

2024 年度 研究開発実施報告

概要

研究開発課題名	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
開発対象データベースの名称(URL)	Joint Open Genome Omics Platform (https://jogo.csml.org/)
研究代表者氏名	長崎 正朗 (90396862)
所属・役職	九州大学 生体防御医学研究所 附属高深度オミクスサイエンスセンター バイオメディカル情報解析分野 教授 (2025 年 3 月時点)



□目次

概要	1
§1. 研究実施体制	3
§2. 研究開発対象とするデータベース・ツール等	4
(1) データベース一覧	4
(2) ツール等一覧	4
§3. 実施内容	5
(1) 本年度に計画されていた研究開発項目・タスク	5
(2) 進捗状況	7
§4. 成果発表等	14
(1) 原著論文発表	14
① 論文数概要	14
② 論文詳細情報	14
(2) その他の著作物(総説、書籍など)	14
(3) 国際学会および国内学会発表	15
① 概要	15
② 招待講演	15
③ 口頭講演	15
④ ポスター発表	15
(4) 知的財産権の出願（国内の出願件数のみ公開）	16
① 出願件数	16
② 一覧	16
(5) 受賞・報道等	16
① 受賞	16
② メディア報道	16
③ その他の成果発表	16
§5. 主要なデータベースの利活用状況	17
1. アクセス数	17
① 実績	17
② 分析	17
2. データベースの利用状況を示すアクセス数以外の指標	17
3. データベースの利活用により得られた研究成果(生命科学研究への波及効果)	18
4. データベースの利活用によりもたらされた産業への波及効果や科学技術のイノベーション(産業や科学技術への波及効果)	18
§6. 研究開発期間中に主催した活動(ワークショップ等)	19
(1) 進捗ミーティング	19
(2) 主催したワークショップ、シンポジウム、アウトリーチ活動等	19

§1. 研究実施体制

グループ名	研究代表者・ 研究分担者 氏名	所属機関・役職名	研究題目
長崎グループ	長崎 正朗	九州大学・教授	JoGoデータベース構築と国際連携
河合グループ	河合 洋介	国立国際医療研究センター・副プロジェクト長	JoGoデータベースRDF化のためのアノテーション情報の整理と外部DBとの連携

§2. 研究開発対象とするデータベース・ツール等

(1) データベース一覧

【主なデータベース】

No.	名称	別称・略称	URL
1	Joint Open Genome Omics Platform	JoGo	https://jogo.csml.org/

【その他のデータベース】

No.	名称	別称・略称	URL
1			

(2) ツール等一覧

No.	名称	別称・略称	URL
1			

§3. 実施内容

(1) 本年度に計画されていた研究開発項目・タスク

①研究開発の構想（全体概要）

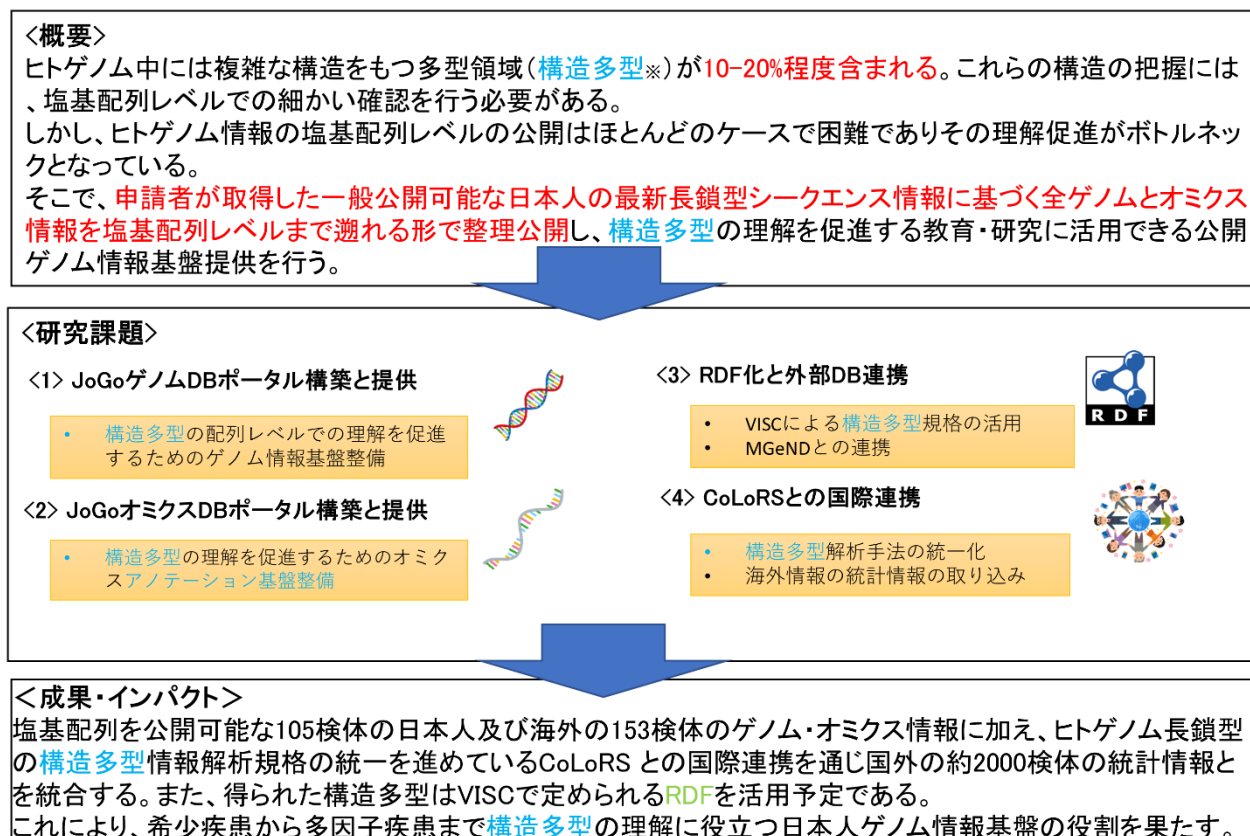


図1に通年の全体計画の概要を示す。

（長崎グループ）

本研究の JoGo データベースの開発と公開の統括を行うとともに有識者との定期的な意見交換、国際連携を主担当として行う。具体的には、実施計画に記載をした長鎖型ゲノムシーケンス情報に関連する【1-1】「JoGo ゲノム DB ポータル構築と提供」と【2-1】「JoGo オミクス DB ポータル構築と提供」のオミクス情報に基づく【2-1-1】「eQTL アノテーション情報の整理」ならびに 【2-1-2】「ATACQTL アノテーション情報の整理」の各研究開発項目について主担当者として研究開発を行う。さらに、国際連携として、【4】「CoLoRS との国際連携」の研究開発項目を推進する。

