

ライフサイエンスデータベース統合推進事業  
(統合化推進プログラム 2017年採択課題)

## 研究開発終了報告書

「疾患ヒトゲノム変異の生物学的機能注釈を目指した多階層  
オーミクスデータの統合」

菅野 純夫 | 東京医科歯科大学 難治疾患研究所  
非常勤講師

研究開発期間： 2017年4月～2022年3月



©2022 菅野 純夫 (東京医科歯科大学) Licensed under CC BY 4.0

## §1. 研究開発実施の概要

本データベースは、ヒトゲノム多型・変異に生物学的機能注釈を与えるべく、ゲノム変異位置の近傍の多層オーミクスデータ(エピゲノム:ヒストン修飾、DNA メチル化パターン;トランスクリプトーム:発現量、スプライスパターン)をあたり限り広範囲で取得し、日本人ゲノム情報への統合を試みたものである。

これまでにヒトゲノム多型情報については、ヒトゲノムバリエーションデータベース(HGVDB)と、実体として統合した。統合対象となった HGVDB データには種々の疾患ゲノムの GWAS 解析結果および HLA 領域関連データが含まれている。特に HLA データは昨今、がん免疫の分野においてその整備が喫緊の課題となっているもののひとつである。また、東北メガバンク、BioBank Japan とも連携を図り、ゲノム情報についても我が国で産生されたデータの横断的な 2 次利用を促進すべくデータの統合を試みた。また健常者多層オーミクスデータについては、CREST-IHEC データの収載として完了した。

一方で当データベースでは、近年進展の著しいゲノム解析関連技術について先行的にデータを公開する枠組みの整備を整えた。具体的には、申請者らは個別研究あるいは製薬企業との共同研究において、ヒトがん細胞薬剤応答オーミクス摂動大規模データ、および各種プラットフォーム(フリーダタイム C1、10X Genomics Chromium、独自に開発した Nx1 システムによる)を用いたシングルセルデータを集積した。これらについて追加解析費用等を活用して、製薬会社等が個別に検索、利用できるデータベース検索プラットフォームとして整備した。またナノポアシーケンサーを用いた長鎖 DNA 解析データ、空間トランスクリプトームデータについても同様の枠組みの整備を行った。最終年度には、新たなる追加解析経費を得て、シングルセル解析を健常者末梢血について行ったデータについての公開を行った。また新学術領域「先進ゲノム支援」から産出されるマウス等モデル生物種データを加え、上記のブラウザで統合表示することを可能とした。全体として、がんを主標的として、臨床検体では直接の集積が困難な多層オーミクス情報を充実させ、さらには国内研究で産生されたデータを統合することでモデル系におけるオーミクス情報の整備を試行したのとなった。

## §2. 研究開発実施体制

### 1. 研究グループ

#### (1) 菅野グループ(医科歯科) (研究代表者グループ)

- 担当項目

データベースについて研究の統括を行う。各グループ間での進捗の調整を行い、連携を促進する。また各グループで共通して必要となる計算機環境を整備し、その環境維持に努める。

#### (2) 菅野グループ(東大) (主たる共同研究者グループ(1))

- 担当項目

各種のデータセットの収集と加工を行い、データベースとして公開する主体を担う。先進ゲノム支援、IHEC データについては、すでに内部的に格納されているものを用いる。dbSNP、ExAC 等の外部参照データについては、前期開発期間内に開発したツール群を用いて、定期的、半自動的作業でデータをできる限り、最新の情報に保つ。また、シングルセル、薬剤摂動オーミクスデータについては、関連解析ツール、ブラウザを開発する。前開発期間におこなった開発を発展的に継続する形で DBCLS の担当者と連絡を密にして、多層オーミクスデータの RDF 化スキーマの試用を進める。特に、実験条件、試料臨床情報(非公開)に関するメタデータオントロジーの整備と RDF 化は集中的に行う。

