

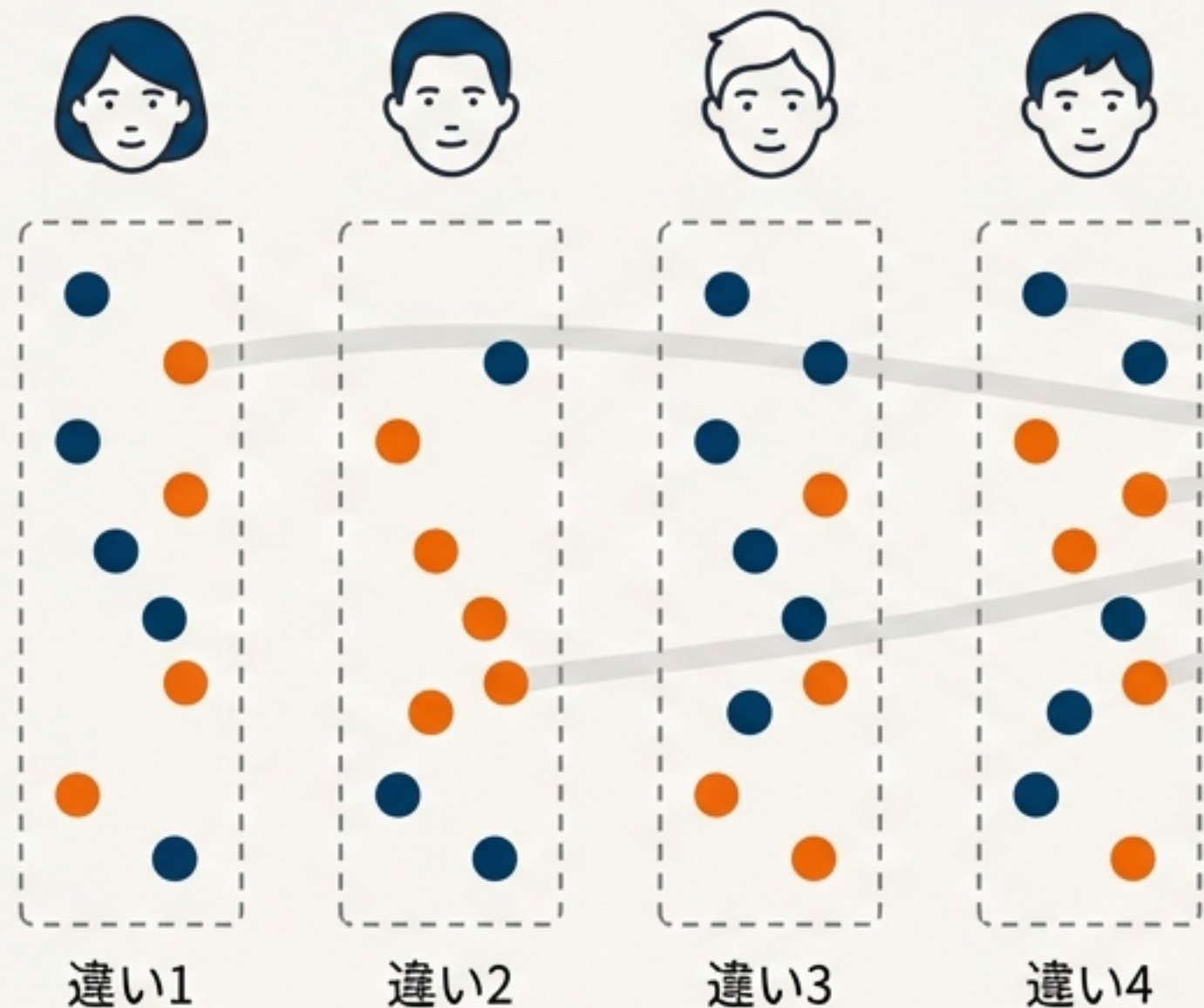


点から線へ：ハプロタイプ解析が 拓くゲノム科学の新時代

ロングリードシーケンスに基づくハプロタイプデータベース「JoGo」の紹介

ゲノム解析の現状：個々の「点」としての遺伝的変異

従来

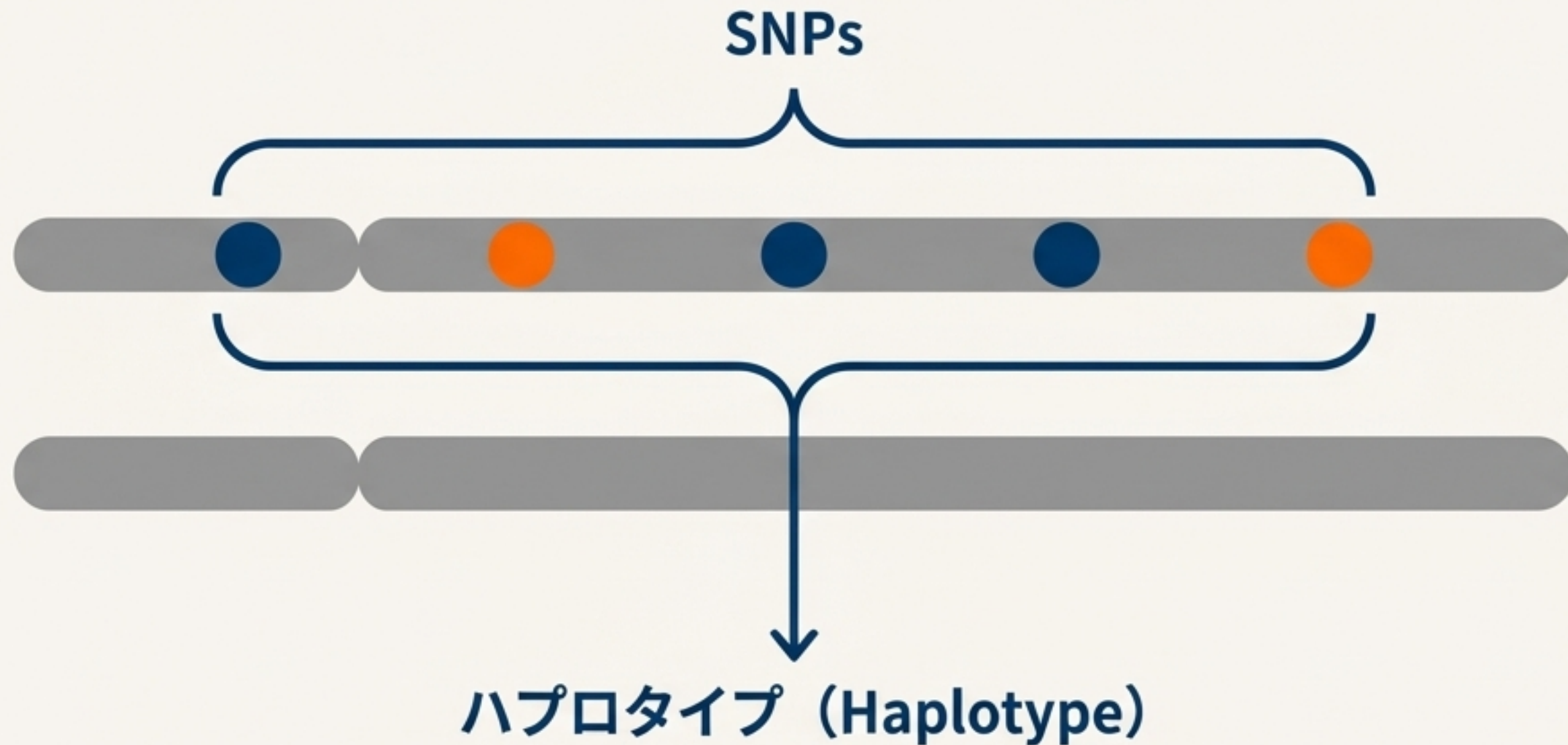


ゲノムワイド関連解析（GWAS）などは、個々の塩基の違い（SNP）という「点」の情報を中心に解析してきた。