当該年度は、

【1】 JoGo ゲノム DB ポータル構築と提供（長崎）

- 【1-1-1-1】 SNV（構築）
- 【1-1-1-2】 SNV（提供開始および更新）
- 【1-1-2-1】 STR（構築）
- 【1-1-2-2】 STR（提供開始および更新）
- 【1-1-3-1】 SV（構築）

- 【1-1-3-2】SV (提供開始および更新)
- 【1-1-4-1】 アラメントデータベース (構築)
- 【1-1-4-2】 アラメントデータベース (提供開始および更新)
- 【1-1-5-1】 phasing データ (構築)
- 【1-1-5-2】 phasing データ (提供開始および更新)

【2】JoGo オミクス DB ポータル構築と提供 (長崎)

- 【2-1-1】eQTL アノテーション情報の整理
- 【2-1-2】ATACQTL アノテーション情報の整理
- 【2-1-3】公共 ISO-SEQ の整備と提供

【4】CoLoRS との国際連携 (長崎)

- 【4】CoLoRS 等との国際連携の継続

の主担当を行う(実施内容詳細は 2 に記載)。

(河合グループ)

実施計画に記載をした【3】「RDF 化のためのアノテーション情報の整理と外部 DB との連携」のサブ研究開発項目について主な担当者として研究開発を行う。また、【1-1】「JoGo ゲノム DB ポータル構築と提供」について、国立国際医療研究センター・中央バイオバンクが中心となって進めている SRS の解析パイプラインを適用することで、長鎖型シーケンサでの各変異の検出精度が評価できる体制を構築する。さらに、DBCLS のコアメンバが定期的に行っているバリエーション情報標準化研究会に参画し構造多型の RDF 規格の意見交換と実装のための連携を進める。他に、MGeND の運営母体であることから、JoGo DB と MGeND との連携実装のための橋渡しの役割を担う。

当該年度は、

【3】RDF 化のためのアノテーション情報の整理と外部 DB との連携 (河合)

【3-1】RDF 連携のための会議参加と定期意見交換 (DBCLS)

の主担当を行う(実施内容詳細は 2 に記載)。

(2) 進捗状況

図2 本年度の研究開発（令和6年度）

【体制図と目的】計画(1)(2)(4) 長崎 (3) 河合を担当し研究計画を推進

(1) JoGo ゲノム DBポータル構築と提供 長崎

- 構造多型の配列レベルでの理解を促進するためのゲノム情報基盤整備

【1-1-1-1】 SNV (構築) DeepVariant / Clair
 【1-1-1-2】 SNV (提供開始および更新)
 JoGo Portalの公開
 【1-1-2-1】 STR (構築) TRGT v0.5
 【1-1-2-2】 STR (提供開始および更新)
 【1-1-3-1】 SV (構築) Sniffles2
 【1-1-3-2】 SV (提供開始および更新)
 【1-1-4-1】 アラメントデータベース (構築)
 【1-1-4-2】 アラメントデータベース (提供開始および更新) Minimap2
 【1-1-5-1】 phasingデータ (構築) Whatsap / longphase
 【1-1-5-2】 phasingデータ (提供および更新)

本年度の達成目標

公開可能な国内外258検体のゲノム情報のJoGoポータル (early access) での公開と更新

(4) CoLoRSとの国際連携 長崎

- 構造多型解析手法の統一化
- 海外情報の統計情報の取り込み

CoLoRSの参画の継続

(2) JoGo オミクス DBポータル構築と提供 長崎

- 構造多型の理解を促進するためのオミクスアノテーション基盤整備

【2-1-1】 eQTLアノテーション情報の整理
 【2-1-2】 ATACQTLアノテーション情報の整理
 【2-1-3】 公共ISO-SEQの整備と提供

本年度の達成目標

公共データのISO-SEQの情報の公開 (early access) と日本人の検体のオミクス情報の公開に向けた整理

(3) RDF化と外部DB連携 河合

- VISCによる構造多型規格の活用
- MGeNDとの連携

【3】RDF化のためのアノテーション情報の整理と外部DBとの連携

【3-1】外部DB連携のための会議参加と定期意見交換

【3-3】外部DB連携方針検討と外部連携

本年度の達成目標

MGeNDとの連携に向けた試験実装
 DBCLSの基盤DBとの連携の試験実装

(長崎グループ)

本年度は通期計画書のロードマップ(本年度の研究開発概要は図2に記載)、特に以下の課題を推進した。令和6年度は、追加計画に基づき、日本人の検体に加え、海外の153検体の解析を行うこととなったため、それらの情報を合わせた解析を進めた。また、JoGoのウェブサイトについてβ版の開発は完了しており、有識者によるアーリーアクセスによる評価と改良を進めた。

図 3

Joint Open Genome and Omics Platform (JoGo)