#### (3) 土原グループ(主たる共同研究者グループ(2))

- 担当項目

培養がん細胞多層オーミクスデータ、薬剤摂動データ、シングルセルデータの収載を行う。またがん臨床データについての整備と秘匿検索についての技術開発を模索する。

#### (4) 徳永グループ(主たる共同研究者グループ(3))

- 担当項目

統合データベースプロジェクトおよび統合化推進プログラムにおいて構築してきた「ヒトゲノムバリエーションデータベース(HGVDB)」(日本人の健常者 SNP-DB、GWAS-DB、健常者 CNV-DB、疾患関連 CNV-DB、HLA-DB、統合化変異 DB を含む)を継承、移管し、登録データを一層充実させるとともに、より活用しやすい DB を目指して機能改善する。

#### (5) 井ノ上グループ(主たる共同研究者グループ(1))

- 担当項目

統合データベースプロジェクトおよび統合化推進プログラム「ヒトゲノムバリエーションデータベース(HGVDB)」、特に国立遺伝学研究所で 1 次データ産生が行われてきたその HLA 多型部分について、徳永グループと共同して DBKERO へとデータを継承、移管し、登録データを一層充実させるとともに、より活用しやすい DB を目指して機能改善する。

## 2. 有識者会議等

### (1) 会議概要

名称	DBKERO 産学利用促進委員会
目的	データコンテンツの充実と閲覧手法の改善をし、DBKERO の産学の利用を促進する。
委員数	7人

### (2) 開催歴

年月日	場所	参加人数	主な議題
2019年1月24日	東京大学医科学研究所・総合研究棟 8F 会議室	9名	データコンテンツの充実と他のデータベースとの連携について
2019年9月25日	東京大学大学院新領域創成科学研究科情報生命科学実験棟 2F 会議室	9名	DBKERO を利用した創薬シーズ探索の可能性について

### §3. 研究開発の目的、実施内容及び成果

#### 1. 研究開発の背景

ヒトゲノム情報、多層オーミクスデータの整備と有効活用化は、国内外で喫緊の課題である。イルミナ HiSeq を中心として、近年の次世代シーケンス関連技術の進展は著しく、ヒトゲノム配列あるいはがんゲノム配列を大量に決定し疾患関連多型・変異を同定、診断・治療に応用しようという試みは数多い。しかしいかに設計図であるゲノム配列とその変異を解析しても、その生物学的な意義を直接読み解くことは依然として困難である。本データベースにはゲノム、エピゲノム、トランスクリプトーム情報が整理統合されている。特にがんゲノムについては、肺腺がん、食道がん、大腸がんゲノム情報に、肺線がん定常状態培養細胞トランスクリプトーム、エピゲノム情報さらにゲノム支援で収集されたマウス等のモデル生物種のデータセットが統合されている。また健常者多層オーミクスデータとして IHEC により集積されたエピゲノム情報が、また健常者末梢血シングルセルデータが統合されている。さらにがんを主標的としてロングリード解析データ、空間トランスクリプトーム解析データ等最新のゲノム計測技術を駆使して得られた先端オーミクス解析データを付記した。これにより、国内研究で見出されつつある種々の日本人疾患ゲノム情報に生物学的機能解析を与える際に、実験モデル系より得られた有用な情報基盤を整備することが可能となったと考えている。

#### 2. 研究開発対象のデータベース・ツール

##### (1) データベース

###### ・主要なもの

正式名称	略称	概要
DBKERO: 疾患ゲノム多層オーミクス機能注釈データベース	DBKERO	本データベースではヒトゲノム多型・変異に生物学的機能注釈を与えるべく、ゲノム変異位置、近傍のエピゲノム(ヒストン修飾、DNA メチル化パターン)、トランスクリプトーム情報(発現量、スプライスパターン)をヒトゲノム情報に統合する。

###### ・上記以外のもの

正式名称	略称	概要
該当なし		

##### (2) ツール等

正式名称	略称	概要
BEACON SCAN	なし	ゲノムデータの統合検索システム
DBKERO RDF viewer	なし	RDF 検索による統合検索を受入れ、検索結果を表示する枠組み
Transcription Factor Binding Site Search	なし	ユーザーが指定するゲノム上のターゲット領域から、転写因子の結合領域を、ChIP-Seq の結果や配列パターンから予測するプログラム

