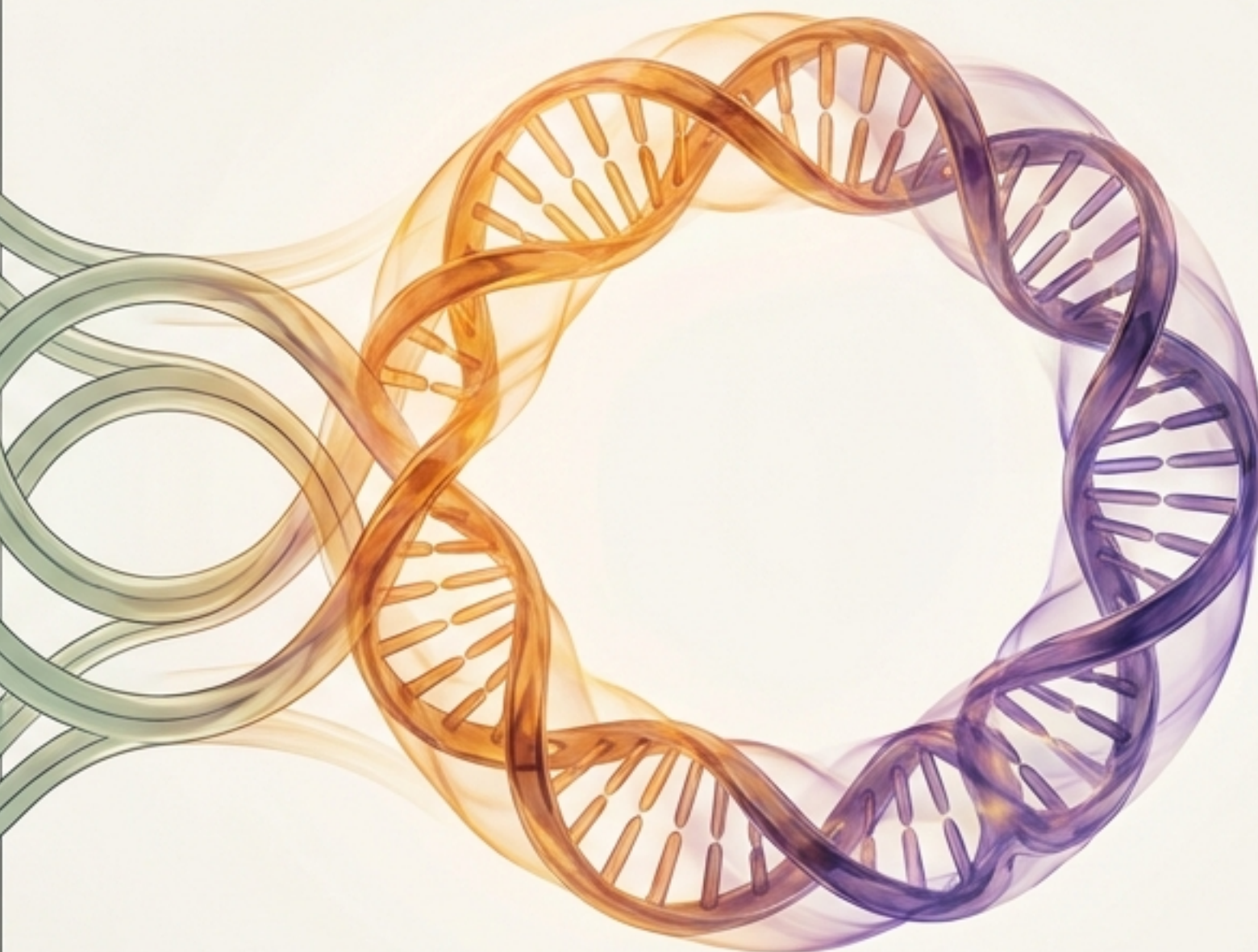
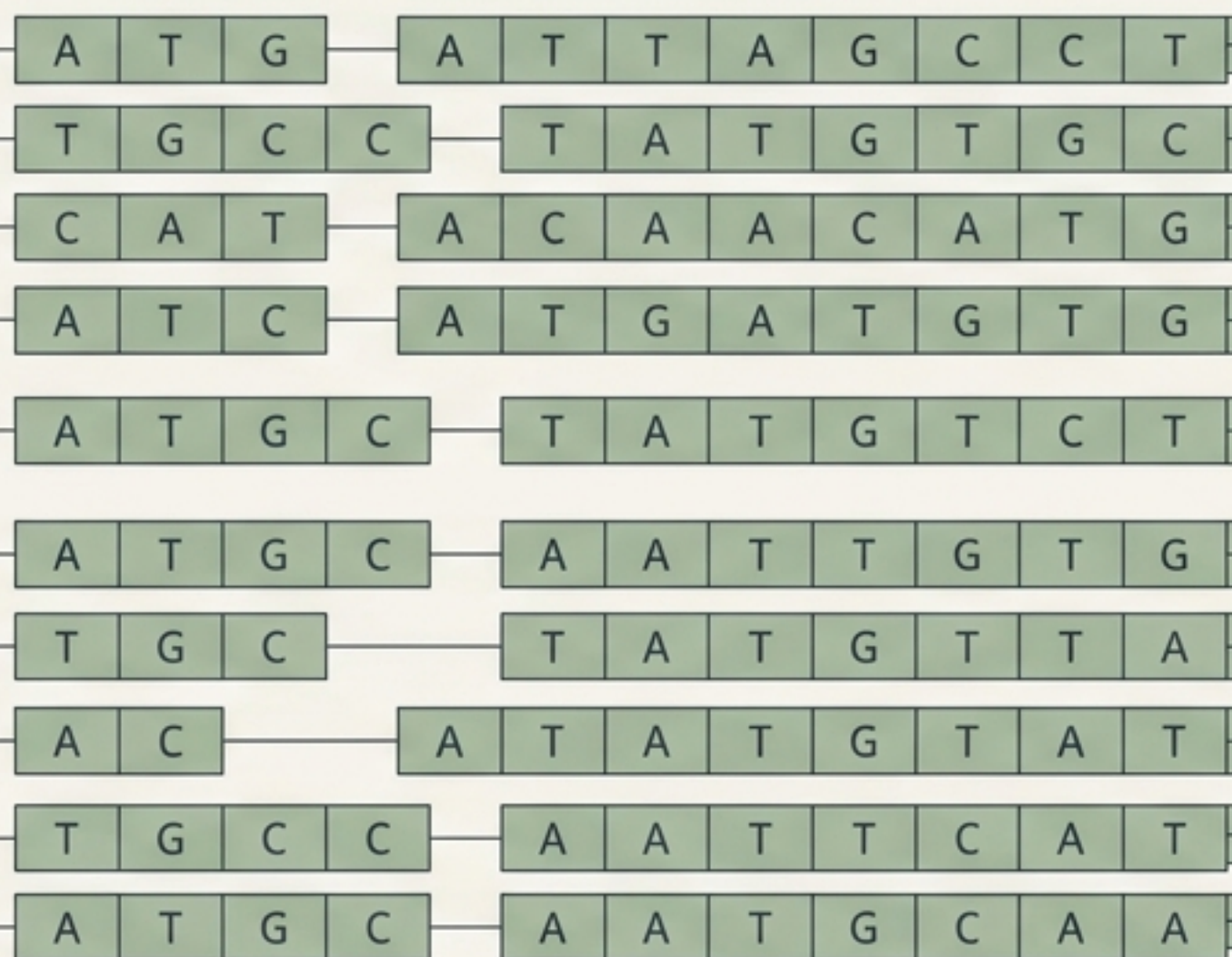
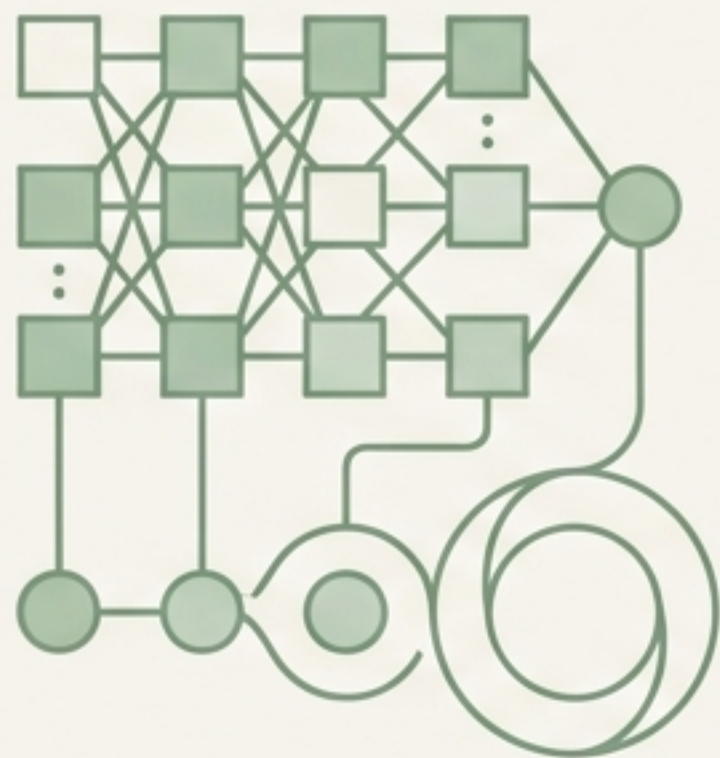