Home > Search by Variant

Search by Variant

Showing 1 to 50 of 629 results

VIEW	REGIONNAME	POS	REF	ALT	ANNOTATION	IMPACT	SAMPLES	AHAPIDS	CHAPIDS	THAPIDS
	A3GALT2_chr1_33301766_33326098	33301803	A	AGGCTT	downstream_gene_variant	MODIFIER	NA18939.hp2	a0001	a0001:c0001	a0001:c0001:t0001
	A3GALT2_chr1_33301766_33326098	33301803	A	AGGCTTGG others(13): Show Hide	downstream_gene_variant	MODIFIER	HG01433.hp2	a0001	a0001:c0001	a0001:c0001:t0001
	A3GALT2_chr1_33301766_33326098	33301808	T	TG	downstream_gene_variant	MODIFIER	NA18972.hp1	a0001	a0001:c0001	a0001:c0001:t0001
	A3GALT2_chr1_33301766_33326098	33301838	TG	T	downstream_gene_variant	MODIFIER	NA18939.hp2	a0001	a0001:c0001	a0001:c0001:t0001
	A3GALT2_chr1_33301766_33326098	33301917	C	CA	downstream_gene_variant	MODIFIER	NA18949.hp2 NA18972.hp1	a0001	a0001:c0001	a0001:c0001:t0001
	A3GALT2_chr1_33301766_33326098	33301917	CA	C	downstream_gene_variant	MODIFIER	NA18939.hp2	a0001	a0001:c0001	a0001:c0001:t0001

【1-1-1】 データベースの構築 (SNV / SV / STR)

【1-1-1-1】 SNV (構築)

【1-1-4-1】258 検体について「アラメントデータベース (構築)」のアライメント結果を進めるとともに、DeepVariant を使うことで変異コールを行った。

【1-1-1-2】 SNV (提供開始および更新)

105 検体から 258 検体の情報に拡充することでデータベースを構築するとともに、JoGo ポータル経由で【1-1-1-1】「SNV (構築)」の結果を提供するためのウェブインターフェースの開発を行った(図3)。

【1-1-2-1】 STR (構築)

【1-1-2-2】 STR (提供開始および更新)

【1-1-4-1】105 検体から 258 検体の情報に拡充した「アラメントデータベース (構築)」のアライメント結果に対して、TRGT のソフトウェアを適用することで STR の変異コールを進めた。また、これらの結果について整備をすることで、令和 7 年度の【1-1-2-2】で必要となるデータベースとして提供できる準備を整えた。特に、TRGT については、昨年度から高頻度でバージョンアップがおこなわれていることから、できる限り新しい版での提供を予定している。

図 4

Joint Open Genome and Omics Platform (JoGo)

Showing 1 to 50 of 330818 results

VIEW	REGIONNAME	POS	REF	ALT	ANNOTATION	IMPACT	SAMPLES	AHAPIDS	CHAPIDS	THAPIDS	THAPIDS	AC	AN	AF	LEN	HGVSC
	TMEM242_chr6_157284025_157328519	15731158	A	AGCCTCAT others(2986): Show Hide	intron_variant	MODIFIER	HG02280.hp2	a0001	a0001:c0001	a0001:c0001:0003	a0001:c0001:0003:g0024	1	316	0.0032	29993	c.327 others(30010): Show Hide
	TMEM242_chr6_157284025_157328519	157311768	A	GGCCTCAT others(29915): Show Hide	intron_variant	MODIFIER	HG02145.hp2	a0001	a0001:c0003	a0001:c0003:0002	a0001:c0003:0002:g0001	1	258	0.0039	29922	c.327 others(29939): Show Hide
	NBPF10_chr1_146059711_146149804	146096665	G	GAATTATT others(29897): Show Hide	intron_variant	MODIFIER	NA19043.hp2	a0105	a0105:c0001	a0105:c0001:0001	a0105:c0001:0001:g0001	1	282	0.0035	29904	c.647 others(29921): Show Hide
	TMEM242_chr6_157284025_157328519	157311760	A	ACTCACCT others(29885): Show Hide	intron_variant	MODIFIER	HG01891.hp2	a0001	a0001:c0001	a0001:c0001:0010	a0001:c0001:0010:g0001	1	178	0.0056	29892	c.327 others(29909): Show Hide
	NBPF10_chr1_146059711_146149804	146101387	G	GAATTATT others(29873): Show Hide	intron_variant	MODIFIER	HG01433.hp1	a0046	a0046:c0001	a0046:c0001:0001	a0046:c0001:0001:g0001	1	281	0.0036	29880	c.573 others(29897): Show Hide
	NBPF14_chr1_148526385_148601011	148572625	T	TAGGAAAA others(29845): Show Hide	intron_variant	MODIFIER	HG02486.hp1	a0062	a0062:c0001	a0062:c0001:0001	a0062:c0001:0001:g0001	1	215	0.0047	29852	c.258 others(29867): Show Hide
	NBPF14_chr1_148526385_148601011	148572586	T	TCTTCTTC others(29842): Show Hide	frameshift_variant others(1): Show Hide	HIGH	NA18952.hp2	a0002	a0002:c0004	a0002:c0004:0001	a0002:c0004:0001:g0001	1	186	0.0054	29849	c.261 others(29858): Show Hide
	NBPF14_chr1_148526385_148601011	148572625	T	TAGGAAAA others(29830): Show Hide	intron_variant	MODIFIER	HG02922.hp1	a0072	a0072:c0001	a0072:c0001:0001	a0072:c0001:0001:g0001	1	215	0.0047	29837	c.258 others(29852): Show Hide
	NBPF14_chr1_148526385_148601011	148572586	T	TCTTCTTC others(29819): Show Hide	stop_gained others(1): Show Hide	HIGH	NA19056.hp1	a0100	a0100:c0001	a0100:c0001:0001	a0100:c0001:0001:g0001	1	186	0.0054	29826	c.261 others(29835): Show Hide
	TMEM242_chr6_157284025_157328519	157311755	T	AGTGGCCT others(29692): Show Hide	intron_variant	MODIFIER	HG01257.hp2	a0001	a0001:c0001	a0001:c0001:0002	a0001:c0001:0002:g0006	1	259	0.0039	29699	c.327 others(29716): Show Hide

【1-1-3-1】SV (構築)

【1-1-3-2】SV (提供開始および更新)

【1-1-4-1】105 検体から 258 検体の情報に拡充した「アラメントデータベース (構築)」のアライメント結果に対して Sniffle2 などのソフトウェアを適用することで遺伝子領域を中心とした SV の変異コールを行う予定であったが、独自のツールでコールできる状況になったことから、そちらを使うこととして解析を進めた。すでに、SV の検出は完了しており、JoGo ポータル経由で検索できる状況となっている(図4)。また、SV に関連した派生成果として、JoGo-LILR ソフトウェアの開発 (Nagasaki *et al. Human Immunology* 2025)を行うことができた。