### 3. 達成目標及び実施計画

#### (1) 当初の実実施計画・達成目標

##### 初年度

- ・ IHEC-CREST において、現在までに日本チームが集積、JGA からその配列情報が公開されているエピゲノム情報について、そのシグナル強度情報を全件収載、即時、公開する。その実務の大半を初年度に完了する。
- ・ HGVDDB については、現在までに徳永グループから公開されているデータの全データを本データベースに移管する。これには遺伝研担当分である HLA 領域データ、現在格納されている希少疾患変異データセットを含む。初年度は HLA 領域データを中心にその 2/3 の移設を行う。これを検討材料に DBKERO との接続部のパイプライン、ブラウザ等についての関連プログラム群の整備を行う。

##### ～第3年次末

- ・ IHEC-CREST データについて、その公開データベースの動作確認および種々の最終確認を行う。この時点での Update を行うことで CREST-IHEC の最終成果とする。また IHEC の他国産出関連データについては IHEC データポータルについてリンクを整備する。データの具体的な統合についてもその実務を開始する。
- ・ HGVDDB について、残り 2/3 のデータを移設し、全データについて最終確認を行う。データベースを公開して最終動作確認を行う。
- ・ 東北メガバンクと、データ公開条件等についての付帯条件を整理し、本データベースに収載する。

##### ～第5年次末

- ・ 来年度公開の予定されている Bio Bank Japan GWAS データおよび 5 年次末までには公開が見込まれる AMED 臨床ゲノムデータについては、JGA より一般公開になった部分の頻度情報について、本データベースでの 2 次利用が可能か、その利用条件について協議を行う。
- ・ 先進ゲノム支援において、その研究期間終了時に論文発表が完了し、配列データが JGA・SRA より一般公開になった、マウス等のモデル生物種多層オーミクスデータについて、データ生産者と協議して、本データベースからの公開を行う。過去の実績に鑑み、多くの支援案件について、公開が可能となると考えている。
- ・ がん細胞における薬剤応答多層オーミクス摂動データ、シングルセルデータについて、論文発表が完了したものから公開を行う。

#### (2) 期間中に追加・削除・変更した実施計画・達成目標

##### 学会ブース出展(2017 年追加実施)

臨床系、ユーザー系の学会での発表を行うことでユーザーを発掘する取り組みを強化した。日本癌学会、日本人類遺伝学会等の各学会において本データベースに関する一連の発表を行い、実務担当者が同行してデータベースのさらに詳細な説明、データベースを活用した共同研究について提案を行った。

##### シングルセルデータ公開インターフェースの開発(2021 年度追加実施)

菅野グループ(東大)では、今回、健常者の日常生活においてゲノム配列がどのように機能しているかを解析するために、柏の葉在住在勤の健常者について末梢血シングルセル解析を実施した。日常生活でのイベントが発生するたびに適宜、採血を行い解析に用いている。生活イベントは、インフルエンザ、コロナウィルスのワクチン接種を含む。今回の追加実施で

このようなデータセットのアノテーションパイプラインを開発し、その可視化を行うことで、現在の DBKERO では表現できなかった個体差や日動差といった多様性を明示的に表現し、これを直感的に理解するための枠組みを開発しており今年度末に全面公開する予定である。

#### 4. 実施内容

##### (1) 実施内容

上記に掲げた研究開発予定項目については概ね当初の目標を達成できたと考えている。全体として、現在、DBKERO のデータベースでは 5,510 件の多層オーミクスデータセットを公開している。これらのデータは全てゲノムブラウザ等を通して可視化されている。さらにこのうち倫理規定がクリアされているもの、あるいはその対象外であって、二次配布が可能であるデータ(多くは培養細胞のデータセットからなる)はダウンロードポータル(Data Portal)に掲載を完了した。具体的には Single cell, IHEC, GWAS, Drug Perturbation, Cancer Cell Line, Long read に関連するデータの、2,573 件(47%)である。残りのデータのうち ICGC, TCGA, ENCODE 等外部のデータベースから得たデータ 2,723 件(49%)については Data Portal には公開できないが、しかるべくリンクサイトを明示した。またヒト正常組織細胞 TSS-Seq(倫理規定の対象外であるもの)203 件(4%)については引き続き公開の準備を進めており、第 5 年次終了時までには完了予定である。また、新規公開データセットの概要をまとめた論文を作製、発表した(Suzuki et al. 2018 NAR DB issue)。また関連学会において、データベースを周知、データ活用例の紹介を行った(下記)。