しかし、これらの「点」の組み合わせが持つ生物学的な意味は？

複数の変異が協調して遺伝子機能に影響を与えるが、「点」だけではその全体像を捉えきれない。

「線」で捉える遺伝情報：ハプロタイプとは



ハプロタイプとは、同じ染色体上で一緒に受け継がれる複数の変異の組み合わせ。

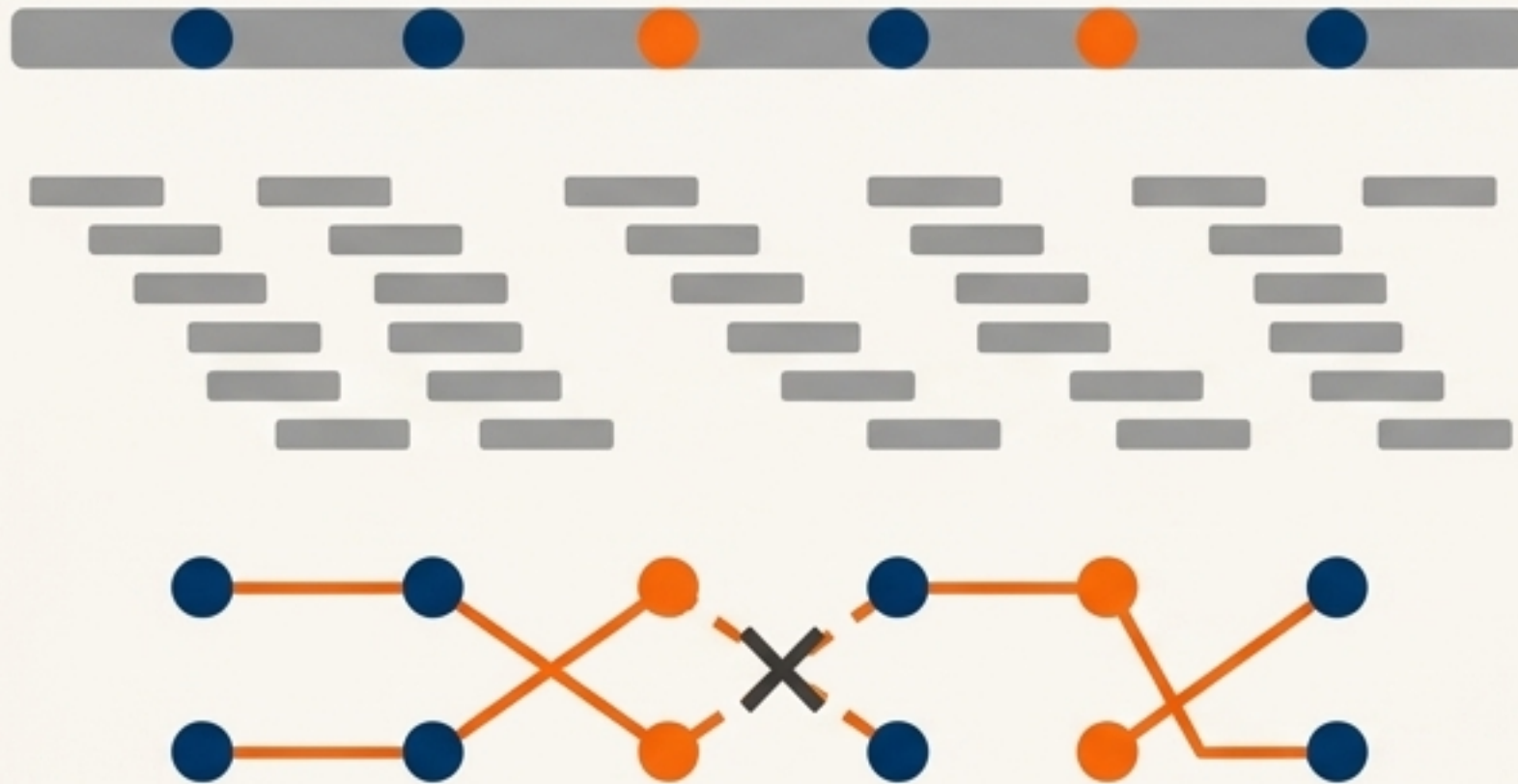
個々の変異（点）ではなく、遺伝子機能を規定する一つの単位の（線）として理解することが重要。

