。 45

本レポートでは次の管理モデルを推奨する。

第一に、**論文本文、集計結果、失敗率、評価コード、非危険なbenchmarkは原則公開**する。これは第三者が「成功例だけを選んでいないか」「モデル優位性が本物か」を検証するために必要である。

第二に、**病原体に直接転用可能なファインチューニング成果、危険な宿主域変更を目的とするモデル、実行可能なwet-lab recipeについては、リスク評価後の段階的アクセス**が望ましい。現在のEvoファージモデルは既に公開されているため、これは将来のより高リスクモデルに対するprospective recommendationである。 30

第三に、モデル安全策だけに依存せず、**合成DNA/RNA側に独立したscreeningを維持**する。モデルの訓練データ除外が破られたり別モデルが使われたりしても、物理化の段階で顧客・配列・用途を確認できるからである。 36

第四に、AI-bio研究では通常のIBC審査に加え、モデルがもたらす**capability uplift**を評価項目に加えるべきである。「研究対象そのものがBSL-1相当だから低リスク」という評価だけでは、方法論が将来どこへ転用可能かを捉えられない。

第五に、出版社と資金提供者は、論文受理時に**デュアルユースcommunication plan**を要求することが望ましい。公開する必要がある科学的結論と、悪用の限界費用を大幅に下げる実行情報を区別する。この考え方は、研究者・機関・資金提供者・出版社をすべてガバナンス主体とするWHO方針と整合する。 46

結論と実務的インパクト

科学的信頼性

本研究の総合的な信頼性は高いが、**一般化可能性はまだ限定的**と評価する。

信頼性を支える最大の理由は、AIモデルのスコアだけで論文を成立させず、物理的にDNAを合成し、宿主内で全ゲノムをrebootし、**宿主細胞内で感染サイクルを完遂して増殖可能なファージ**を回収したことである。その上で、配列確認、競争適応度、宿主域、耐性突破、cryo-EMという複数の異なる測定系を用いている。コード、モデル、配列、構造データが広く公開されていることも強い。 47

一方、「Evo 2が生命の設計原理を理解した」と結論づける証拠は不足している。5.61%の実験hit rateは事前に強く選別された候補での値であり、成功配列もΦX174/Microviridaeの進化空間を基盤にする。原論文内には単純な制約付き進化モデルとの直接比較が不足する。もっとも、Black et al.の独立再解析は、Evo 2尤度と生存性の関連、およびランダム変異・継代進化に対する設計効率を支持した。17, 52

ただし、その効率の多くは天然配列近傍の探索と選別フィルタに由来し、全く新しい配列空間を自由に開拓した証拠ではない。したがって、原論文のベースライン不足という批判は、後続再解析によって一部緩和されたものの、完全には解消していない。52

したがって、現段階で最も正確な表現は、

Evoは、既知の進化的足場とwet-lab selectionを利用しながら、従来人間が容易には列挙できなかった全ゲノム候補空間を探索し、その一部について実際に機能する組合せを発見した

である。

「完全de novo生命設計」より控えめだが、それでも十分に重要な成果である。

再現性

計算再解析可能性：高め。 モデルウェイト、コード、生成配列、viable/nonviable配列が公開されている。48

構造再解析可能性：高め。 EMDBにcryo-EM mapが登録され、第三者が構造品質を再解析できる。14

研究チーム内の実験検証：高い。 配列機能、競争適応度、宿主域、耐性突破、cryo-EMを用いて多面的に検証している。

独立wet-lab再現性：未確認。 研究チームは方法を公開しているが、2026年8月10日時点で別研究室によるend-to-end再現は確認できない。計算再解析可能性の高さと、独立実験再現性は区別して評価する必要がある。

モデルの設計上の有用性：中程度。 Black et al.の独立計算再解析は、モデル尤度と生存性の関連、およびランダム変異・継代進化に対する設計効率を支持する。ただし査読前で、効率には天然配列近傍の探索と後段フィルタが大きく寄与し、異なるウイルス群への一般化は未確認である。52