個々の開発項目の成果については下記に図示する(図 1)。

##### • IHEC-CREST データの収載

下図に示すようにエピゲノム関連データ閲覧システムを開発した。収載された IHEC-CREST データは表に示す。

チーム	器官	検体数	データセット数			
			BS-seq	RNA-seq	ChIP-seq	合計
金井班	肝臓	8	8	8	48	64
金井班	腸管	12	12	12	72	96
金井班	胃	16	14	16	33	63
白髭班	血管内皮	54	0	50	188	238
佐々木班	子宮内膜	42	42	30	108	180
合計		132	76	116	449	641

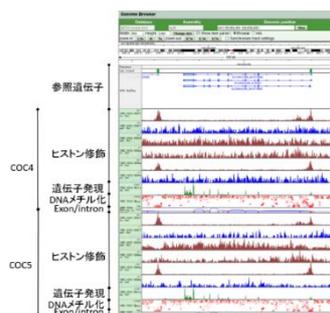


図 1 IHEC ブラウザ(左表はデータの一覧、右図は表示例)

##### • HGV-DB の統合

HGVDB については、これまでに徳永グループから公開されていたデータで移管可能な全データを本データベースに移管した(図 2)。

疾患名	検体数	
	Case 数	Control 数
スティーブensonンズ症候群	117	691
アルツハイマー(study 1)	397	964
アルツハイマー(study 2)	752	752
気管支喘息	565	940
脳動脈瘤	288	194
子宮内腫瘍	696	825
胃がん	752	752
C型肝炎(study 1)	90	190
C型肝炎(study 2)	99	187
C型肝炎(study 3)	85	76
高血圧	752	752
ナルコレプシー	222	389
パニック障害	170	167
2型糖尿病	752	752
情動脱力発作を伴うナルコレプシー患者	409	1562
HLA-DQB1*06:02 陰性真性遺眠症	125	562
日本人原発性胆管性肝硬変	487	476
B型肝炎ウイルスキャリア、および、健康対照者	181	184
B型肝炎ウイルスキャリア、および、B型肝炎ウイルス排除群	181	185
B型肝炎ウイルス排除群、および、健康対照者	185	184
HLA 配列データ 1	7	
HLA 配列データ 2	10	
HLA 配列データ 3	1	

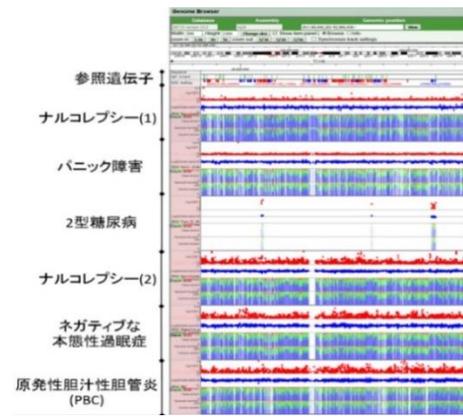


図 2 HGVDB のブラウザ(左表は搭載データの一覧、右図は表示例)

● 関連データベースとの統合

統合 DB human variant DB と相互リンクを形成した。また統合 DB プロジェクトの別班で開発/運用する ChIP Atlas とのデータリンクを実装した。さらに東北メガバンクと、データ公開条件等についての付帯条件を整理し、より包括的な連携の枠組みを模索している。