【1-1-4-1】アラメントデータベース (構築)

長鎖型由来 (lrWGS) の全ゲノム配列情報 (lrWGS) に対して、アライメントソフトウェアを用いて cram 形式のデータを構築した。GRCh38 のリファレンス配列を選択し、アライメントした結果を提供することで新たに提供される国際参照配列、各ユーザーが興味のある領域について個人の配列レベルで確認できるデータベースを構築することを目標に進めている。令和 6 年度は、アノテーション情報が充実していることから、GRCh38 を選択し 105 検体から 258 検体に拡充し、大規模電算機環境を用いて計算を行った。また、比較が可能なように短鎖型由来 (srWGS) のアライメント結果についてもポータル上で試験的に提供できる体制を整えた。

図 5

Joint Open Genome and Omics Platform (JoGo)

Home > Summary of Gene Haplotype Catalog by Sample

Summary of Gene Haplotype Catalog by Sample

Q Search

Showing 1 to 100 of 258 results

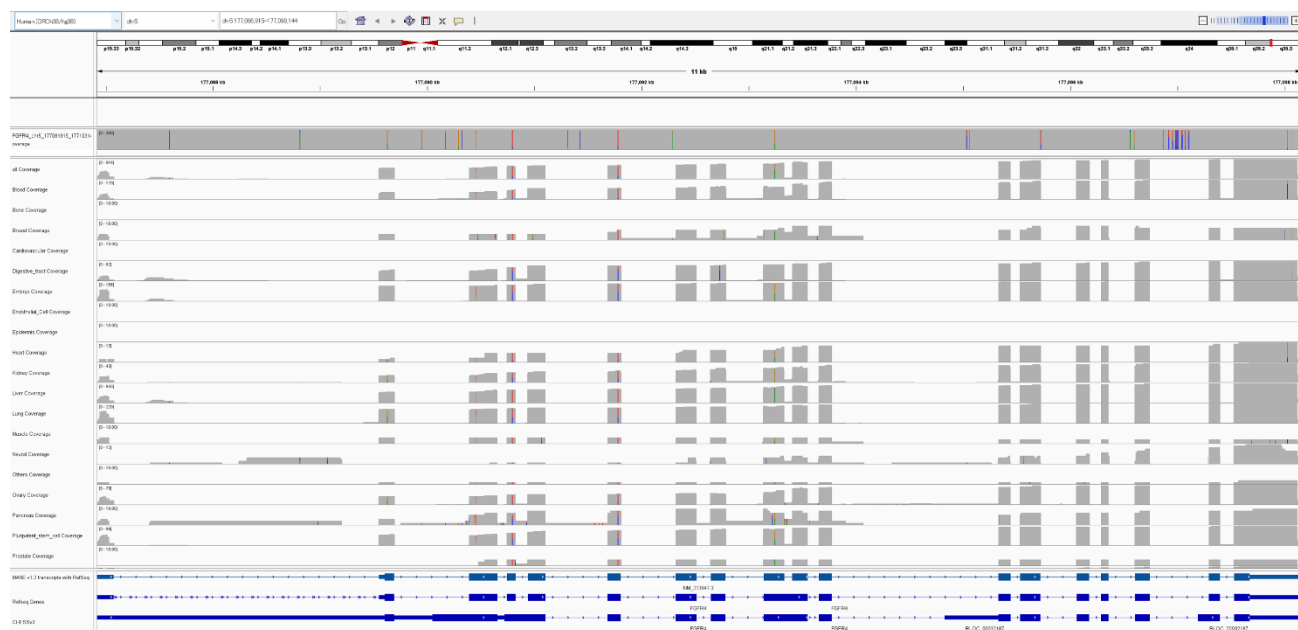
VIEW SRWS	SAMPLEID	POPNAME	GPOPNAME	BOTH	SINGLE	MULTI	NONE	BOTH	SINGLE	MULTI	NONE
srWGS	HG00099	GBR	EUR	14880	3184	76	1099	0.77	0.17	0.00	0.06
srWGS	HG00140	GBR	EUR	14286	3988	61	904	0.74	0.21	0.00	0.05
srWGS	HG00280	FIN	EUR	14160	4027	50	1002	0.74	0.21	0.00	0.05
srWGS	HG00323	FIN	EUR	13751	4328	69	1091	0.71	0.22	0.00	0.06
srWGS	HG00408	CHS	EAS	13788	4242	66	1143	0.72	0.22	0.00	0.06
srWGS	HG00423	CHS	EAS	13313	4495	80	1351	0.69	0.23	0.00	0.07
srWGS	HG00438	CHS	EAS	14619	3540	61	1019	0.76	0.18	0.00	0.05
srWGS	HG00544	CHS	EAS	13549	4379	67	1244	0.70	0.23	0.00	0.06
srWGS	HG00558	CHS	EAS	14227	3882	53	1077	0.74	0.20	0.00	0.06
srWGS	HG00597	CHS	EAS	13669	4263	62	1245	0.71	0.22	0.00	0.06
srWGS	HG00609	CHS	EAS	13569	4589	51	1030	0.71	0.24	0.00	0.05
srWGS	HG00621	CHS	EAS	14015	4184	48	992	0.73	0.22	0.00	0.05
srWGS	HG00639	PUR	AMR	15616	2736	50	837	0.81	0.14	0.00	0.04
srWGS	HG00642	PUR	AMR	14717	3560	62	900	0.76	0.19	0.00	0.05
srWGS	HG00673	CHS	EAS	13554	4455	82	1148	0.70	0.23	0.00	0.06
srWGS	HG00733	PUR	AMR	12361	5300	42	1536	0.64	0.28	0.00	0.08
srWGS	HG00735	PUR	AMR	15721	2486	95	937	0.82	0.13	0.00	0.05
srWGS	HG00738	PUR	AMR	14845	3450	70	874	0.77	0.18	0.00	0.05
srWGS	HG00741	PUR	AMR	15843	2427	99	870	0.82	0.13	0.01	0.05
srWGS	HG01069	PUR	AMR	13291	4787	29	1132	0.69	0.25	0.00	0.06