ゲノム言語モデルの衝撃：Evoによる機能性生命のゼロからの生成

AIはいかにして「タンパク質の予測」から「全ゲノムの自律設計」へと移行したか。
その技術的機序とバイオセキュリティにおける新たなパラダイム。

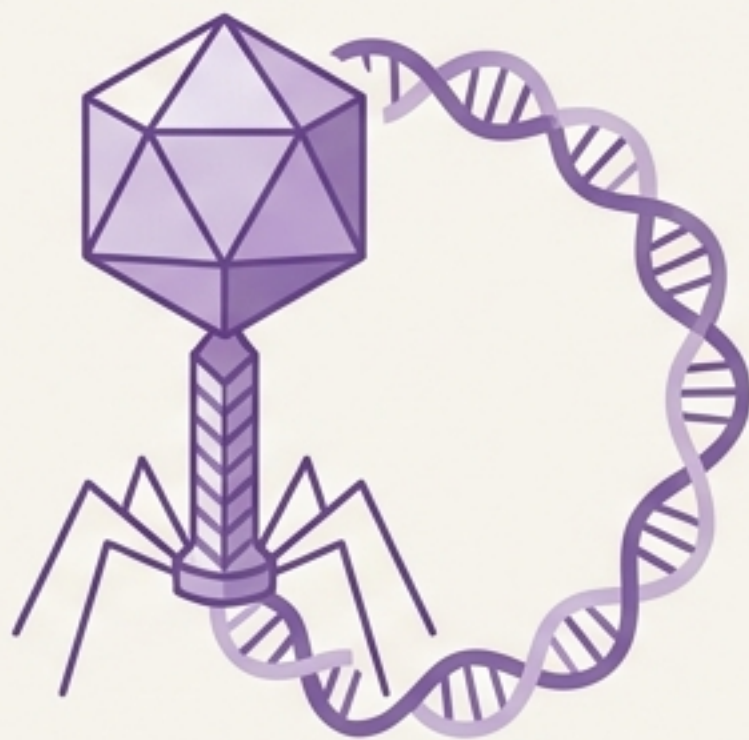


エグゼクティブ・サマリー



ゲノム生成の実現

スタンフォード大学とArc Instituteによる革新。単一タンパク質の構造予測（Alpha Fold等）から、生命システム全体の基盤となる「全ゲノム」のDe Novo生成へ移行。



実験的な完全証明

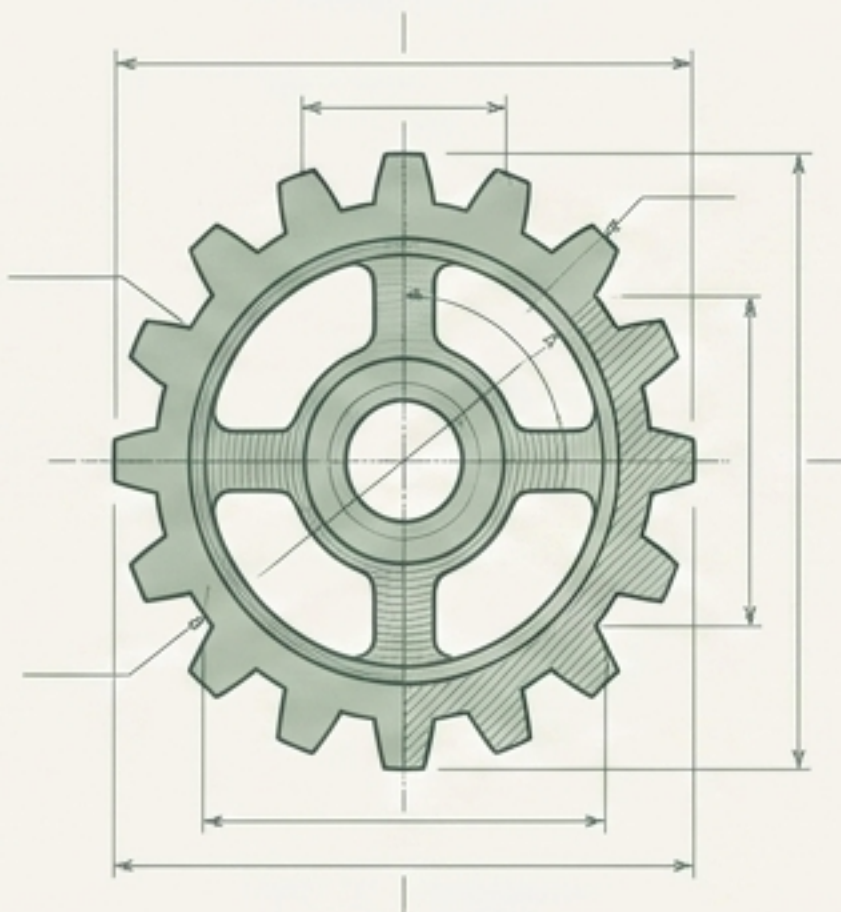
100万塩基のコンテキストウィンドウを持つEvoモデルを用い、試験管内で機能し自己増殖する「16種類の全く新しいバクテリオファージ」の生成に成功。