- 免疫応答（HLA）
- 薬剤代謝（CYP）
- 疾患感受性

なぜ「線」の解読は難しかったのか？

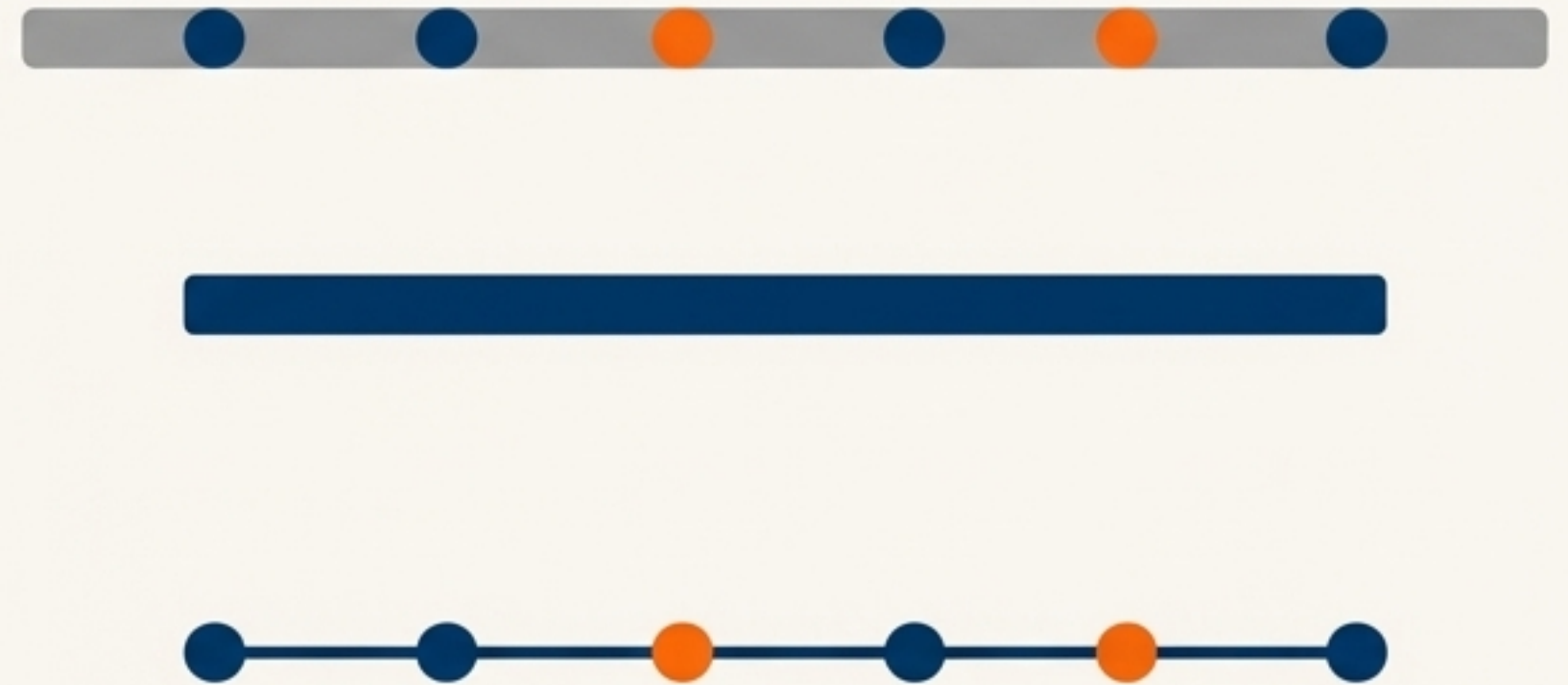
短鎖型シーケンサの限界

短鎖型シーケンサ (Short Reads)



位相決定のエラー（スイッチエラー）が発生し、
正確なハプロタイプの再構築が困難。

長鎖型シーケンサ (Long Reads)