産業への影響

短期的に最も現実的なのは、**ファージ探索の候補供給エンジン**としての利用である。天然環境から適切なファージを採取し続けるだけでなく、既知ファージ群から大規模な候補プールを生成し、物理実験で絞り込むdesign-build-test loopを作れる可能性がある。耐性実験で、多様なAI由来ゲノムが組換えの素材となった結果は、この「多様性供給装置」としての用途を支持する。 25

ただし商用化のボトルネックはモデル推論ではなく、**合成コスト、機能hit rate、宿主域評価、製造品質、安定性、規制対応**になる可能性が高い。Stanford自身もDNA合成が高価であることを認めている。 32

農業では植物病原細菌向けファージ、工業では特定の細菌制御、研究用途では遺伝子ネットワーク・ウイルス進化の探索に展開できる可能性がある。Arcは今後の候補としてより大型のDNA phageや植物病原菌を挙げているが、これは将来計画であり、**実証済み成果ではない**。 25

医療への影響

本研究はファージ治療にとって興味深いproof of conceptだが、**治療法を実証した研究ではない**。実験対象は実験室E. coliであり、動物感染モデル、患者、薬物動態、免疫反応、毒性、製剤安定性、体内での細菌・ファージ共進化は評価されていない。研究が示したのは、「遺伝的に多様な人工ファージ集合が、単一天然ファージの耐性を超える進化的素材を提供できる」という段階である。 49

特に重要なのは、耐性突破が**最初のAI出力そのものではなく、AI設計群の組換えと追加進化の結果**だった点である。将来の医療応用では「患者ごとにAIが完成薬を出力する」というより、AIが候補多様性を提供し、高スループット評価・進化・製造・臨床判断が後段に続く形が現実的である。

公衆衛生への影響

正の側面では、抗菌薬耐性菌に対する新しい探索プラットフォームとなり得る。また、ゲノム生成モデルを使って病原体の可能な進化空間を先回りして研究することが、防御的なワクチン、診断、治療開発に寄与する可能性もある。Stanfordは、AIを自然発生・人工的生物脅威に対する防御側の技術として利用できると主張している。³²

負の側面では、今回の研究が「**whole-genome generative designは現実の機能性ウイルスに到達できる**」ことを実証した能力デモンストレーションである以上、病原体設計ガバナンスは「まだ架空の将来能力だから後で考えればよい」という段階を過ぎた。Inglesby/Hankeが同号でbiosafety/biosecurityを前面に出した理由はここにある。¹⁶

その意味で最も妥当な政策は、研究を一律に閉じることでも、無条件に全てを公開することでもない。**安全なファージ研究・失敗データ・評価手法のオープン性を保ちつつ、高リスク病原体への能力移転、危険なモデル特化、物理核酸合成、外部放出に近づくにつれて監督を強くする「段階的ガバナンス」**である。WHOのshared responsibility、米国の2026 high-risk life sciences policy、核酸合成screeningの三者は、この方向で相互補完的に利用できる。⁵⁰

最終評価として、本研究は「AIが生命を創造した」というセンセーショナルな話ではなく、より技術的かつ重要な転換を示している。すなわち、ゲノムを“読む”モデルが完全ゲノムの候補を“書き”、化学合成とwet-lab selectionを介して、**宿主細胞内で感染サイクルを完遂し、増殖する生物学的システム**まで到達した。現在は小型・非病原性のファージで成功率も低い。しかし、モデル性能向上、DNA合成費低下、高スループット自動実験が同時に進めば、現在wet-labが担う強い選択ボトルネックは縮小する可能性がある。Arc Institute自身も、合成コスト低下とモデル改善によってwhole-genome designの探索範囲が拡大すると予測している。したがって、**科学的には非常に有望、臨床的にはまだ初期、バイオセキュリティ政策上は今から制度設計すべき能力**、という三つを同時に認識するのが最も妥当である。⁴⁹

主要参考資料

1, 7, 13

King SH et al., “Generative design of bacteriophages with genome language models,” Science 393 (2026).
<https://doi.org/10.1126/science.aec2657>

