



JoGo: 機能階層を表す470万型のACTGハプロタイプ辞書とJoGo-Nextへの展開



○長崎正朗 (九州大)、河合洋介 (遺伝研)、黒木陽子 (NCCHD)、片山俊明 (DBCLS)、守屋勇樹 (DBCLS)、関谷弥生 (九州大)、川島秀一 (DBCLS)、寺岡凌 (九州大)、町田宗隆 (九州大)、松原太一 (九州大)、橋本洋希 (九州大)、浅倉章宏 (九州大)、大平正貴 (東京大)、山下理宇 (NCC)、高田豊行 (理研BRC)、三橋信孝 (DBCLS)、鎌田真由美 (北里大)、大川恭行 (九州大)、徳永勝士 (JHS)、Variant Information Standardization Collegium

① 研究開発の構想 (背景)

2023-2025年度 DICP開発課題(当時)

日本人塩基配列情報の公開可能なゲノム・オミクス情報基盤による双方向型研究教育データベース開発と国際連携

【目標】塩基配列を公開可能な100体級の日本人のゲノム・オミクス情報に加え、ヒトゲノム長鎖型の構造多型情報解析規格の統一を進めているCoLoRSとの国際連携を通じ国内外の約2,000体級の統計情報とを統合する。また、得られた構造多型はVISCで定められるRDFを活用予定である。これにより、希少疾患から多因子疾患まで構造多型の理解に役立つ日本人ゲノム情報基盤の役割を果たす。

JoGo1.0ポータルの開発と一般公開

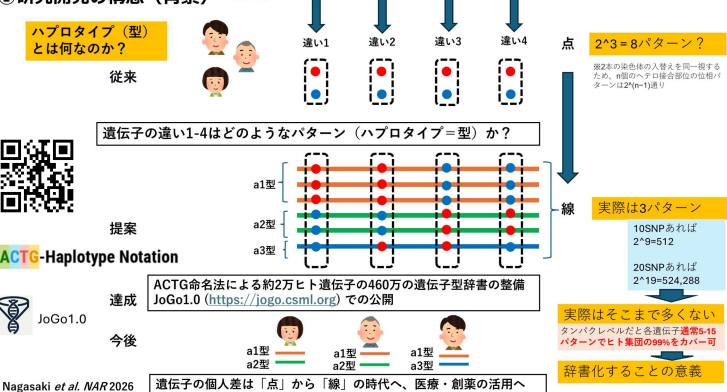
ACTG-Haplotype 命名法の開発

https://jogo.cslm.org/ 2025.8一般公開

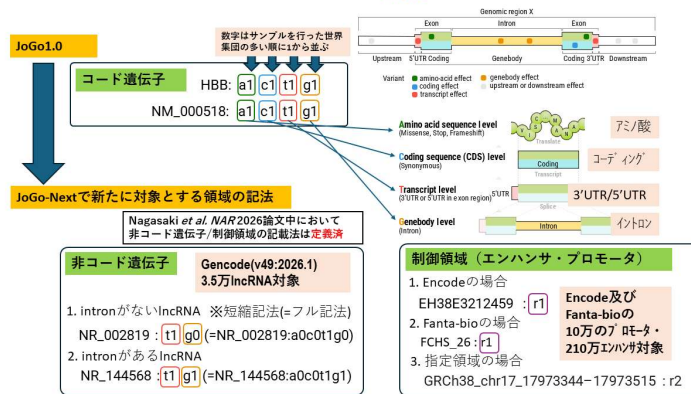


NAR DB Issue 2026.1掲載 (雑誌社により上位1%以内が選ばれるBreakthrough Articleの称号付与を受けた)

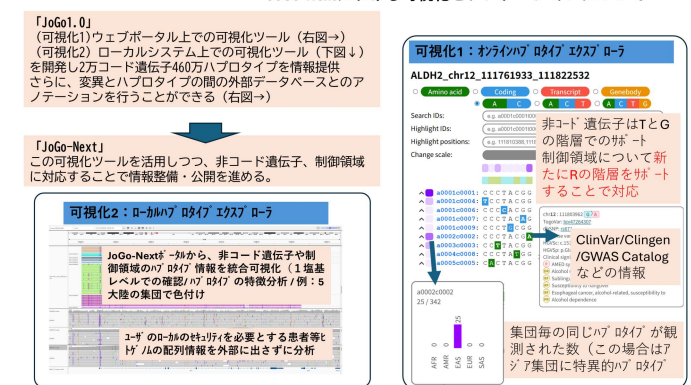
① 研究開発の構想 (背景)



③ 研究開発計画 (実施計画) JoGo-NextにおけるACTG-ハプロタイプ法について

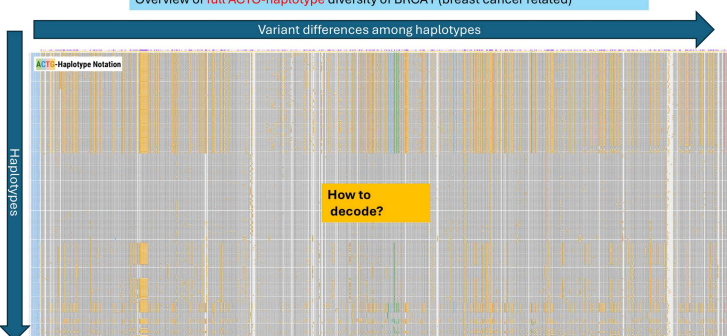


JoGo-Nextにおける可視化とアノテーションについて



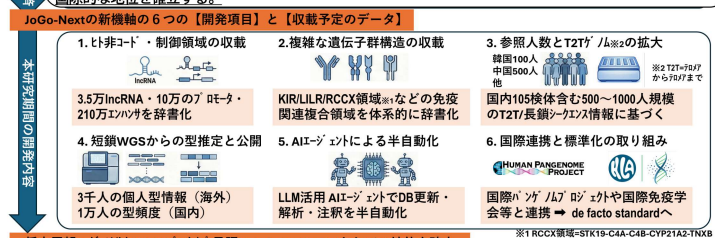
JoGo 1.0: BRCA1 ACTG-Haplotype Diversity Map