1分子で遺伝子全体をカバーし、
正確な「線」を決定可能。

技術革新がゲノム解読の新たな扉を開く

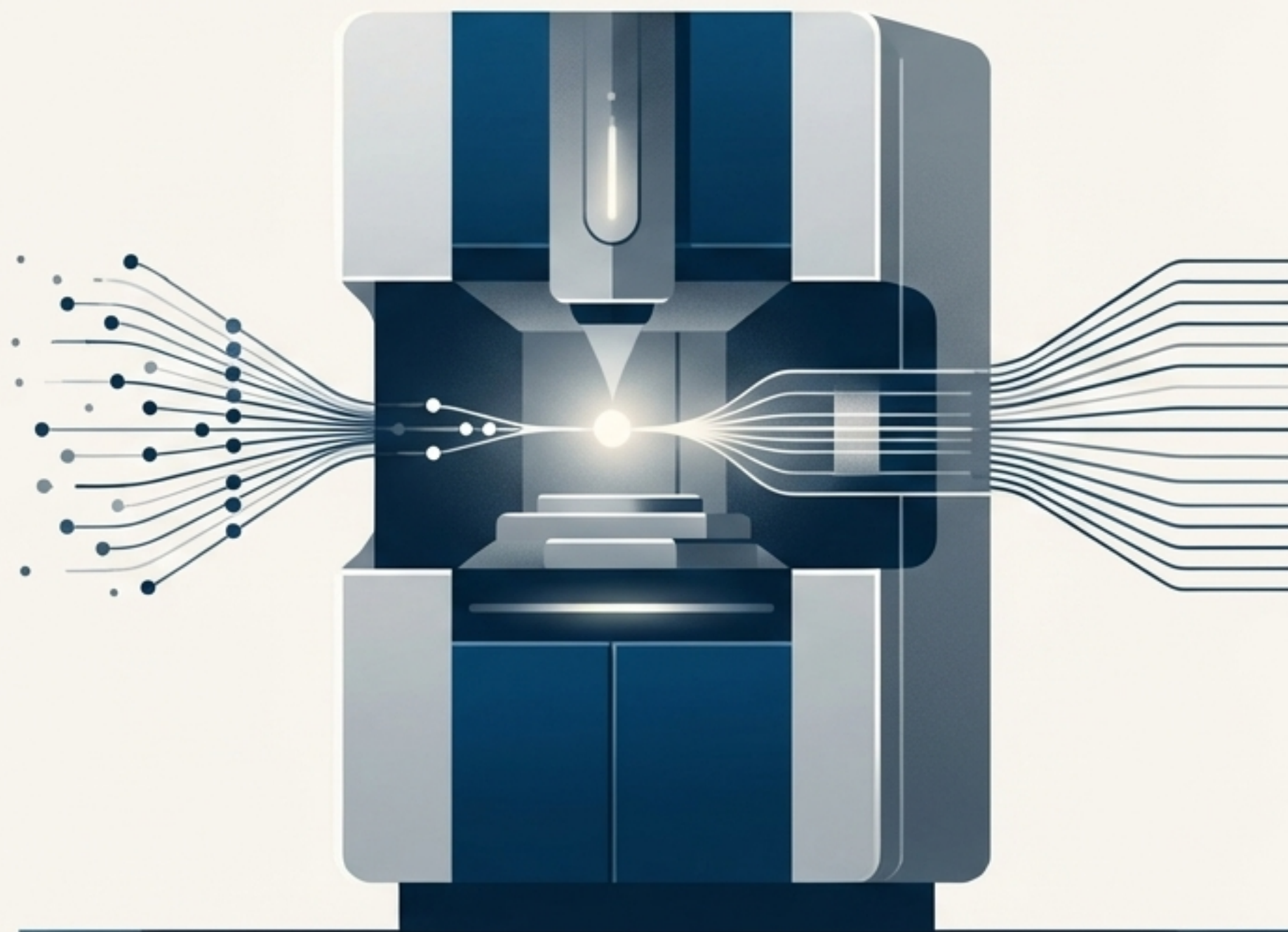
高精度長鎖型シーケンサ（HiFi-LRS）の登場

Accuracy: 1塩基あたりのエラー率が大幅に改善

Length: 数万塩基のDNA断片を一度に解読

Impact: これにより、ヒト遺伝子全体のハプロタイプを網羅的かつ正確に決定することが初めて可能になった。

「遺伝子全体の塩基配列型を正確に決めることが可能になりつつあります。」



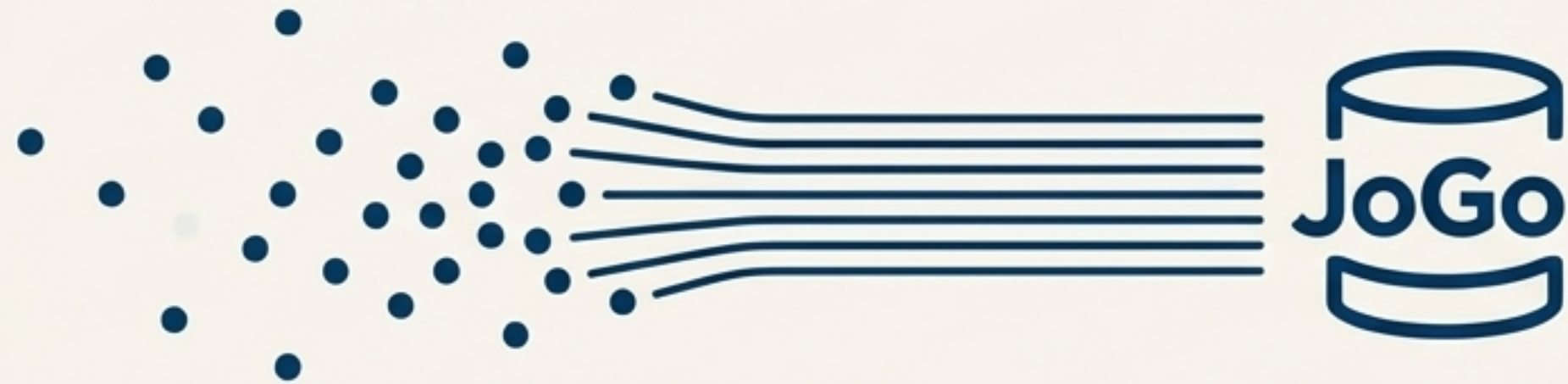
世界初の遺伝子型辞書「JoGo」の構築



JoGo

Joint Open Genome and Omics Platform

ヒト全遺伝子の“型”を統一規則で記述する共通言語を確立する。

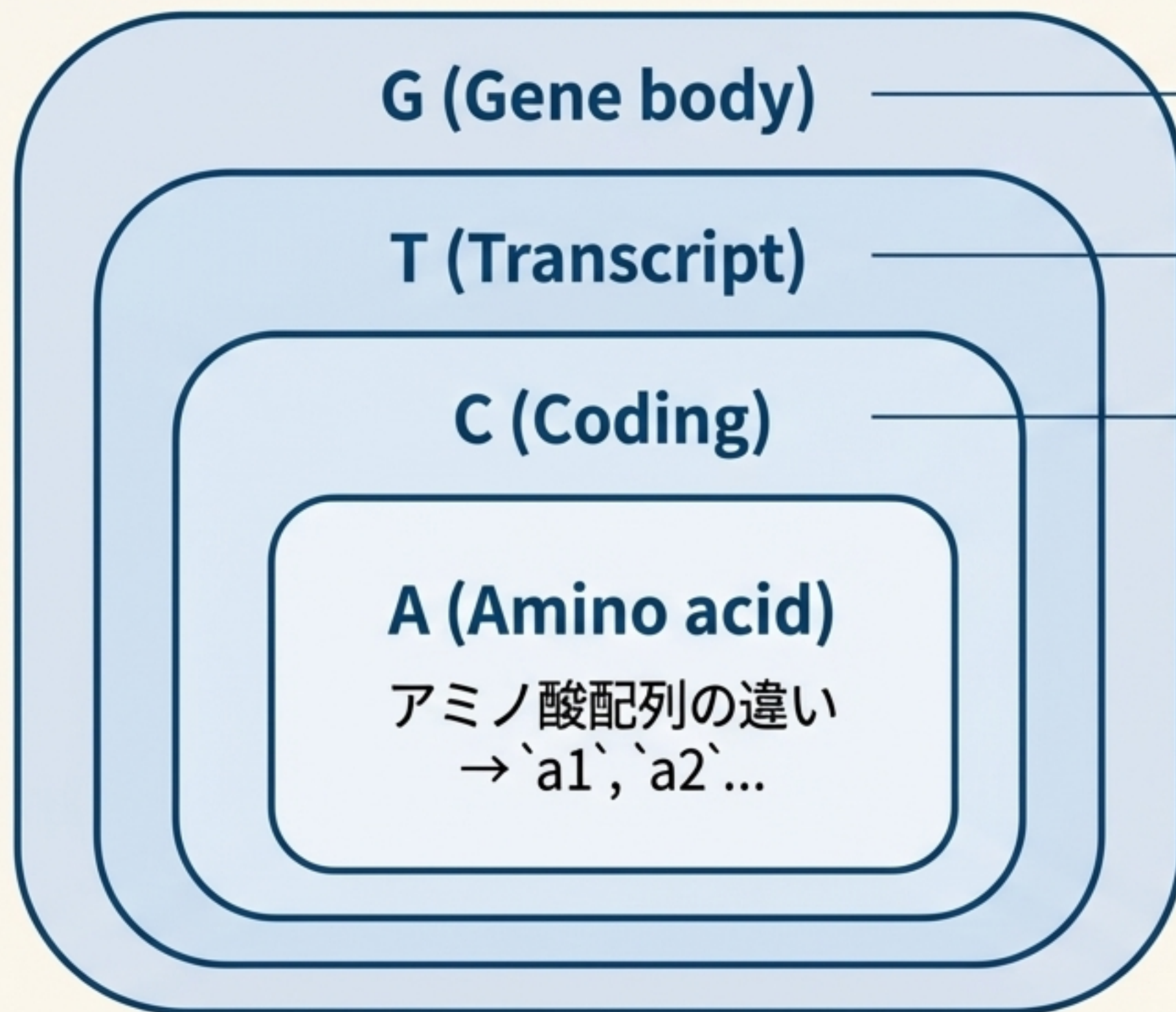


- ロングリードシーケンサ情報に基づく世界初の網羅的ハプロタイプデータベース

- 国際的な学術誌 *Nucleic Acids Research* にて、特に重要性の高い論文（上位2-3%）に与えられる **Breakthrough Article** として採択。

<https://jogo.csml.org>

遺伝子型を体系化する「ACTG階層命名法」