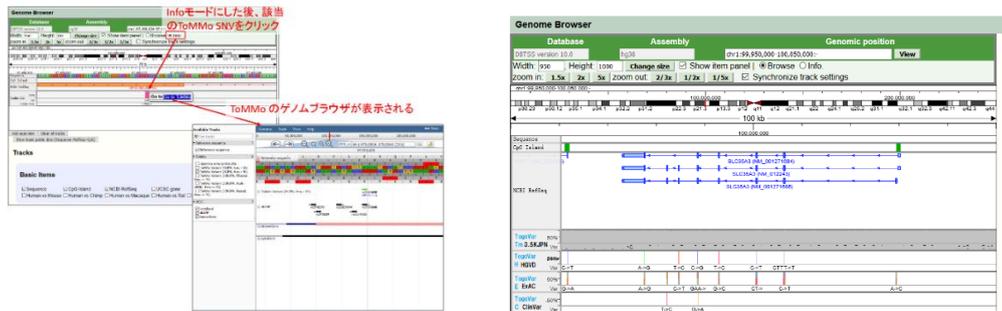


図 3 左図 ToMMo、右図 NBDC human variant DB との相互データ参照例  
※ToMMo の公開許諾が得られないため現在は内部閲覧のみ

● 「先進ゲノム支援」データの取載

新学術領域「先進ゲノム支援」においてマウス等モデル生物において、ヒトの病態関連解析、生物学的特性解析を指向したものを中心に、ヒトデータと関連付けて公開した。

● データポータル整備

ゲノム支援関連データを中心に多様であるが雑多であるデータセットについて視覚的に検索、データ入手が可能ならデータポータルの整備を行った(図 4)。

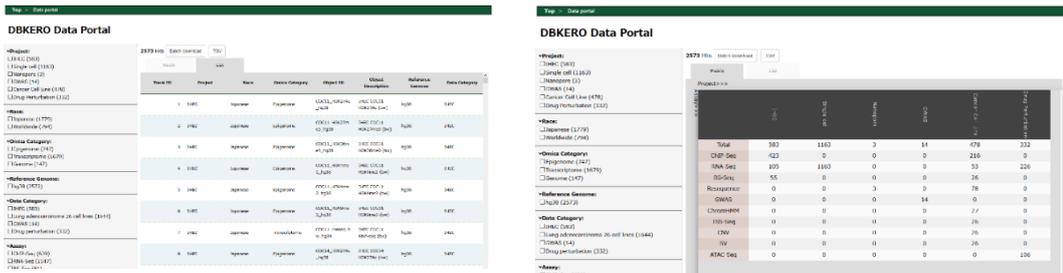


図 4 左図データポータル一覧表示、右図カテゴリーによるデータの検索

• **RDF 化の推進**

RDF 化について、トランスクリプトーム、エピゲノムの部分についてのデータ加工を完了した。変更された RDF データベースは NBDC RDF データベースポータルへ登録した。ゲノムデータについてはフォーマット等について引き続き NBDC 側担当者と協議するものとし、さらに先進ゲノム支援関連データを模擬データに、メタデータについての RDF 化の枠組みを構築している。

( <https://integbio.jp/rdf/dataset/kero> )

• **先進計測データに関するデータ配布**

ロングリードデータ、シングルセルデータ、空間トランスクリプトーム解析データについて、研究代表者ら自身が産出したデータを模擬データセットに、ブラウザの開発、データの公開を行った(図 3-5)。

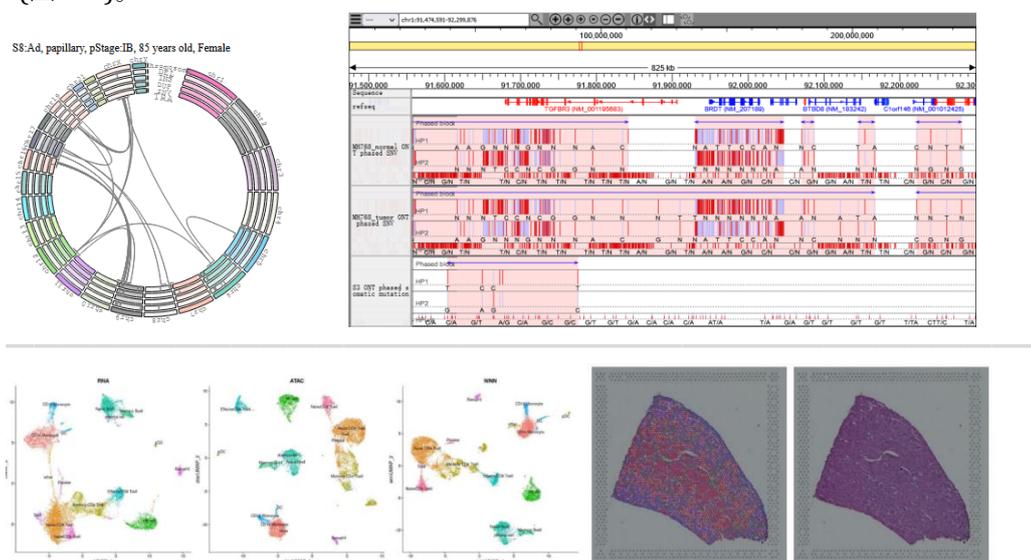


図 5 ロングリードブラウザ(SV 及びフェージングブラウザ(上図)、シングルセルデータの UM AP、空間トランスクリプトーム解析(下図))