【1-1-5-1】phasing データ(構築)

【1-1-1-1】「SNV (構築)」の情報を活用することで **phasing** 情報を作成可能である。特にポータル上での可視化を行う際に **phasing** されていることでシーケンスされた配列を色で分けることやグループ化することが可能である。特に構造多型においては、複雑なパターンであることも多く、**phasing** された情報が対象とする染色体領域を確認する上で有用なことが多い。そのため、【1-1-1-1】「SNV (構築)」の **cram** ファイルに **phasing** のタグ情報を埋め込むことで **phasing** 情報を付与した **cram** ファイルの作成を進めている。今年度は、105 検体から 258 検体の情報に拡充したことからこれらの追加検体についても同様の解析を進めた。

なお、昨年度にも報告をしているが、特に、DBCLS の関係者やゲノムサイエンスの有識者と研究会を通じて意見交換を行い、現在の取得されているデータの品質から、特に各遺伝子領域における個人毎のハプロタイプの局所 *denovo* アセンブリにより、常染色体について、両アレルを確定することができた情報が精度が最も高い(スイッチングエラーの可能性が低い)ことから、JoGo v1 については、これらを中心に整備を行い、遺伝子単位での情報を検索して公開することに注力した(図5:各検体の遺伝子のペアハプロタイプ全長を確定できた数をポータルで表示したスクリーンショット)。また、JoGo β 版を VISC の研究会メンバ(DBCLS の関係者やゲノムサイエンスの有識者)等や授業などで活用することでフィードバックを継続して受け概ね公開できる状況となっている。

図6



【2-1-3】【2023 年度追加実施項目】 公共 ISO-SEQ の整備と提供

高精度長鎖型シーケンスに基づく ISO-SEQ (PacBio HIFI) の倫理申請を必要としない公開情報について JoGo オミクスポータルで統一して解析を行うことでユーザーが実際の実験において得られる ISO-SEQ のシーケンス情報のアライメント結果を可視化および取得できるようにすることが目標となっていた。同データベースは、令和5年度において、20臓器200検体のダウンロードと GRCh38 基準の配列へのアライメントが完了していた。令和6年度は、同情報を JoGo のポータル上から可視化アクセスできるように整備を進めた。これにより、JoGo ポータルから提供できる体制となっている(図6:EGFR4 領域の組織毎の発現情報を JoGo 経由で起動した iGV で表示した結果)。

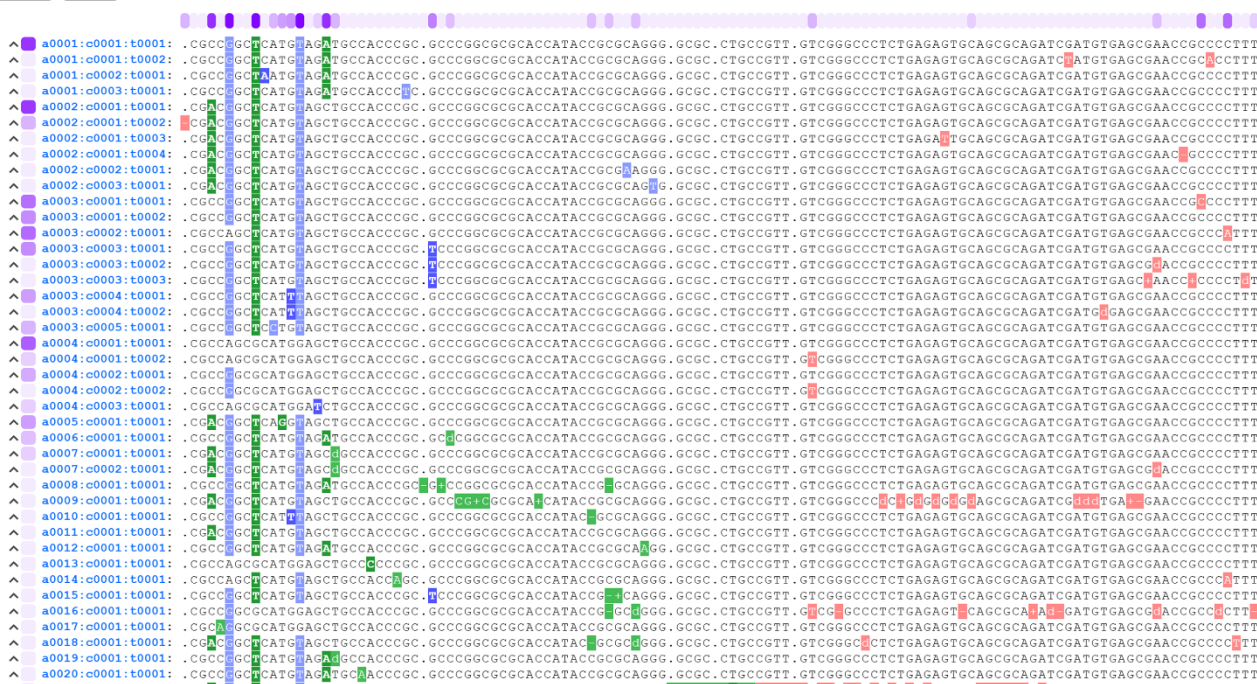
【4】 CoLoRS との国際連携

CoLoRS への参画を申請し、2023 年 7 月に CoLoRS への参画を行った(<https://colordsdb.org/>)。CoLoRS に提供を行ったデータについては取り込まれ、令和 6 年度内に CoLoRS が目標としていた 1,427 検体の SNP および SV の統計情報が利用できるようになった。当初目的の国際連携は達成したが、来年度は、同情報を JoGo ポータル上でどのように利用をするか検討を進めていく。

図 7

FGFR4_chr5_177081915_177103144

Show: ☐ All level ☐ Amino acid level ☐ Coding level ☒ Transcript level ☐ Genebody level



(河合グループ)

[3] RDF 化のためのアノテーション情報の整理と外部 DB との連携 (河合)