イントロンを含む遺伝子全体の配列の違い → `g1`, `g2`...

5'/3' UTRを含む転写産物全体の違い → `t1`, `t2`...

同義置換を含むコーディング領域の違い → `c1`, `c2`...

The Rule of Frequency

IDは世界集団における
出現頻度順で付番される



`a1`は`a2`よりも頻度が高いか同等

JoGo 1.0が網羅するゲノム多様性のスケール

258

人のグローバルゲノム

5大陸の世界集団をカバー
うち108人分は本プロジェクトでシーケンス

19,194

個のヒト遺伝子

(MANE-standardized)

470万

の総ハプロタイプ

A-level: 174,376 | C-level: 300,610
T-level: 486,288 | G-level: 3,695,204

国際参照配列（GRCh38, CHM13v2）も同一の番号空間に収載し、直接比較が可能。

Case Study 1: *HBB*遺伝子と遺伝性疾患

a1 — M V H L T P E E K S ... K K T E S K E M

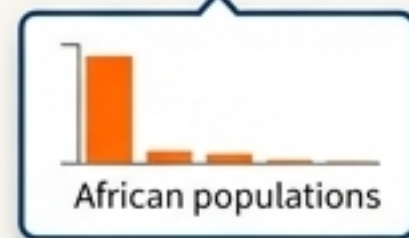
a1 : 正常型ヘモグロビン

a2 — M V H L T P E E **V** S ... K K T E S K E M

a2 : 鎌状赤血球症の原因
- (p.Glu7Val)