次世代のガバナンス

自然界の相同性に依存しない未知の配列を生成できるため、既存のDNA合成スクリーニングを無効化するパラドックス。プログラマブルな防衛体制の構築が急務。

従来のアプローチ



AlphaFold / ESM

- 対象: 単一の遺伝子・タンパク質の3次元構造予測
- 限界: 生命機能は単一の部品では決定されない。

新たなパラダイム

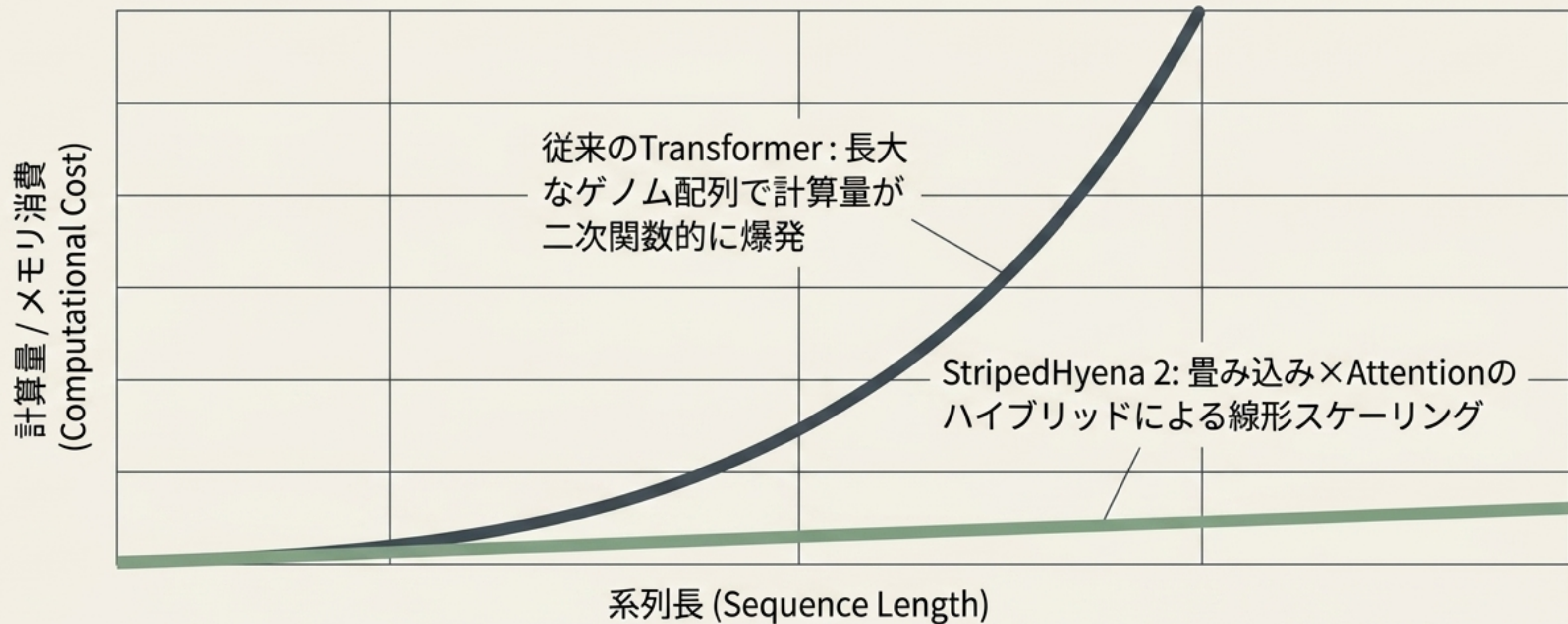


ゲノム言語モデル (Evo)

- 対象: ゲノム全体（システム）の協調設計
- 特徴: 遺伝子、プロモーター、非翻訳領域、エンハンサーなど、生命活動に必要な全ての相互作用を学習・生成する。

The Evo Foundation: ゲノム言語モデルの進化

モデル	パラメータ数	学習データセット	コンテキストウィンドウ	アーキテクチャ特性
Evo 1	7B	OpenGenome (原核生物・ファージ 270万種 / 約3000億 トークン)	131,072 塩基対	ゲノム言語モデルの初 期実装、原核生物・ ファージ特化。
Evo 2 (7B / 40B)	7B & 40B	OpenGenome 2 (全生命ドメイン / 2.4兆~9.3兆トークン)	最大 1,000,000 塩基対	StripedHyena 2 アーキテクチャ採用。 全ドメイン対応の高精 度なDe Novo生成。



文脈の壁の突破

段階的な長文脈ミッドトレーニングにより、単一塩基解像度を保持したまま「100万塩基対」のコンテキストウィンドウを実現。離れた調節領域と遺伝子コード領域の長距離相互作用の学習が可能に。

