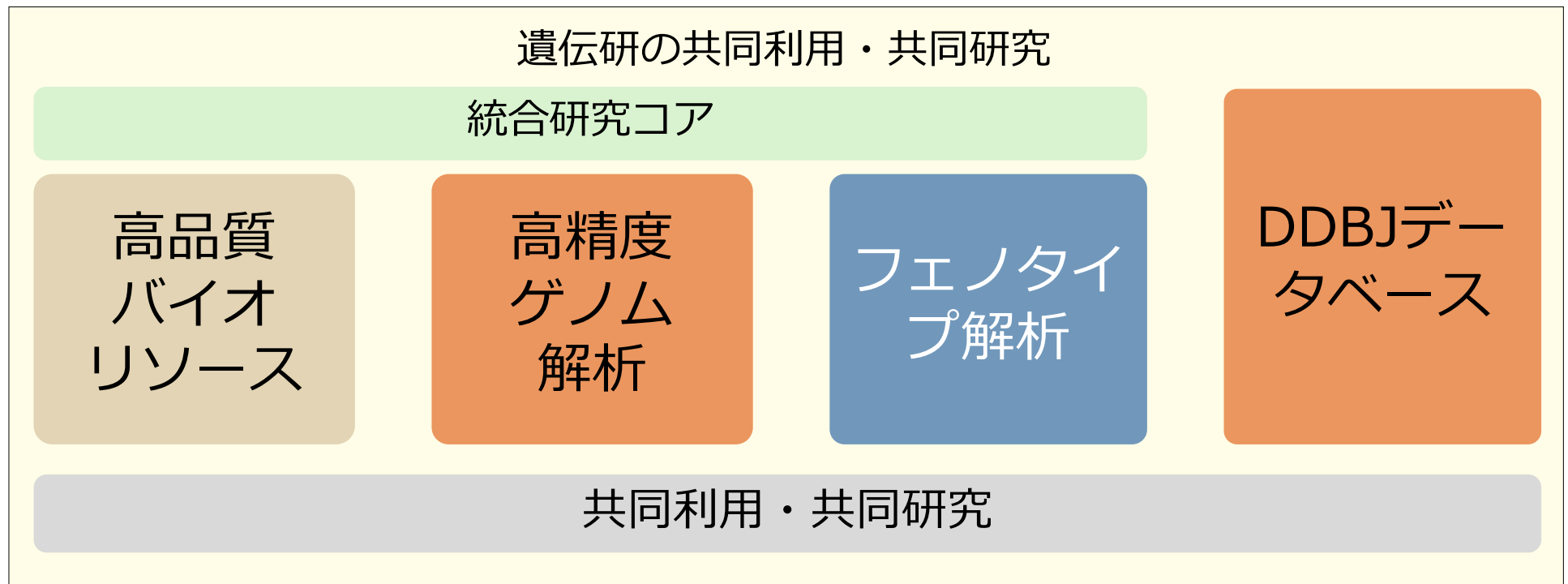


“観測から制御へ” 新時代の生命研究をリードする
AI駆動型生命科学研究拠点の形成
(AI-Empowered Life Science Initiative: ALIS)

国立遺伝学研究所

- ◆ 国立遺伝学研究所は、遺伝学的アプローチを基盤に生命現象の本質解明に挑む研究機関
- ◆ DNAの構造・機能解析、ゲノム情報の大規模解析、発生や進化のメカニズムの解明など、生命科学における基礎研究を推進
- ◆ 大学共同利用機関として



上記機能を通じ、国内外の研究者による先端的かつ挑戦的な研究を支えるハブ拠点としての役割を果たしている。



生成AIを中核とした新たな展開

生成AI

- ◆ 深層学習（Deep Learning）技術により、大量データに潜む複雑なパターンや関係性をモデル化し、そこから新たなデータを生成することが可能となった
- ◆ 膨大なデータから内在する構造や特徴を確率的に抽出し、それに基づいて類似した性質を持つ新規データを創出できる技術