a3 — M V H L T P E E **K** S ... K K T E S K E M

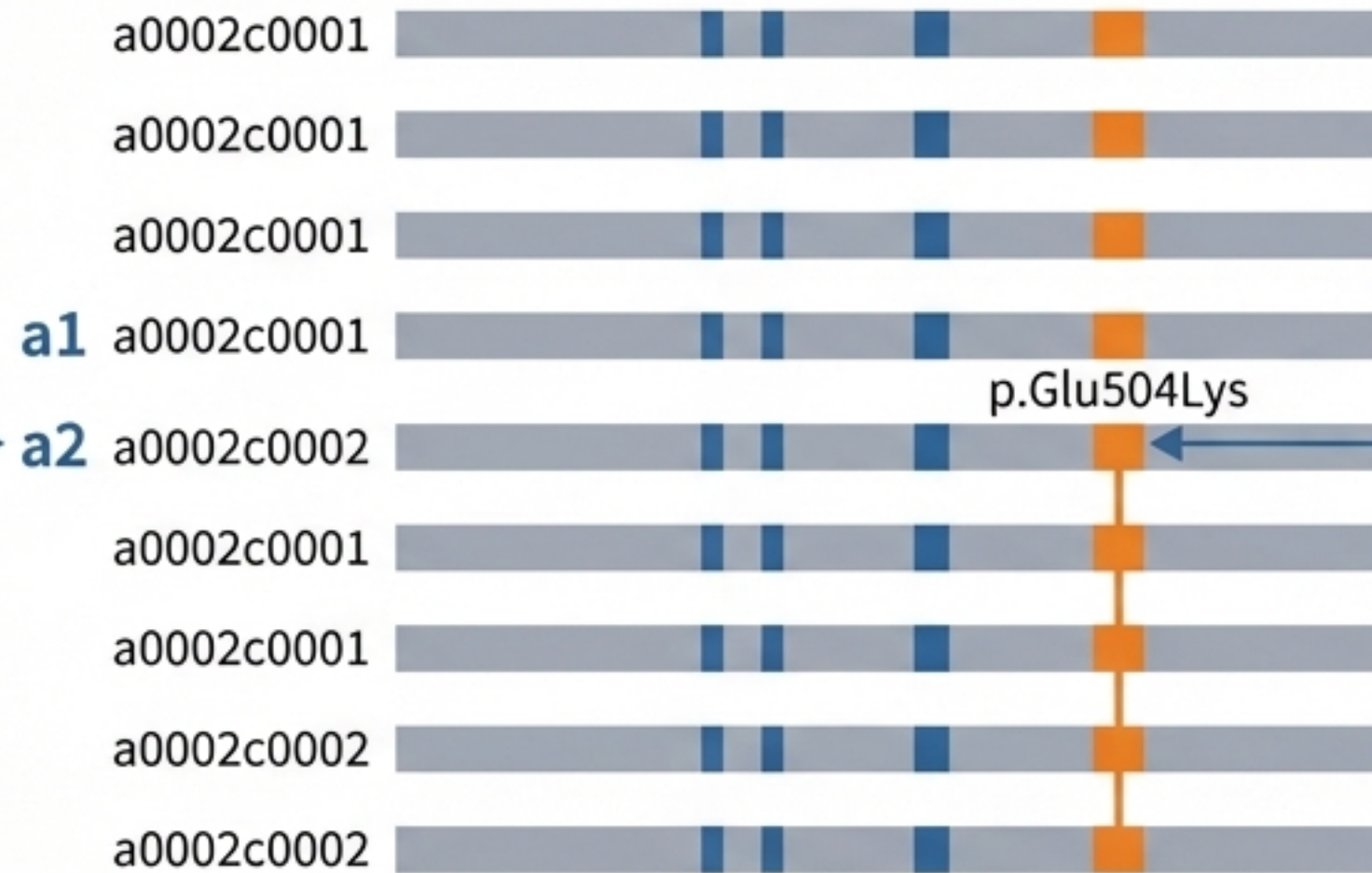
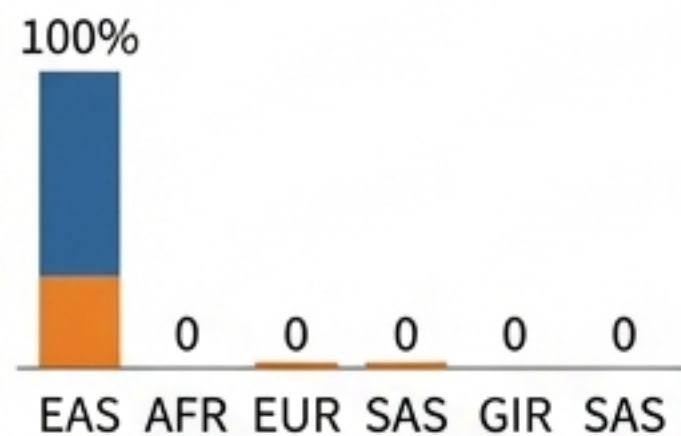
a3 : β サラセミアなどに関連
- (p.Glu7Lys)