• **追加実施**

5 年間の開発期間中に、以下の 3 項目について追加実施を行った。

1) **大規模薬剤刺激オーミクスデータ**

がん細胞における薬剤反応性についての多層オーミクス摂動を記載するために、新規のブラウザを開発し、その公開を行った。これは肺腺がん細胞株パネルに対する96化合物刺激条件で取得された RNA Seq/ATAC Seq データセット合計 1 万セットよりなる。また製薬会社等の民間企業が、機密性をもってデータベース検索および個別登録を行えるシステムを開発し、運用を開始した(表 1, 図 6)。

表 1 大規模薬剤摂動オーミクスデータセット

データセット	
95 化合物	肺腺癌 23 セルライン (各 RNA-seq, ATAC-seq)
23 化合物	肺腺癌 5 セルライン×4 濃度×3 タイムコース (各 RNA-seq, ATAC-seq)



- 講演(60分):がん遺伝子等、テーマを決めて、ヒトに関するデータベースを広く検索した。
- 演習(60分):その中でマルチオーミクスデータベースとしての DBKERO の特色を解説し、そのウェブベースでの利用をガイドした。
- 演習(120分):DBKERO データポータルに収載されるデータについて、その加工と利活用を推進する初等プログラミング講習を行った。

人類遺伝学会、分子生物学会で講演し、データベースの紹介を行った。また NBDC の主催する AJACS 講習会でも講演を行った。

ただし、前述のように、第 4 年次には、コロナ禍の状況下、対面での説明会、ハンズオンでの講習会を開催することが容易ではなかった。上記のようにいくつかのオンライン講習会を開催したが、それを補完する意味で、DBKERO 上にデータベース利用を解説する動画を加えた。これは各ブラウザのページごとその利用法を実際に解説したものである。5 分程度の short 動画 10 本程度からなる(図 8 左)。また、特にシングルセル解析については、実際にどのようにデータポータル上のデータを加工、自身の目的に利活用ができるのか、という質問が多数寄せられた。これに対応するために、DBKERO 収載のマウス肺の scRNA データを材料にデータプロセス手法を解説した Tutorial Page を開設した。本ページでは、同じくマウス肺の scATAC データとのオーミクス階層統合手法も解説されている(図 8 右)。これらのウェブコンテンツの充実により、従来、ハンズオン講習会により行うべきであった新規利用者への周知と既存利用者の便宜のさらなる向上に努めた。

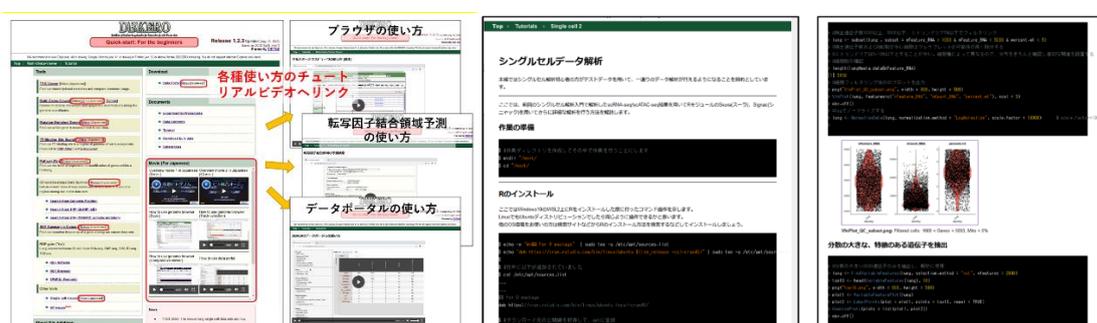


図 8. チュートリアル動画の追加:赤枠内等、各ページでそのページの使い方をガイドした動画を配した(左)。シングルセルデータ解析処理手順の解説ページ。DBKERO 収載データについて、それぞれの加工手順を詳細にチュートリアルした(右)。