自然言語に適用

大規模言語モデル
LLM

文章や会話を生成
構造化されていない情報から
知識を抽出することが可能

ゲノム配列に適用

ゲノム言語モデル
GLM

ゲノム配列を生成
DNA配列の機能や特徴を推
定・設計することが可能

GLMがライフサイエンスに与える影響

「研究プロセスの革命」 「圧倒的な効率化」



「研究対象の質的な拡大」 「生命情報の根本的理解」
といった次元の違う変革をもたらし得る

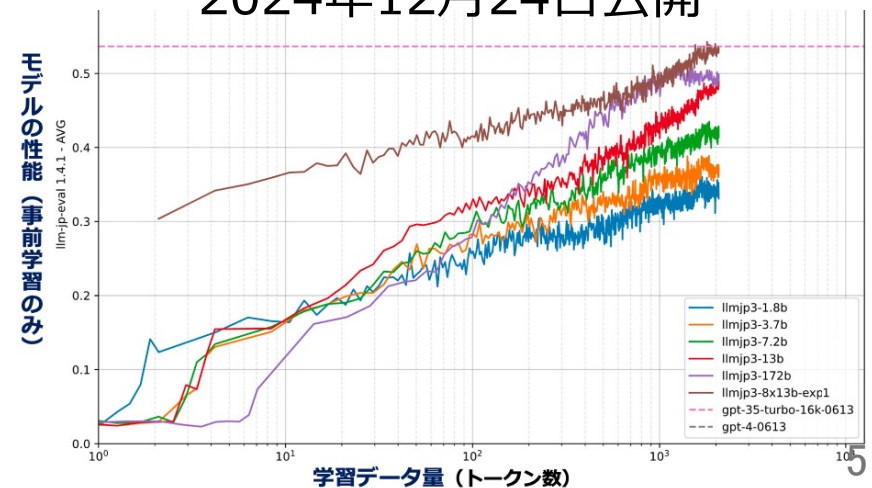
- ゲノム配列情報を活用することで、遺伝子機能予測や特異的変異の抽出が瞬時に可能となり、従来数ヶ月を要していた解析作業を数分～数時間に短縮することが期待される
- また、配列の最適化により、これまで試行錯誤を重ね、膨大な労力とコストを要していた遺伝子設計や変異体スクリーニングを大幅に省略できる
- さらに、配列に内在する未知の法則性や関連性を抽出することで、従来の仮説駆動型研究では考慮されなかった新たな知識や概念を自動的に発見でき、研究者が新規発見に至るプロセス自体が劇的に効率化される
- すなわち、AIが新たな科学的仮説を生成するプロセス自体が革新されることにより、研究分野全体の発展と国際競争力の強化が期待される

ROIS with NIG+NIIによる ゲノム生成AI

- ROIS は配下にNIIを持ち、遺伝研と情報研の連携が肝
- NIIはスクラッチ1700億モデルを開発済み（最近は 30B x 8 MoEを開発中）
- 言語処理を中心とするアカデミアが結集して「共同研究」として、巨大LLM開発を進めるという**新しい大規模共同研究のアプローチ**
- 巨大デジタルシステムは一研究者、一大学では不可能との判断で多数の研究者が集う