2, 5, 10, 18, 19, 21, 22, 24, 25, 35, 47, 49

Arc Institute, “How We Built the First AI-Generated Genomes.”
<https://arcinstitute.org/news/hie-king-first-synthetic-phage>

3, 8, 29

King SH et al., “Generative design of novel bacteriophages with genome language models,” bioRxiv (2025).
<https://www.biorxiv.org/content/10.1101/2025.09.12.675911v1>

4, 17

Fraser J; Bondy-Denomy J., PREview of the King et al. preprint.
<https://zenodo.org/records/17178008>

52

Black JRM et al., “Quantifying evolutionary novelty and design efficiency in generative genome design,” bioRxiv (2026), DOI 10.64898/2026.06.12.731871.
<https://www.biorxiv.org/content/10.64898/2026.06.12.731871v1>

9, 20, 23, 26, 32

Stanford Report, “AI designs a novel E. coli killer” (2026-08-06).
<https://news.stanford.edu/stories/2026/08/evo-2-ai-tool-e-coli-killer-bacteriophages>

51

AAAS / EurekAlert!, “AI system designs functional bacteriophages from scratch” (2026-08-06).
<https://www.eurekalert.org/news-releases/1138470>

11, 41, 48

ArcInstitute/evo2 GitHub, phage_gen.
https://github.com/ArcInstitute/evo2/tree/main/phage_gen

12, 30

evo-design/evo-2-7b-8k-microviridae (model weights).
<https://huggingface.co/evo-design/evo-2-7b-8k-microviridae>

配列

NCBI GenBank, accessions PZ509151–PZ509168.
<https://www.ncbi.nlm.nih.gov/nucleotide/>

14

EMDB EMD-77390, Evo-Φ36 single-particle map / fitted model 36CQ.
<https://www.ebi.ac.uk/emdb/EMD-77390>

15

Evo 2 / Nature (2026), DOI 10.1038/s41586-026-10176-5.
<https://doi.org/10.1038/s41586-026-10176-5>

16

Inglesby TV; Hanke MS., “AI-designed viral genomes,” *Science* (2026).
<https://doi.org/10.1126/science.aej8512>

27, 42, 43

Science Media Centre, expert reaction to the Science paper (2026-08-06).
<https://www.sciencemediacentre.org/expert-reaction-to-generative-design-of-bacteriophages-with-genome-language-models/>

28

AlphaLab, 二次解説（9/7内訳に関する未確認情報の参照元）。
<https://www.alphalab.site/ai-designed-bacteriophage-genomes>

31, 33, 36

HHS, 2023 Screening Framework Guidance for Providers and Users of Synthetic Nucleic Acids.
<https://aspr.hhs.gov/S3/Pages/Synthetic-Nucleic-Acids.aspx>

53

HHS, Frequently Asked Questions — 2023 Screening Framework Guidance.
<https://aspr.hhs.gov/S3/Pages/Synthetic-Nucleic-Acid-Screening-FAQ.aspx>

34

OSTP / HHS, 2024 Framework for Nucleic Acid Synthesis Screening.
<https://aspr.hhs.gov/S3/Pages/OSTP-Framework-for-Nucleic-Acid-Synthesis-Screening.aspx>

38

HHS, History of Research Oversight Policies.
<https://aspr.hhs.gov/S3/Pages/History-of-Research-Oversight-Policies.aspx>

39, 40, 44

HHS / USG, High-Risk Life Sciences Research Oversight.
<https://aspr.hhs.gov/S3/Pages/High-Consequence-Research-Oversight.aspx>

6, 37, 45, 46

WHO, Global guidance framework for the responsible use of the life sciences (2022).
<https://www.who.int/publications/i/item/9789240056107>

50

WHO, Technical Advisory Group on the Responsible Use of the Life Sciences and Dual-Use Research.
[https://www.who.int/groups/technical-advisory-group-on-the-responsible-use-of-the-life-sciences-and-dual-use-research-\(tag-ruls-dur\)](https://www.who.int/groups/technical-advisory-group-on-the-responsible-use-of-the-life-sciences-and-dual-use-research-(tag-ruls-dur))