JoGoの命名法とビューアにより、ハプロタイプの配列差と臨床的意義 (ClinVar情報) が直感的に結びつく。

Case Study 2: *ALDH2*遺伝子型と集団差

ALDH2:a2型は東アジア
集団に高頻度。



アルコール感受性



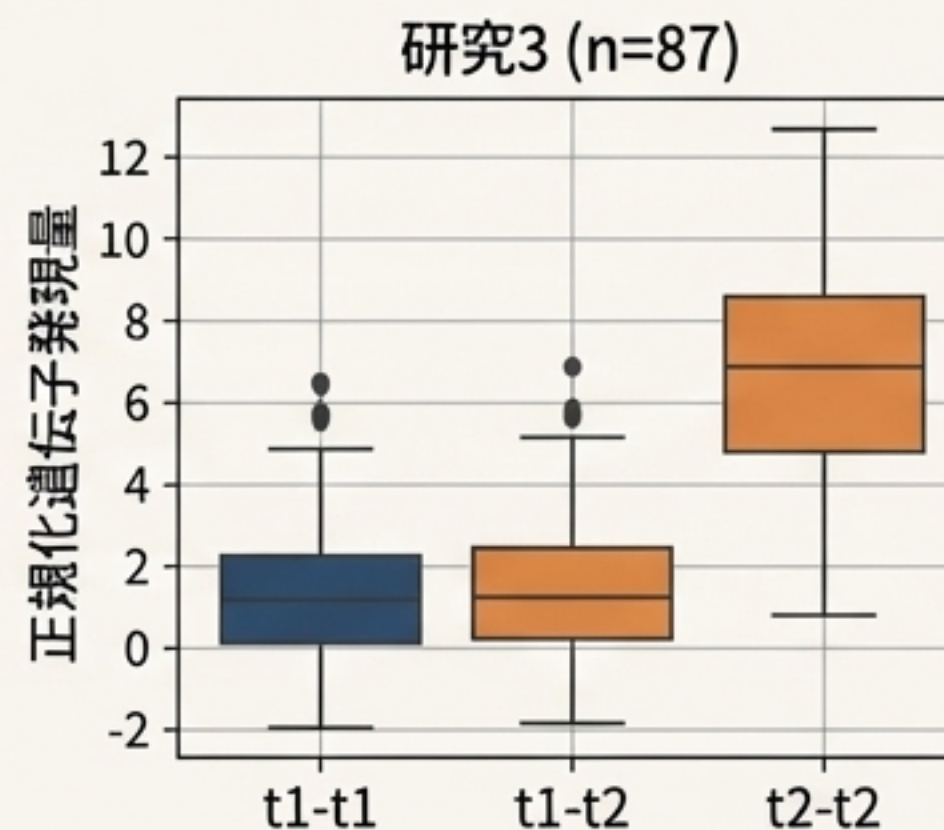
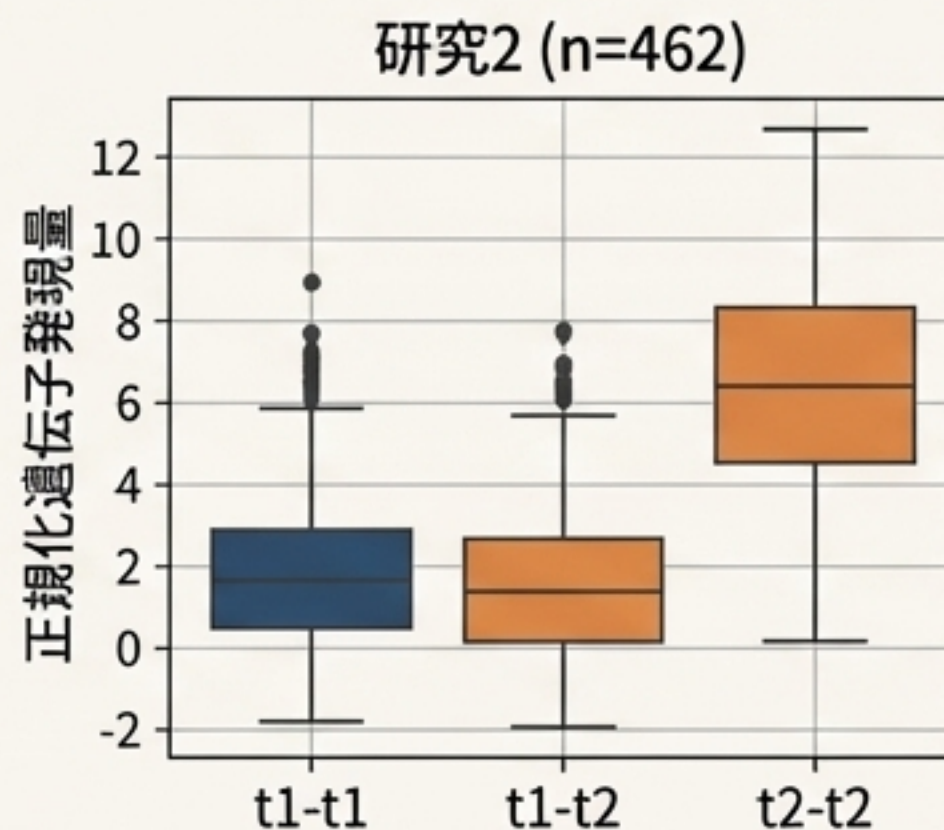
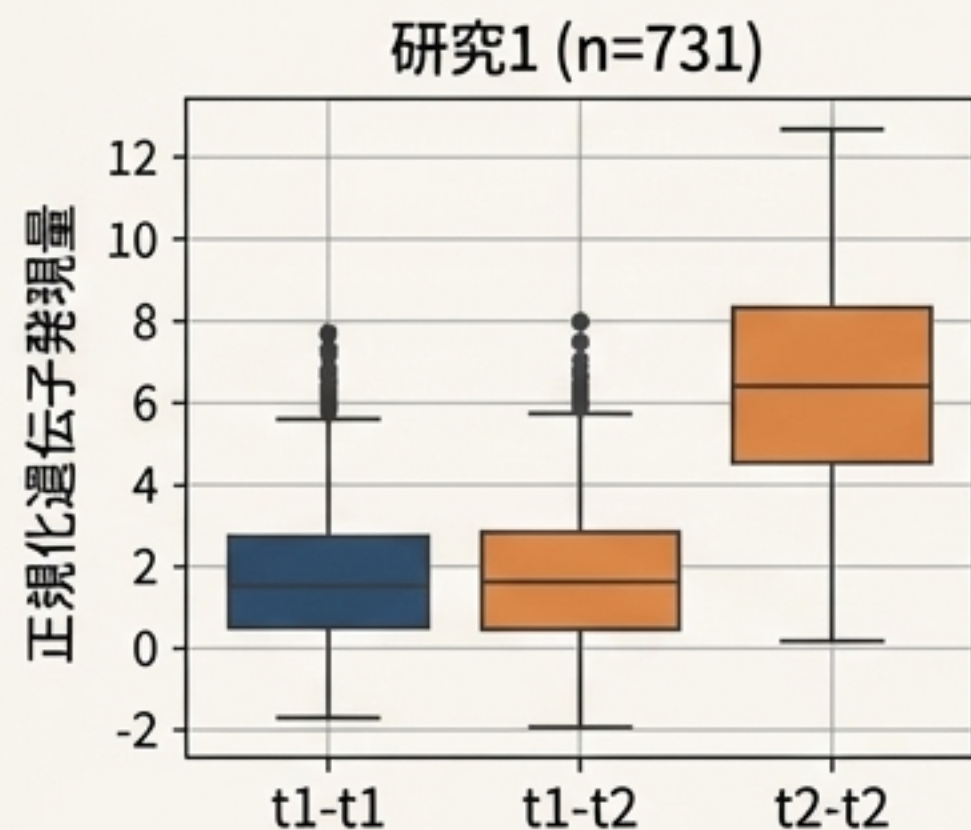
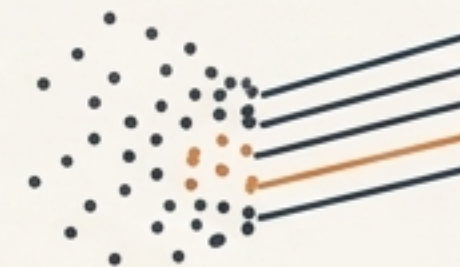
食道がんリスクな
どと関連。

ハプロタイプを軸に、集団特異的な遺伝的特徴とその機能的・臨床的な帰結をワンストップで探索できる。

カタログから発見へ：ハプロタイプQTL解析

遺伝子型（線）と遺伝子発現量の関連を直接解析（*METTL27*遺伝子）

t1 – GCCCCAAAAGATGGGGGGGGCCCTG-TTTTC**C**AAGATTGCTCCGGCCGGGTGG –
t2 – GCCCCAAAAGATGGGGGGGGCCCTG-TTTTC**A**AAGATTGCTCCGGCCGGGTGG –
t3 – GCCCCAAAAGATGGGGGGGGCCCTG-TTTT**C**AAGATTGCTCCGGCCGGGTGG –