1700億パラメタLLM
2024年12月24日公開

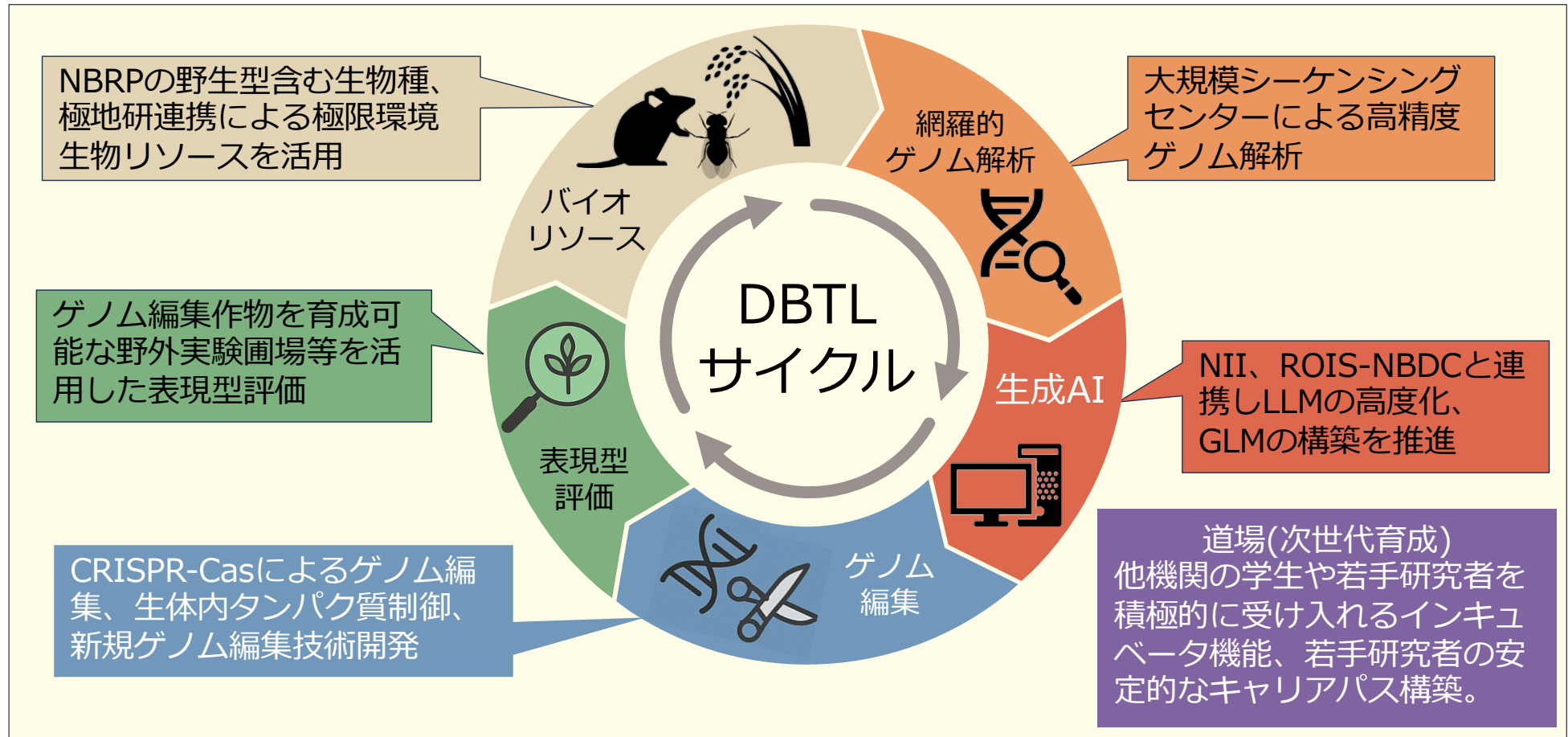
- 2024.4 LLMセンタ始動
- NII-NIGは、研究所レベルの**最初の共同研究**



「観測」から「制御」へ、生命科学の未来を切り拓く研究拠点

AI駆動型生命科学研究拠点

(AI-Empowered Life Science Initiative: ALIS)