生成ターゲット：バクテリオファージ ΦX174

- **Microviridae科の溶菌性ファージ**

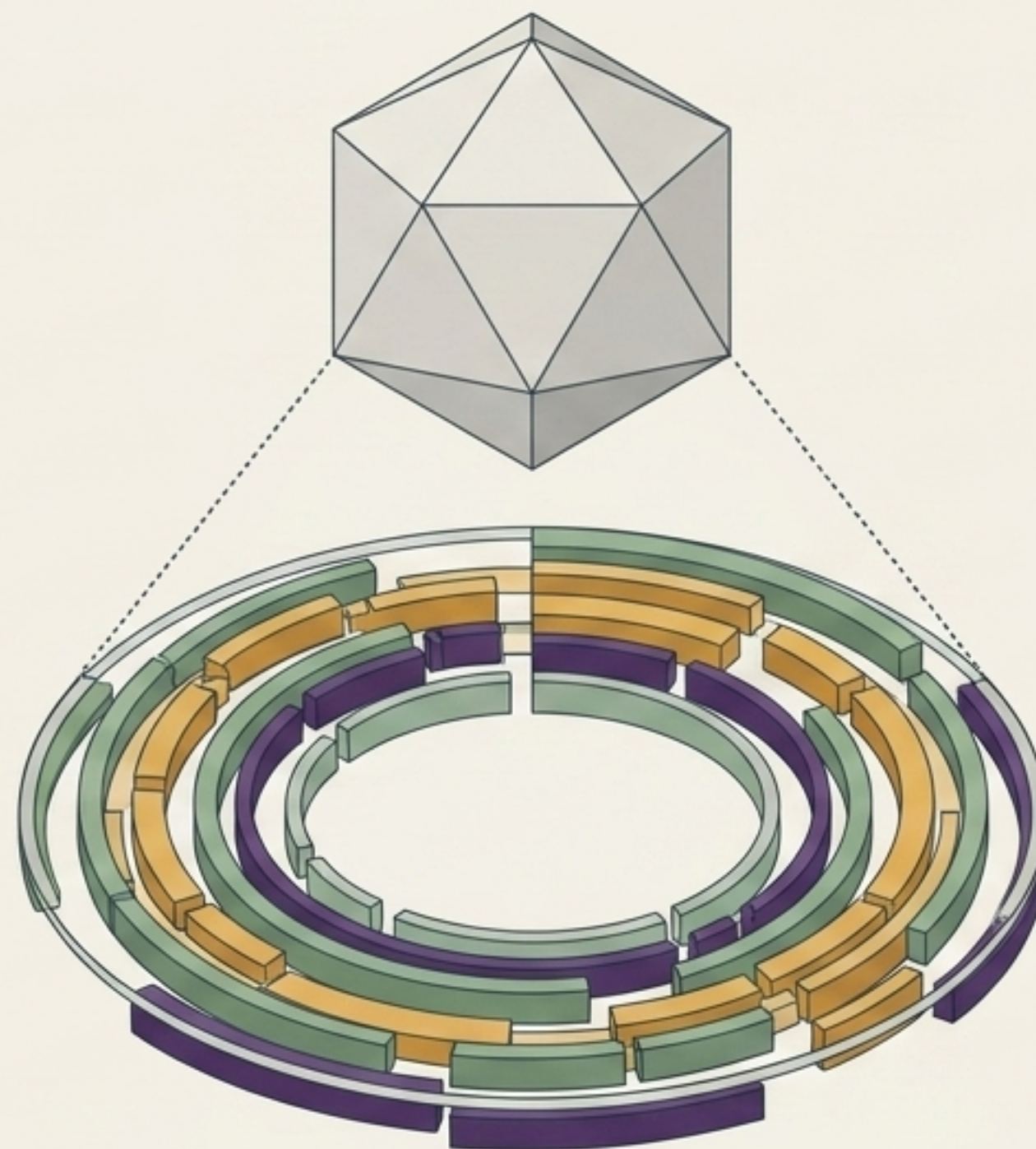
一本鎖DNAゲノムを持つ複雑なシステム。

- **高密度な遺伝的アーキテクチャ**

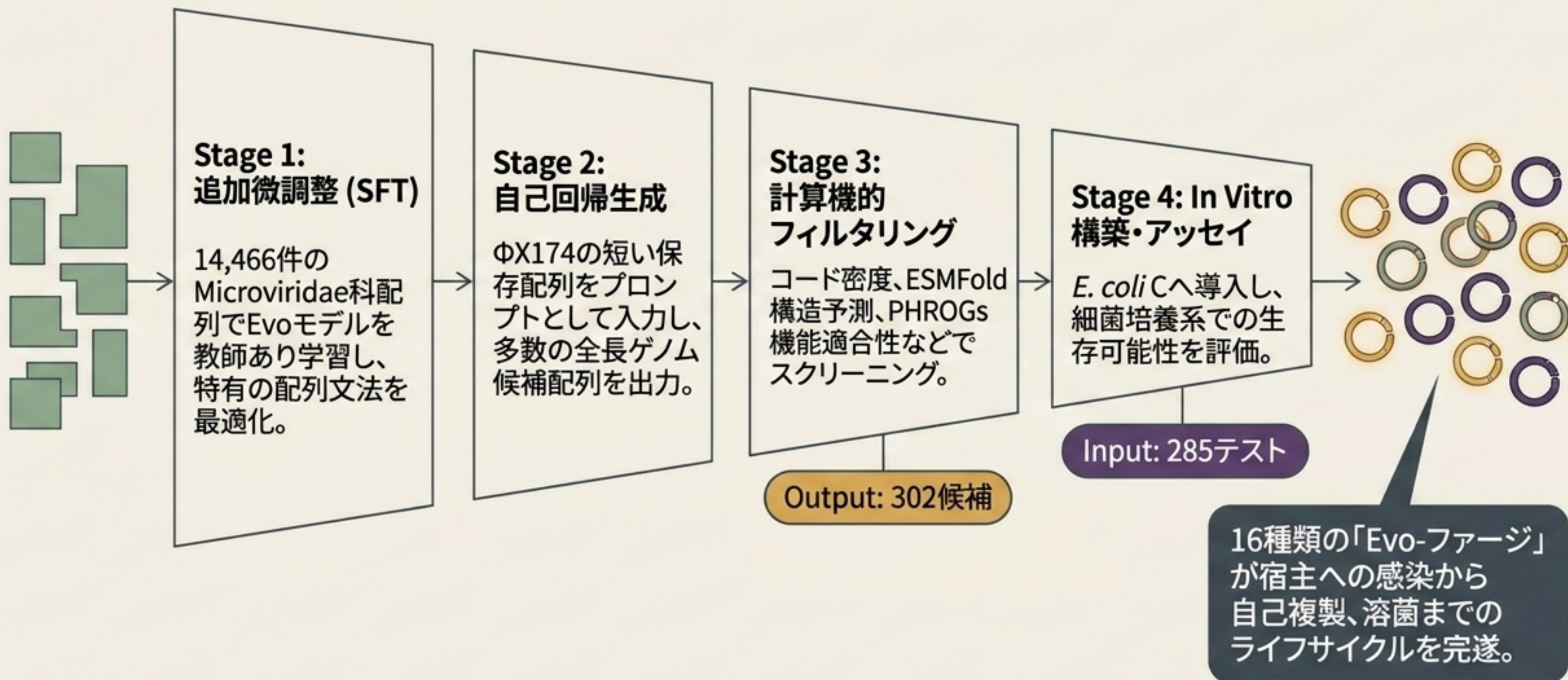
約5,386塩基対という極めてコンパクトなサイズに、11個の遺伝子と調節エレメントが密集。