[3-1] 外部 DB 連携のための会議参加と定期意見交換 (DBCLS 他)

[3-2] RDF での SV 情報提供

JoGo ポータルで整備を進めるデータベースについて、すでに国内で確立しているデータベースと積極的に連携を継続して進めた。特に、JoGo ポータルサイトで提供を行う配列情報の可視化に対応する URL についてどのように生成すればよいか、また逆にどのようなニーズがあるかどうかについて意見交換を行った。例えば、外部サイト A で提供している変異周辺の配列情報を JoGo ポータル側に外部連携で表示してもらったり、JoGo ポータルでの各ソフトウェアから得られた変異情報やフィルタリング条件での結果などの情報を登録してもらい必要がある。特に【1-1】「JoGo ゲノム DB ポータル構築と提供」と【2-1】「JoGo オミクス DB ポータル構築と提供」で得られる情報は、個別の解析ソフトウェアから得られる情報について整理をして提供することから、得られた結果についてどのレベルまで外部データベースと連携するかについては NBDC / DBCLS などの有識者との意見交換を進めた。その結果、DBCLS が開発を進めている Stanza に JoGo で整備を進めている情報を可視化するためのコンポーネントの開発が行われた(図7:EGFR4 の領域を JoGo の Stanza を用いて JoGo ポータル上で可視化した結果)

本年度は、関連して、バリエーション標準化研究会 (VISC)が主体となって変異情報の表現方法を進めているが、その中で、DBCLS の片山博士ら DBCLS 有識者と継続して意見交換会をする中で、現在の方針としては RDF 化にはこだわらずに、JoGo で構築するハプロタイプ情報を Stanza などのコンポーネントで可視化して表現できるようにすることで検討が進められることとなった。

他に、テロメアからテロメアレベルの日本人の配列情報を試験的に構築することで得られた難読領域の配列情報を JoGo 側にフィードバックすることで今後より複雑な領域をどのように表現するか継続して検討を進めている。特に、沖縄で 2025 年 1 月に開催された **Graph Summit** にも参加(長崎、河合が参加)し、複数の遺伝子をまたぐ構造を表現するためのグラフゲノムによるソフトウェアを試験利用するなどしてより適切な表現形式の検討を進めた。

§4. 成果発表等

(1) 原著論文発表

① 論文数概要

種別	国内外	件数
発行済論文	国内(和文)	0 件
	国際(欧文)	4 件
未発行論文 (accepted, in press 等)	国内(和文)	0 件
	国際(欧文)	0 件

② 論文詳細情報

1. Seik-Soon Khor, Kouyuki Hirayasu, Yosuke Kawai, Hie Lim Kim, Masao Nagasaki, Katsushi Tokunaga. LILR genotype imputation with attribute bagging (LIBAG): leukocyte immunoglobulin-like receptor copy number imputation system.. *Frontiers in immunology* (2025) 16: 1559301-1559301 (10.3389/fimmu.2025.1559301)
2. Masao Nagasaki, Kouyuki Hirayasu, Seik-Soon Khor, Ryoko Otokozawa, Yayoi Sekiya, Yosuke Kawai, Katsushi Tokunaga. JoGo-LILR caller: Unveiling and navigating the complex diversity of LILRB3-LILRA6 copy number haplotype structures with whole-genome sequencing.. *Human immunology* (2025) 86: 3 111272-111272 (10.1016/j.humimm.2025.111272)
3. Kouyuki Hirayasu, Seik-Soon Khor, Yosuke Kawai, Mihoko Shimada, Yosuke Omae, Gen Hasegawa, Yuko Hashikawa, Hiromu Tanimoto, Jun Ohashi, Kazuyoshi Hosomichi, Atsushi Tajima, Hiroyuki Nakamura, Minoru Nakamura, Katsushi Tokunaga, Rikinari Hanayama, Masao Nagasaki. Identification of the hybrid gene LILRB5-3 by long-read sequencing and implication of its novel signalling function.. *Frontiers in Immunology* (2024) 15: (10.3389/fimmu.2024.1398935)
4. Takeo Naito, Ryuya Osaka, Yoichi Kakuta, Yosuke Kawai, Seik-Soon Khor, Junji Umeno, Katsushi Tokunaga, NCBN Controls WGS Consortium, Hantsue Ishibashi-Ueda, Tsutomu Tomita, Michio Noguchi, Ayako Takahashi, Yu-ichi Goto, Sumiko Yoshida, Kotaro Hattori, Ryo Matsumura, Aritoshi Iida, Yutaka Maruoka, Hiroyuki Gatanaga, Masaya Sugiyama, Satoshi Suzuki, Kengo Miyo, Yoichi Matsubara, Akihiro Umezawa, Kenichi Hata, Tadashi Kaname, Kouichi Ozaki, Haruhiko Tokuda, Hiroshi Watanabe, Shumpei Niida, Eisei Noiri, Koji Kitajima, Yosuke Omae, Reiko Miyahara, Hideyuki Shimanuki, Yosuke Kawai, Katsushi Tokunaga, Hiroshi Nagai, Yusuke Shimoyama, Rintaro Moroi, Hisashi Shiga, Masao Nagasaki, Yoshitaka Kinouchi, Atsushi Masamune. Genetically Predicted Higher Levels of Caffeic Acid Are Protective Against Ulcerative Colitis: A Comprehensive Metabolome Analysis. *Inflammatory Bowel Diseases* (2024) : (10.1093/ibd/izae143)

(2) その他の著作物(総説、書籍など)

なし。

(3) 国際学会および国内学会発表

① 概要

種別	国内外	件数
招待講演	国内	1 件
	国際	3 件
口頭発表	国内	1 件
	国際	1 件
ポスター発表	国内	1 件
	国際	1 件

② 招待講演

〈国内〉

1. 長崎正朗、オミクスサイエンスセンター大規模情報解析統合システム開発、先駆的科学計算に関するフォーラム 2024、福岡、2024/5/17

〈国際〉

1. Masao Nagasaki, JoGo v1: Joint Open Genome and Omics Platform v1, 2nd International Meeting on Cancer•Aging Biology and Bioinformatics, Osaka, 2025/4/16
2. Masao Nagasaki, Jogo, Graph Summit 2025, Okinawa, 2025/3/11
3. Masao Nagasaki, Jogo, The Barbados meeting on Pangenome graphs: tools and applications, Bellairs, Barbados, 2025/1/6