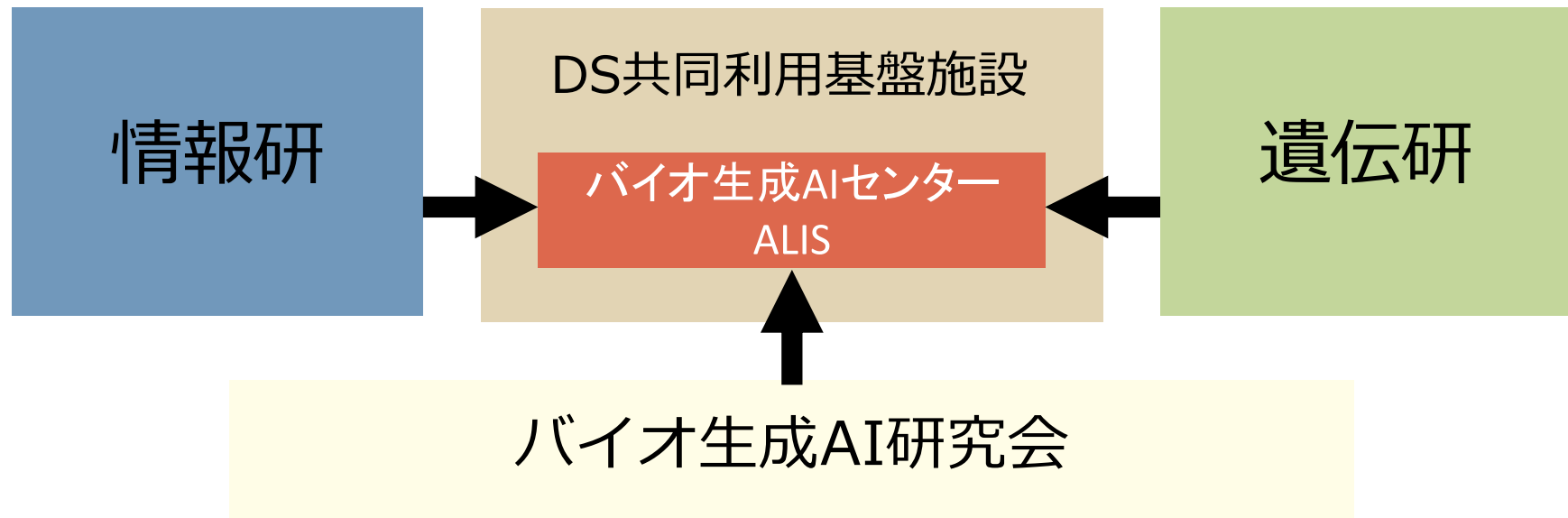


- 「AI × バイオリソース × 表現型解析」による実験プラットフォームを確立し、生命現象や環境応答の制御を可能
- 次世代型の共同研究を全国の研究者とともに推進



バイオ生成AIセンターの始動

ROIS・DS施設内にインキュベーションセンターとして「バイオ生成AIセンターALIS」を新設し、ROIS内外から研究者に集まって頂き、アルゴリズム開発および小規模モデルの構築を開始する予定。



バイオ生成AI研究会の発足

本研究会では、国内関連研究者が集結し、バイオ生成AIの可能性やゲノム言語モデル構築の課題、さらには実験検証との統合など、新しい研究領域を大規模共同研究として促進する場。

第1回研究会 6/2(月)14:00～17:00 @ROIS会議室 ハイブリッド開催

■プログラム

14:00～14:05	バイオ生成AI研究会発足の経緯（近藤@遺伝研）
14:05～14:15	ROISにおける生成AIの取り組みLLM.jp（黒橋@情報研）
14:15～14:25	バイオ生成AI(GLM)の研究動向（東@遺伝研）
14:25～15:25	バイオ生成AIに関する意見交換（自己紹介とともに）

【研究発表】

15:25～15:40	浜田道昭（早稲田大学）
15:40～15:55	白石友一（国立がん研究センター）
15:55～16:05	文部科学省の取組み（文部科学省）
16:05～16:10	閉会の辞（喜連川ROIS機構長）
16:10～16:15	事務連絡

◆ 開催場所：ROIS共用会議室 + NINS会議室

参加者	現地	オンライン	不参加	合計
メンバー	25	12	5	42
研究者（研究室メンバー）	0	6		
理研事務（TRIP関係・経営企画）	0	5		
遺伝研	0	7		
文科省（メンバー以外）	1	3		
ROIS	3	5		
合計	29	38		

<https://genaibio.netlify.app/>

合計67名の参加者

議論して頂きたいこと

- ・ どのようなモデルをつくれば良いか？
- ・ 計算機資源
- ・ 実験研究との連携
- ・ 人材育成

設置する研究部会（案）

■ データ基盤グループ

データ収集・統合・前処理

■ モデル開発グループ

アーキテクチャ設計と実装

■ 計算基盤・インフラグループ

計算機環境の設計、構築

■ ベンチマーク・評価グループ

評価指標の開発と性能測定

■ アプリケーション開発グループ

実用化に向けた応用開発

■ 生物学的検証グループ

モデル出力の生物学的妥当性評価

■ 技術統合委員会？

各グループ間の技術的調整