- **重複遺伝子領域 (Overlapping Genes)**

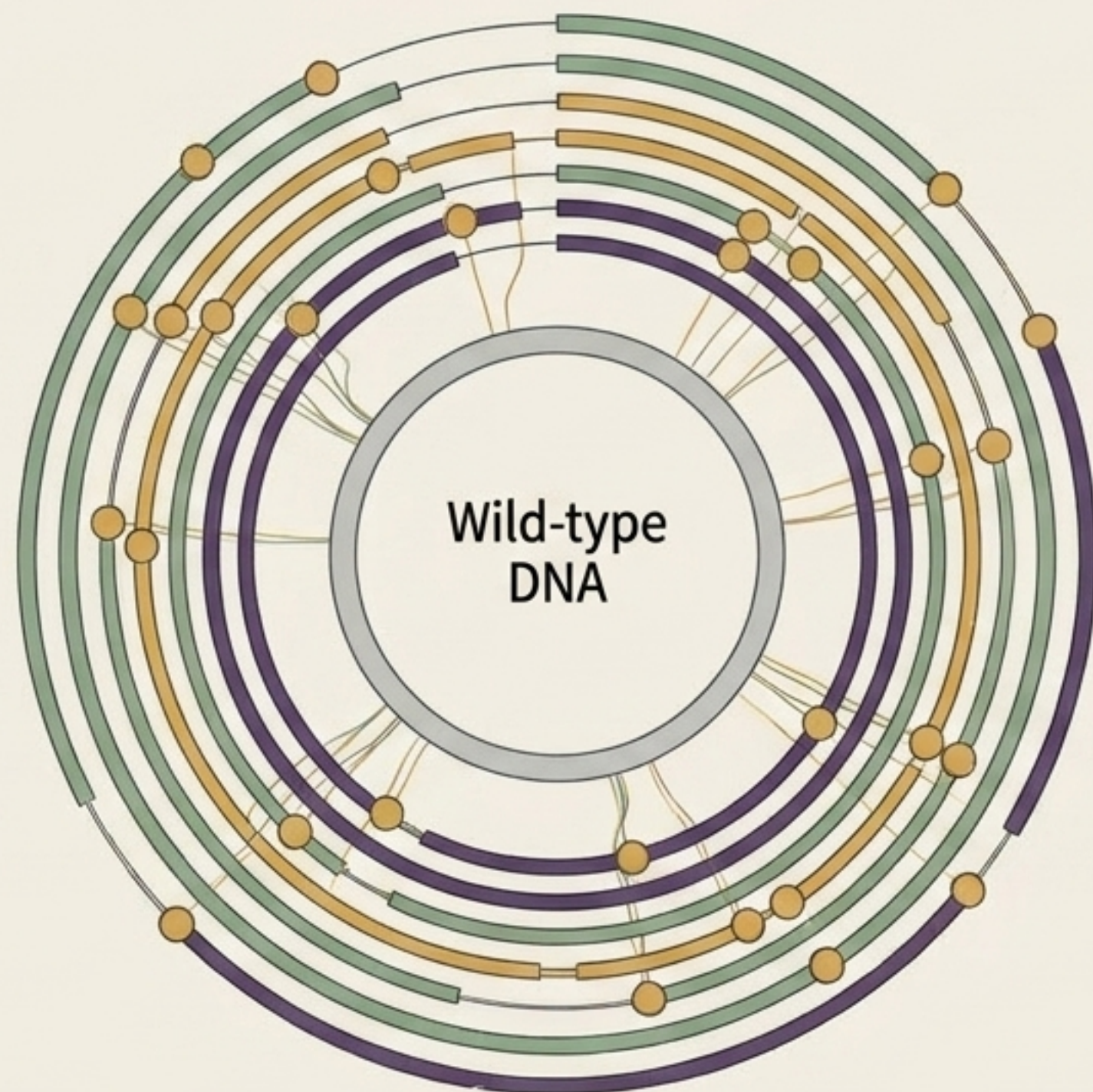
遺伝子配列が重なり合う複雑な構造を持つため、単一塩基の変更がシステム全体に致命的な影響を与える。AIの「システムレベルの設計能力」を証明する究極のテストベッド。



The Digital-to-Biological Loom



単なる模倣ではない「プログラマブルな多様性」



大規模な遺伝子変異

生成された16種類は、自然界の最近縁ゲノムと比較して「67~392か所」の塩基変異を有する。自然界の単一変異体を凌駕する多様性。

動的なゲノム再編成

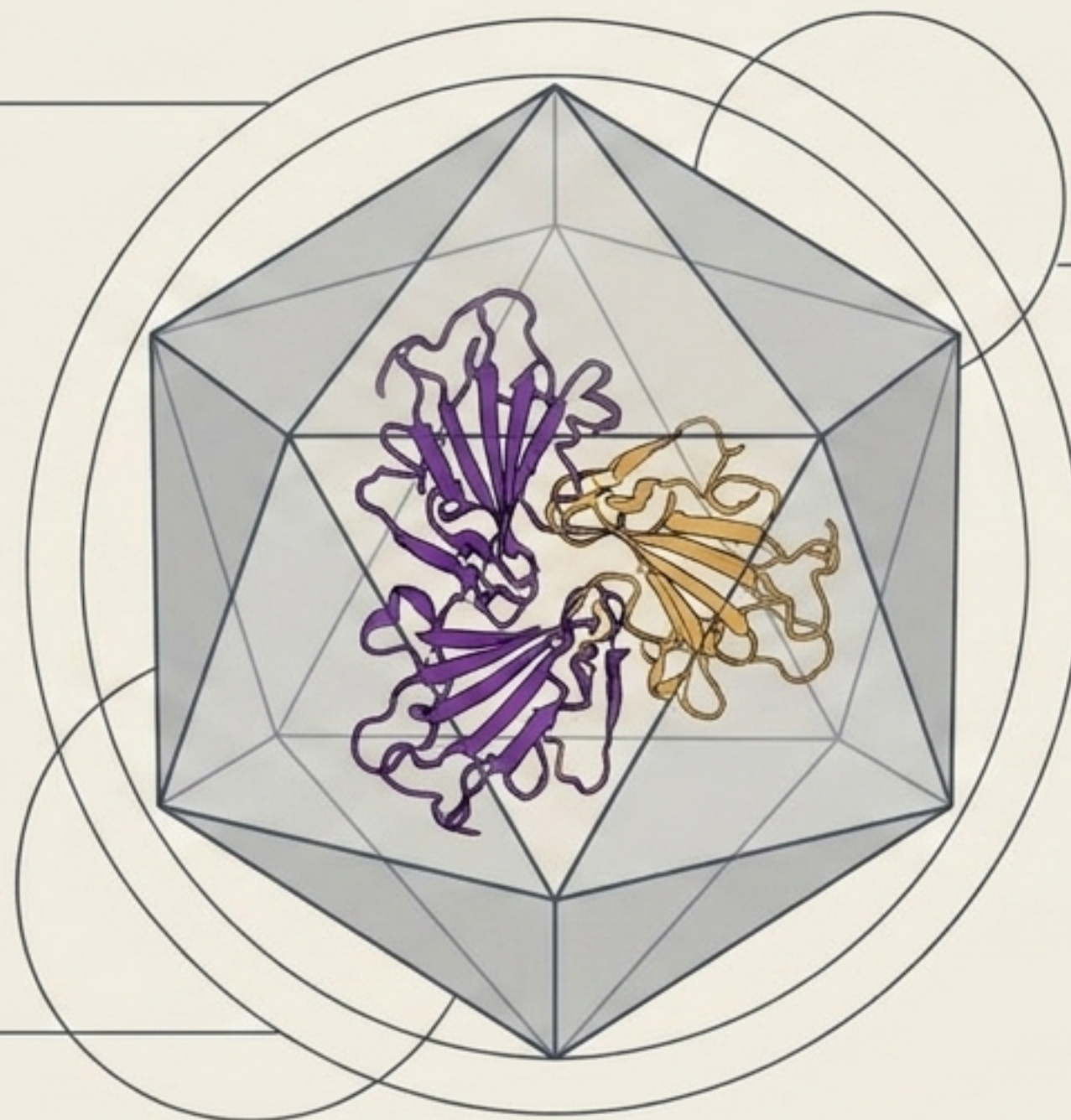
Evo-ファージにおける遺伝子Kの欠損、遺伝子C・Bの短縮化、遺伝子J領域への新規挿入など、大胆な構造改変を自律的に実行。