③ 口頭講演

〈国内〉

1. 長崎正朗、ハイブリッドクラウドを用いたゲノム情報に基づく構造多型パネルの構築とアノテーション、学際大規模情報基盤共同利用・共同研究拠点 第 16 回シンポジウム、オンライン(東京)、2024/7/11

〈国際〉

1. Masao Nagasaki, High-Resolution Genomics and Omics Integrative Analyses of NCVC Biobank Dataset Towards Discovery of Cardiovascular Disease-Related Factors, Yonsei University and NCVC Collaborative Conference for Cardiomyopathy Genomics, Osaka, 2024/6/21

④ ポスター発表

〈国内〉

1. 長崎正朗、河合洋介、JoGo: Japanese Open Genome and Omics Platform 1.0 に向けた取り組みについて、トーゴの日シンポジウム 2024、東京、2024/10/5

〈国際〉

1. Masao Nagasaki, Kouyuki Hirayasu, Seik Soon Khor, Ryoko Otokozawa, Yayoi Sekya, Yosuke Kawai, Katsushi Tokunaga, JoGo-LILR Caller: Unveiling and Navigating the Diversity of LILRB3-LILRA6 Copy Number Structures with Whole-Genome Sequencing,

(4) 知的財産権の出願（国内の出願件数のみ公開）

① 出願件数

種別		件数
特許出願	国内	0 件

② 一覧

1) 国内出願

なし。

(5) 受賞・報道等

① 受賞

なし。

② メディア報道

なし。

③ その他の成果発表

なし。

§5. 主要なデータベースの利活用状況

1. アクセス数

① 実績

表 1 研究開発対象の主要なデータベースの利用状況

名称	種別	2024 年度(月間平均値)
JoGo (beta release: restricted access)	訪問者数	41
	訪問数	137
	ページ数	9,576

※ β 公開後(利用者限定)2024 年 9 月以降

② 分析

β 版であり、公開が限定されているためにまだアクセスは少ない。来年度が最終年度であるため論文が採択され次第速やかに公開を進めていく。状況によっては preprint とデータベース公開にする可能性がある。

2. データベースの利用状況を示すアクセス数以外の指標

JoGo の β 版に含まれる一部の情報を当開発メンバが活用することで採択論文 3 件、来年度投稿予定の論文が1件成果として得られている。

(論文)

1. 1.9K views / 716 Download / 1 Citation

K. Hirayasu, S.S. Khor, Y. Kawai, M. Shimada, Y. Omae, G. Hasegawa, Y. Hashikawa, H. Tanimoto, J. Ohashi, K. Hosomichi, A. Tajima, H. Nakamura, M. Nakamura, K. Tokunaga, R. Hanayama, M. Nagasaki. Identification of the hybrid gene LILRB5-3 by long-read sequencing and implication of its novel signaling function, *Front Immunol.* 15:1398935, 2024.

2. 912 views / 315 Downloads

S.S. Khor, K. Hirayasu, Y. Kawai, H.L. Kim, M. Nagasaki, K. Tokunaga. LILR genotype imputation with attribute bagging (LIBAG): leukocyte immunoglobulin-like receptor copy number imputation system, *Front Immunol.* 16, 2025.

3. 1 Citation

M. Nagasaki, K. Hirayasu, S.S. Khor, R. Otokozawa, Y. Sekiya, Y. Kawai, K. Tokunaga. JoGo-LILR caller: Unveiling and navigating the complex diversity of LILRB3-LILRA6 copy number haplotype structures with whole-genome sequencing, *Human Immunology.* 86(3):111272, 2025.

(Preprint)

4. Kulmanov M, Ashouri S, Liu Y, Abdelhakim M, Alsolme E, Nagasaki M, Ohkawa Y, Suzuki Y, Tawfiq R, Tokunaga K, Katayama T, Abedalthagafi M, H Robert, Kawai Y, Phased genome assemblies and pangenome graphs of human populations of Japan and Saudi Arabia, bioRxiv, <https://doi.org/10.1101/2024.12.18.628902>.

また、同ウェブサイトは教育コンテンツとしての役割も有する。2024 年度は、JoGo の β 版のデータベースを用いた授業を、九州大学の生命医科学 I (40 名参加)、生命医科学研究入門 (111 名参加)で行い、学生にヒトゲノムのオープンデータベース、また、そのデータベースでどのようなことができるか等の授業を行うことで、将来の利用者となる可能性が高いシーズへの教育活動に役立てることができた。

3. データベースの利活用により得られた研究成果(生命科学研究への波及効果)

1. 金沢大学の平安准教授らと、JoGo データベースを対象に解析することで、いままで報告がされていない日本人の一般集団に存在する LILRB5 と LILRB3 の新規融合遺伝子を同定した。同融合遺伝子を平安が機能解析をすることで実際にタンパク質レベルで発現をして生体内で機能していることが判明した。今後、同融合遺伝子をもつ集団の形質や疾患との理解を深めることが可能となった

関連:K. Hirayasu, S.S. Khor, Y. Kawai, M. Shimada, Y. Omae, G. Hasegawa, Y. Hashikawa, H. Tanimoto, J. Ohashi, K. Hosomichi, A. Tajima, H. Nakamura, M. Nakamura, K. Tokunaga, R. Hanayama, M. Nagasaki. Identification of the hybrid gene LILRB5-3 by long-read sequencing and implication of its novel signaling function, *Front Immunol.* 15:1398935, 2024.