JoGoは、ハプロタイプ構造が遺伝子制御や表現型にどのように影響するか、仮説を立てるための強力なツールとなる。

あなたの研究データを安全に：プライバシー保護型ローカビューア Noto Serif JP Bold

Problem: 患者ゲノムなど、公開できない機密性の高いデータを扱いたい

Solution: ローカルハプロタイプビューア

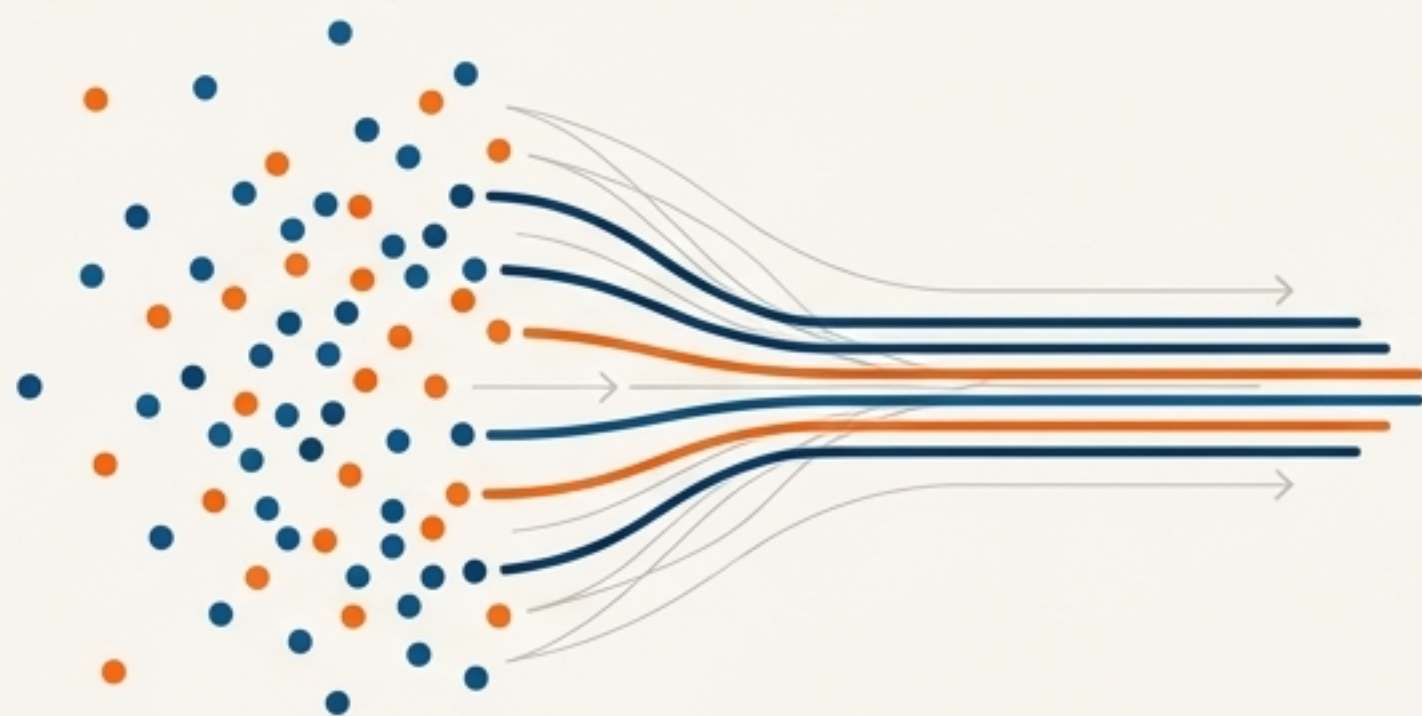
JoGo 1.0のハプロタイプ辞書
(ダウンロード)

あなたの非公開ゲノムデータ



個人データを外部サーバーにアップロードすることなく、
JoGoの参照ハプロタイプと安全に統合・比較・解析が可能。

ゲノム研究は「点」から「線」の時代へ、そしてその先へ

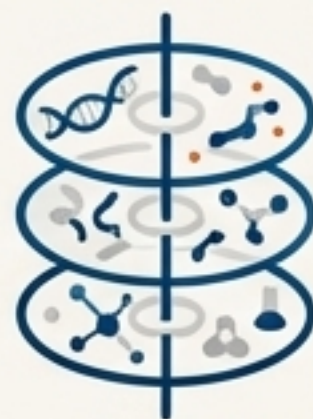


個々の変異（点）からハプロタイプ（線）へ。JoGoはこのパラダイムシフトを加速させる基盤となる。



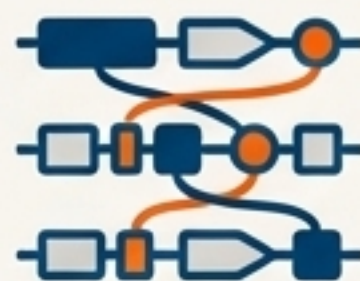
多様性の拡大

未解明な希少ハプロタイプ同定のため、参照集団の多様性向上と人数拡大



多階層オミクス統合

臓器ごとの遺伝子発現（Haplotype-QTL）やプロテオーム、メタボロームとの関連解析



パンゲノムへの展開

T2Tアセンブリを骨格とし、遺伝子単位の構造や順序を表現

「遺伝子型の辞書」が拓く個別化医療の未来

高解像度な疾患解析

疾患関連遺伝子の真の機能単位を特定し、病態解明を深化させる

創薬研究の加速

薬剤応答性に関わるハプロタイプを同定し、効果的な治療標的を発見する

ゲノム医療の基盤確立

全遺伝子を対象とした「遺伝子型に基づく個人特性の理解」を支える

ヒト遺伝子の辞書を、 あなたの研究に。

JoGo: <https://jogo.csml.org>

本研究は多くの共同研究者の協力のもとで行われました。