新種の可能性

最も優れた Evo-ΦX174-1 は、最近縁天然ファージ (NC51) との配列同一性が93.0%にとどまり、新種に相当する可能性を示唆。

クライオ電子顕微鏡による構造的裏付け

インサイト: Evoモデルは、単に塩基を並べるだけでなく「**物理的な構造の整合性**」も理解している。

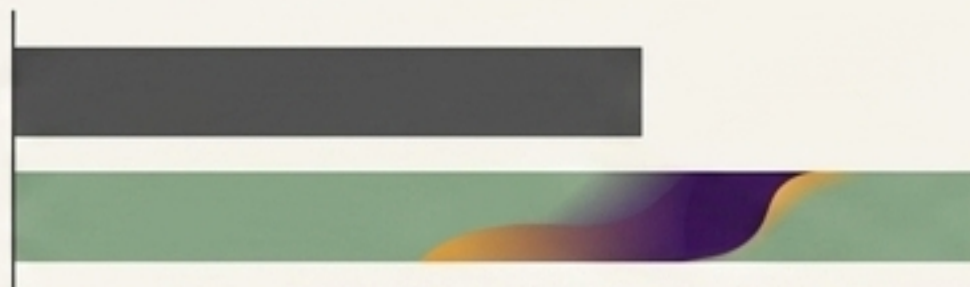


AIが設計したEvo-ファージのカプシド内部には、進化的に遠く離れた系統 (Escherichia phage G4など) に由来するDNAパッケージングタンパク質が組み込まれており、それが物理的・機械的に正常に機能していることがCryo-EMによって確認された。

Natural vs. Artificial Fitness

溶菌速度
(Lysis Kinetics)

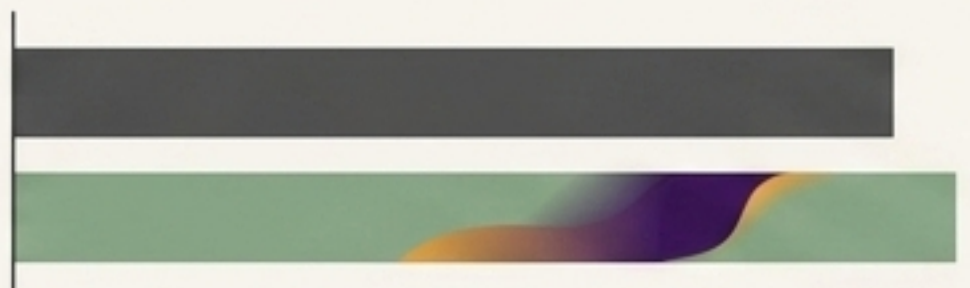
Wild-Type
Evo-Phages



一部のEvo-ファージは、天然の野生型と同等、あるいは「それを上回る」溶菌速度を記録。

増殖適応度
(Competitive Fitness)

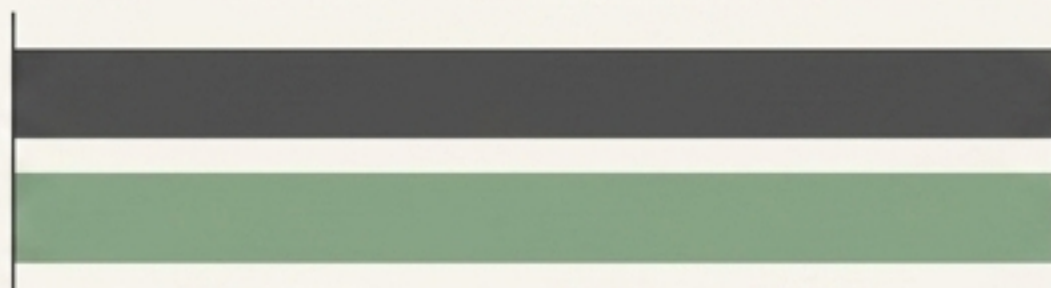
Wild-Type
Evo-Phages



競合増殖実験において、AI生成ゲノムが高適応度バリエーションとして機能。

宿主域
(Host Specificity)