2. JoGo データベースの遺伝子のうちコピー数構造が複雑なことが知られている LILRB3 と LILRA6 に着目することで、長鎖型シーケンスの情報を鋳型とすることで、短鎖型シーケンスから LILRB3 と LILRA6 の複雑なコピー数構造を判定できるソフトウェア JoGo-LILR Caller を開発した。このソフトウェアの開発により、今後は、このソフトウェアを用いて短鎖型シーケンスから得られた情報を再解析することで、LILRB3 と LILRA6 の正確なコピー数構造を分離して解析できるようになる。この成果により、LILRB3 および LILRA6 のコピー数と形質や疾患との解釈が進むと考えられる。

関連:M. Nagasaki, K. Hirayasu, S.S. Khor, R. Otokozawa, Y. Sekiya, Y. Kawai, K. Tokunaga. JoGo-LILR caller: Unveiling and navigating the complex diversity of LILRB3-LILRA6 copy number haplotype structures with whole-genome sequencing, *Human Immunology.* 86(3):111272, 2025.

3. 2 で開発を行った JoGo-LILR Caller を活用することで、さらに SNP アレイの情報から、LILRB3 と LILRA6 のコピー数を正確に推定することができるソフトウェア LIBAG の開発を Nanyang Technological University の Khor 博士と行った。今後は、このソフトウェアを用いて SNP Array から得られた情報を再解析することで、LILRB3 と LILRA6 の正確なコピー数構造を分離して解析できるようになる。この成果により、LILRB3 および LILRA6 のコピー数と形質や疾患との解釈が進むと考えられる。

関連:S.S. Khor, K. Hirayasu, Y. Kawai, H.L. Kim, M. Nagasaki, K. Tokunaga. LILR genotype imputation with attribute bagging (LIBAG): leukocyte immunoglobulin-like receptor copy number imputation system, *Front Immunol.* 16, 2025.

4. データベースの利活用によりもたらされた産業への波及効果や科学技術のイノベーション(産業や科学技術への波及効果)

1. 公開対象はなし

§6. 研究開発期間中に主催した活動(ワークショップ等)

(1) 進捗ミーティング

年月日	名称	場所	参加人数	目的・概要
2024 年 4 月 22、23 日	進捗に関する研究打合せ	九州大学	2 人	進捗に関する研究打合せを行った。
2024 年 6 月 17、18 日	進捗に関する研究打合せ	九州大学	2 人	同上
2024 年 9 月 12、13 日	進捗に関する研究打合せ	九州大学	2 人	同上
2024 年 11 月 25、26 日	進捗に関する研究打合せ	九州大学	2 人	同上
2025 年 2 月 17、18 日	進捗に関する研究打合せ	九州大学	2 人	同上

(2) 主催したワークショップ、シンポジウム、アウトリーチ活動等

年月日	名称	場所	参加人数	目的・概要
2024 年 4 月 22、23 日	バリエーション情報標準化研究会 VISC (Variant Information Standardization Collegium)4.4	九州大学	14 人	DBCLS およびゲノムサイエンスを中心とした有識者が参加し、Joint Open Genome and Omics (JoGo)の開発状況と進捗状況の共有、開発課題の洗い出し、活発な意見交換を現地で密に行った。
2024 年 6 月 17、18 日	バリエーション情報標準化研究会 VISC (Variant Information Standardization Collegium)4.6	九州大学	10 人	DBCLS およびゲノムサイエンスを中心とした有識者が参加し、Joint Open Genome and Omics (JoGo)の開発状況と進捗状況の共有、開発課題の洗い出し、活発な意見交換を現地で密に行った。
2024 年 9 月 12、13 日	バリエーション情報標準化研究会 VISC (Variant Information Standardization Collegium)4.9	九州大学	12 人	DBCLS およびゲノムサイエンスを中心とした有識者が参加し、Joint Open Genome and Omics (JoGo)の開発状況と進捗状況の共有、開発課題の洗い出し、活発な意見交換を現地で密に行った。

年月日	名称	場所	参加人数	目的・概要
2024 年 11 月 25、26 日	バリエント情報標準化研究会 VISC (Variant Information Standardization Collegium)4.11	九州大学	9 人	DBCLS およびゲノムサイエンスを中心とした有識者が参加し、Joint Open Genome and Omics (JoGo)の開発状況と進捗状況の共有、開発課題の洗い出し、活発な意見交換を現地で密に行った。
2025 年 2 月 17、18 日	バリエント情報標準化研究会 VISC (Variant Information Standardization Collegium)5.2	九州大学	12 人	DBCLS およびゲノムサイエンスを中心とした有識者が参加し、Joint Open Genome and Omics (JoGo)の開発状況と進捗状況の共有、開発課題の洗い出し、活発な意見交換を現地で密に行った。
2024 年 5 月 28 日、6 月 4 日	生命医科学 I	九州大学	40 人	「生命医科学 I」の講義の中で、JoGo データベースの開発とその目的について説明を行った。
2025 年 1 月 23 日	生命医科学研究入門	九州大学	111 人	「大規模計算機とヒトゲノム解析の最前線」というタイトルで JoGo を使ったヒトゲノムの理解に関係した講義を行った。

以上

別紙 1 既公開のデータベース・ウェブツール等

No.	正式名称	別称・略称	概要	URL	公開日	状態	分類	関連論文
1	JoGo-LILR-Caller	JoGo-LILR-Caller	LILR領域の構造多型が含まれるLILRB3-A6の領域のハプロタイプ構造を短鎖型シーケンサから高精度に同定することができるソフトウェア	https://nagasakilab.csml.org/en/jogo-lilr_caller_v1	2025/2/28	新規	ツール等	同ソフトウェアを利用した論文 K. Hirayasu, S.S. Khor, Y. Kawai, M. Shimada, Y. Omae, G. Hasegawa, Y. Hashikawa, H. Tanimoto, J. Ohashi, K. Hosomichi, A. Tajima, H. Nakamura, M. Nakamura, K. Tokunaga, R. Hanayama, M. Nagasaki. Identification of the hybrid gene LILRB5-3 by long-read sequencing and implication of its novel signaling function, Front Immunol. 15:1398935, 2024. ソフトウェア論文 M. Nagasaki, K. Hirayasu, S.S. Khor, R. Otokozaawa, Y. Sekiya, Y. Kawai, K. Tokunaga. JoGo-LILR caller: Unveiling and navigating the complex diversity of LILRB3-LILRA6 copy number haplotype structures with whole-genome sequencing, Human Immunology. 86(3):111272, 2025.