Overview of full ACTG-haplotype diversity of BRCA1 (breast cancer related)



① 研究開発構想 (概要・開発項目・収載予定のデータ)

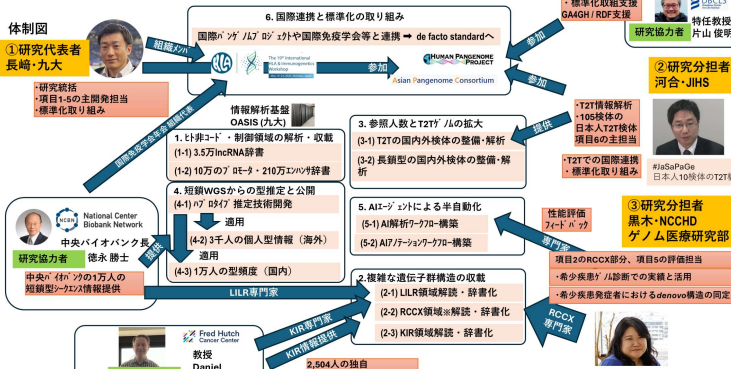
【題目】JoGo-Next: 国際標準に向けたヒトハプロタイプデータベース構築とAI活用
【目標】申請者は、世界で唯一の2万のヒトコーディング遺伝子を対象とした460万種からなるハプロタイプ辞書を構築2025年8月にJoGo1.0として公開をした。DBの国際標準NAR DB Issue 2026においてBreakthrough Articleとしてtop3%のDBとしての称号が付与された。本研究提案「JoGo-Next」では、非コード遺伝子・制御領域を対象とし、さらに対象サンプル数を大幅に拡大することで、唯一のヒトハプロタイプ辞書としてのde facto standardとしての国際的地位を確立する。



研究者に加えAIも活用できる「ゲノムハプロタイプ」言語: 数百万〜数千万ハプロタイプのハプロタイプ情報で創薬や診断の仮説生成を加速

創薬・医療現場への貢献
疾患リスクや個人差をより正確に解釈

③ 研究開発計画 (6つの実施計画と研究開発体制)



④ 利用者像

ヒト
ヒトの変異とハプロタイプの関係、ハプロタイプに関する知識ドメインに関する情報取得。また、そのハプロタイプを得るために付随する元のデータの情報取得。一中学、高校、大学、大学院の教育、研究、医療においてもハプロタイプ辞書として使うことを想定。
【使用者の主な研究領域としては以下を想定】
ヒト遺伝子を対象としたゲノムサイエンス研究者全般、希少疾患、多因子疾患、単一遺伝子疾患など幅広い疾患に関連する研究者、バイオインフォマティクス、システム生物学、集団遺伝学など

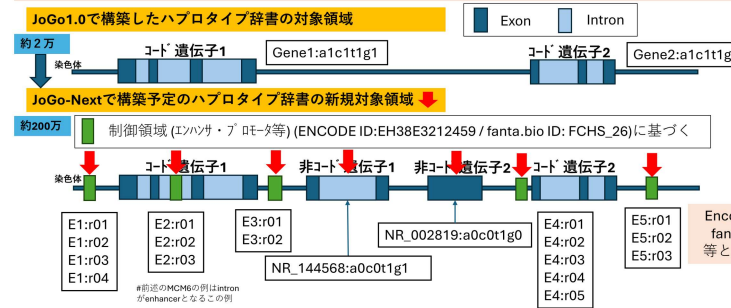
AI
赤色の部分がJoGo1.0に加えてJoGo-Nextを開発することで追加できる情報
ChatGPT、Claude、Geminiとの連携、また、これらAI情報に取り込まれることでAIの中でハプロタイプ辞書の単語やアノテーションが活用される。local LLMの機能拡充にともないlocalでの利用も視野に入る。

【例1】MCM6のintronにあるrs4988235がCT (ラクターゼ) の転写制御にエンハンサーとして関連することが知られています。このrs4988235をもつハプロタイプはJoGo-Nextでは、r2型で、MCM6遺伝子のa1:c1:t1型にみられ、その約半数でこのパターンを持ちます。

【例2】APOE遺伝子は抗アミロイドベータ (Aβ) 抗体薬投与の副作用がハプロタイプによって違うことが知られています。rs429358がリスクになることが知られていますが、このハプロタイプはa2型です。さらに、このa2型は、JoGo-Nextでは、EH9999.r02の場合、発現が高値になることがわかっています。EH999.r02のAPOC1:a1c1とAPOC4:a2c1の組み合わせが日本では80%、ヨーロッパでは50%、アフリカでは確認されません (解析が進むことでこのようアノテーションが可能となる)。

⑤ 学術的・社会的に期待されるインパクト

「ヒトハプロタイプ辞書によるヒトとAIへの網羅的な情報基盤」としての貢献
ヒトとAIの双方にとって、膨大な塩基配列の差分を意味のある機能単位 (ハプロタイプ) へ縮約できることは、新しい「言語」を獲得することに等しい。これにより、①タバコ質全長の配列差に基づく機能差に加え、②EIA/βアプテ等の調節要素の組み合わせに基づく発現差、③遺伝子本体と調節領域の組み合わせに基づく表現型差といった、従来以上に付加価値の高いハプロタイプ単位の関係性を、ゲノム・オミクス測定情報の統合から導出する基盤となる。



想定される活用事例