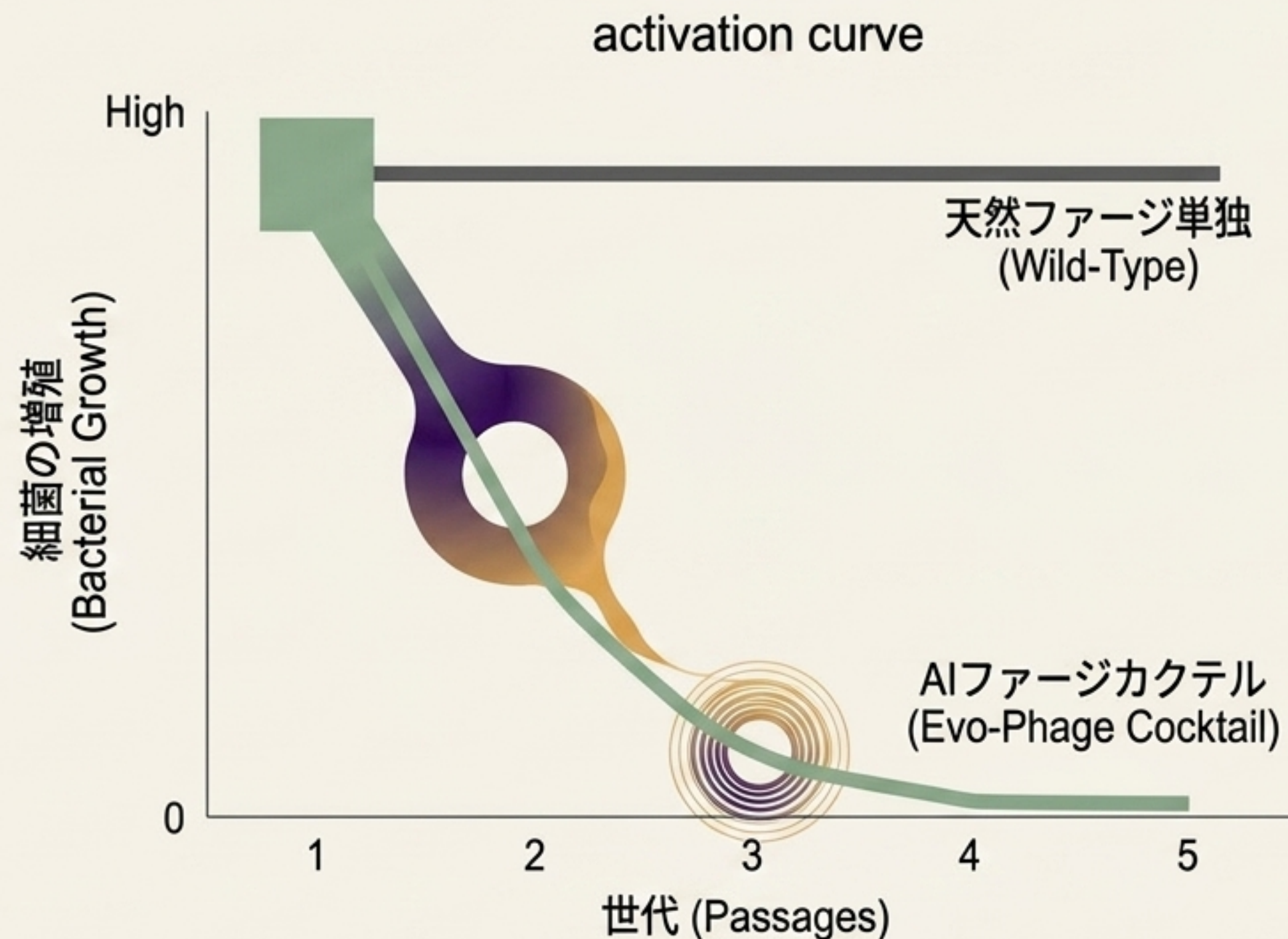
Wild-Type
Evo-Phages



標的株 (E. coli C / W) に特異的に感染し、その他の6株 (E. coli K-12等) では増殖せず。in vitroでの特異性を維持。

結論: AIは生存可能な配列を探すだけでなく、
「自然界よりも高効率なゲノム」を探索・生成できる。

抗菌薬耐性 (AMR) の克服への道筋



天然ファージの限界

実験室で作製されたΦX174耐性大腸菌に対し、天然型の単独投与では増殖を抑制できず。

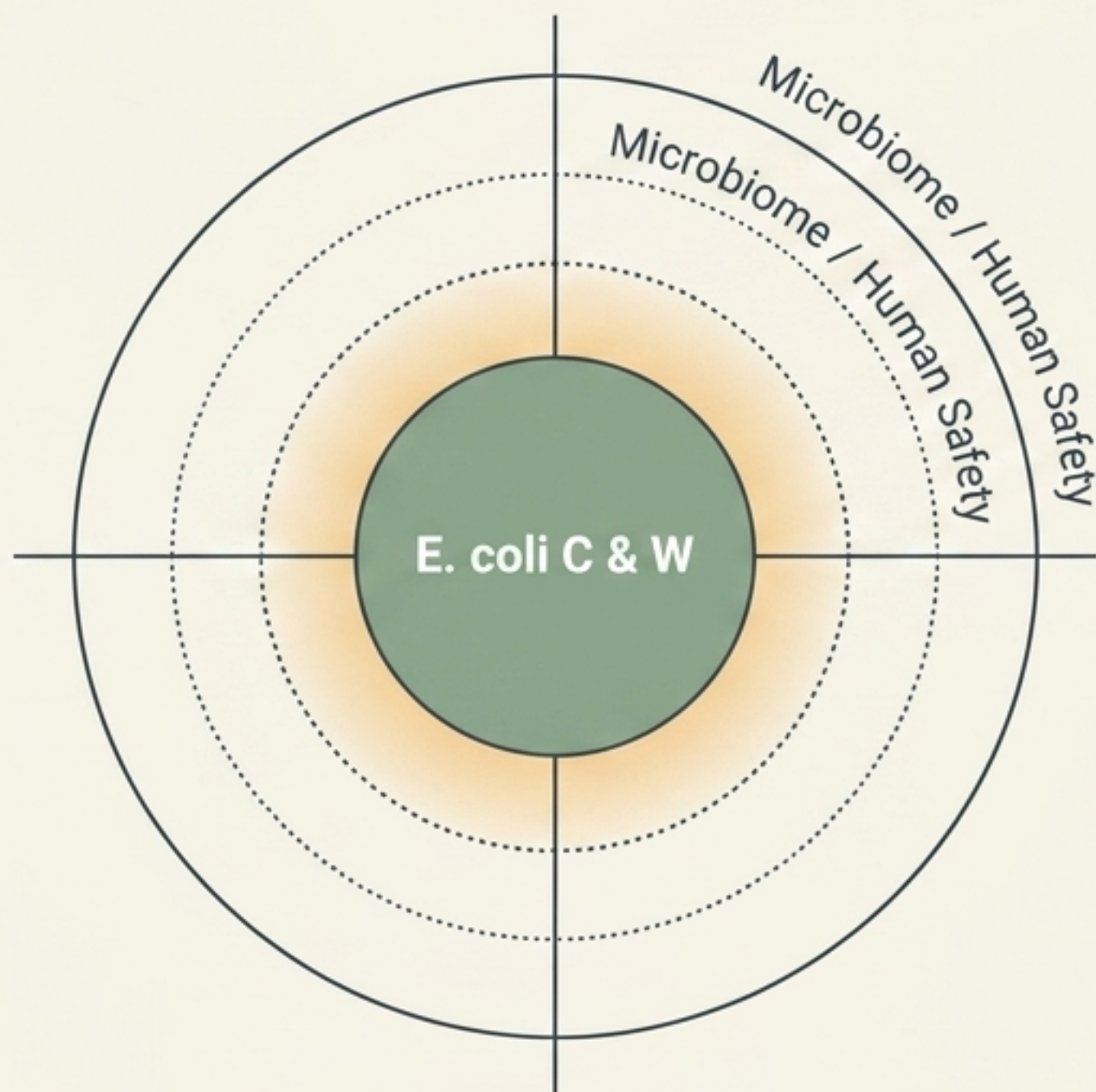
AIファージカクテルの威力

多様な塩基配列構造を持つEvo-ファージ群を組み合わせたカクテルを投与した結果、わずか1～5回の継代培養で大腸菌の耐性を突破し、完全な溶菌作用を確認。

未来の展望

感染症に対する「オーダーメイド型プログラマブル・ファージ療法」の基盤を実証。

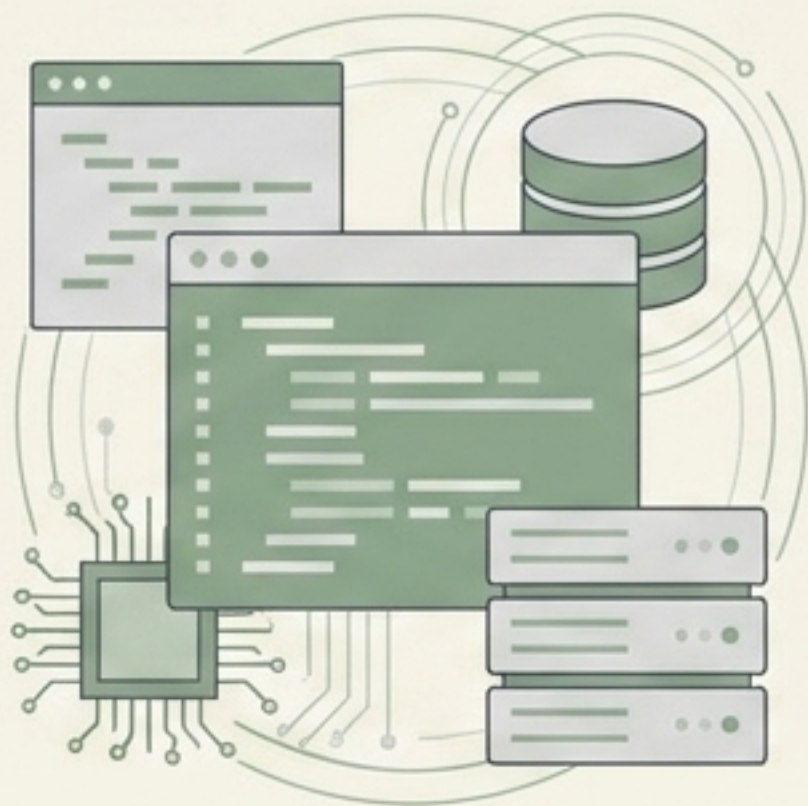
In Vivo 実装へ向けた未解決の境界線



● 実験室の限定的な菌株パネル (in vitro) における特異性と安全性は実証された。

● しかし、これは「複雑な人体内の腸内細菌叢 (マイクロバイオーーム) 全体に影響を与えない」ことを証明するものではない。

● 臨床的有効性、動物個体やヒトでの安全性評価は、今後の重要な検証課題として残されている。



オープンソースによるイノベーション推進
(Apache 2.0)



悪意ある主体による病原体改変
(デュアルユース)

事前学習フィルタリングの限界

モデル構築時にヒト病原性ウイルスのデータを意図的に除外しているが、これは数学的な生成不能性を保証するものではない。悪意ある第三者によるタスク特異的な追加学習 (Post-training / Fine-tuning) が行われれば、安全対策は容易に突破される。

バイオセキュリティのパラドックス



パラドックスの核心

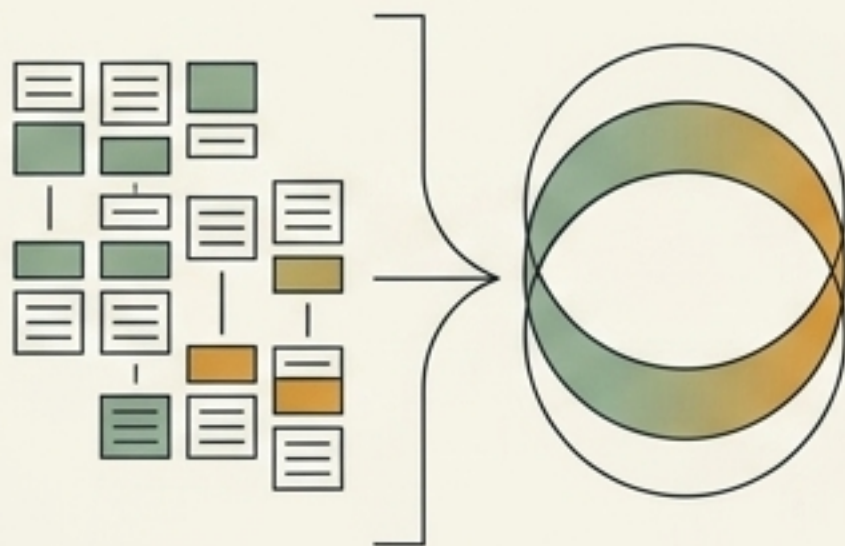
Evoモデル最大の革命である「自然界に存在しない新規機能性DNAの生成能力」。これは同時に、現在の世界のバイオ防衛網を無効化する。

検知ルール限界

既存の商用DNA合成事業者のスクリーニングは、既知データベースとの「配列相同性（BLAST等）」に依存している。AIがゼロから配列を設計し、相同性を著しく低下させた場合、致死的な機能を持つDNAであっても、現在の検知ルールをすり抜ける理論的危険性が生じている。

機能予測型スクリーニング (AI vs AI)

単純な配列の一致率ではなく、AIモデルの尤度スコアやタンパク質構造予測（ESMfold等）を用いて、DNAの「潜在的な毒性・病原性機能」を自動検知するシステムへの移行。



ハードウェア認証の 義務化

卓上型DNA合成機（Benchtop Synthesizers）におけるデジタル証明書認証と、リアルタイムスクリーニングのハードウェアレベルでの組み込み。



緊急防衛アプローチ (CRISPR)

未知のAI生成病原体が流出した際に、即座にガイドRNAを計算・設計し、標的ゲノムを切断・無効化する緊急配備型防衛パイプラインの確立。



次世代ガバナンス構想

プログラマブルな生命の時代には、プログラマブルな防衛が必要である。