2017 年 1 月以降の AWSTATS による Web アクセス統計から、講習会等の実施に従い、最近ではアクセス数、特に訪問者数(新規アクセス者)が増えつつある。ページビュー数については、アクセスするファイル数を減らす工夫が逆にページビュー数を減らす影響を与えることもあり、評価が難しく、月によるばらつきも大きい。概ね上昇傾向にあると思われた(図 9)。今後もコンテンツの充実とともに広報活動をより活発にして、さらなる利用者の拡大に努める。

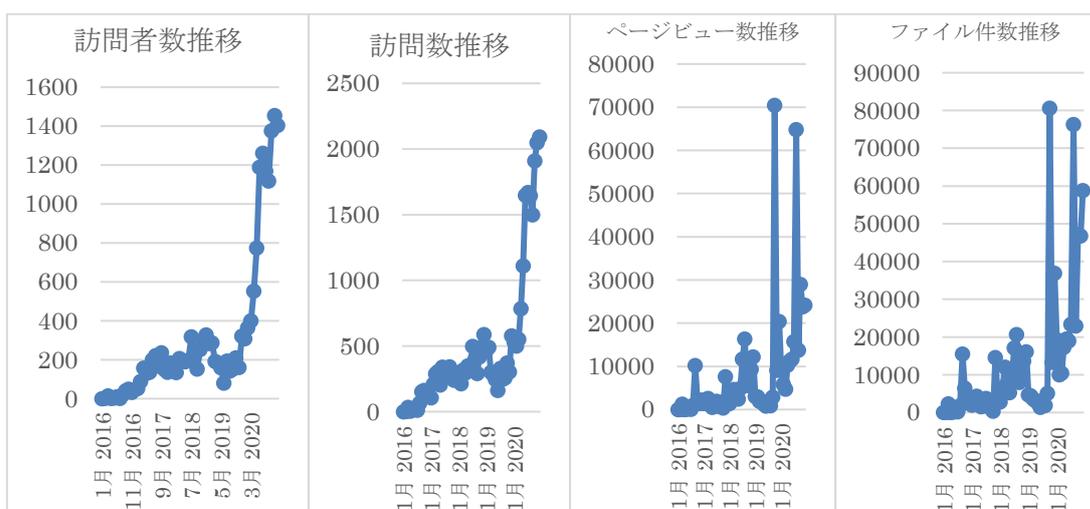


図. 9 訪問者数・訪問数・ページビュー数・ファイル件数の推移

## (2) 実施内容のうちの特定項目の詳細

### ① 研究コミュニティを含むデータ提供者との連携・協業

前述の通り、本データベースの構築は、新学術領域「ゲノム支援」、CREST「日本人標準ゲノム」の各プロジェクトとの密接な連携に基づいて行われた。これらの大型プロジェクトから産生されたデータは、倫理等の制約上、公開形式に差はあるものの、その全てを DBKERO に収載している。また、今年度、CREST「多細胞」とのデータ連携についての協議を開始した。

### ② データベース利用者への周知、利用者との連携・協業

利用者の拡大を目指して、定期的に学会での発表を行い、特に日本分生生物学会ではブースを展示した。また年に数回の利用者講習会を開催している。国内出版社(羊土社、薬研社等)とも連携して、総説を公開している。特に薬剤摂動データについては、製薬企業(協和キリン、富士フイルム、小野薬品工業等)と連携して、そのデータの有効活用に努めた。

### ③ 利用者にとって有用なデータ基盤の構築

今次、NBDC 統合 DB 計画で開発されている「human variation DB」および「ChIP Atlas」との積極的な連携を模索した。特にヒトの多型データに関しては、本計画終了後も発展的な連携を進めて行くべく協議を進めている。ロングリード解析、シングルセル解析、空間トランスクリプトーム解析データについては先進的な可視化手法についての技術開発を行うこともできた。これらの技術は本格化するデータ産生に向けて、重要なデータベース基盤となると考えている。

### ④ 持続的なデータベース運用体制の構築に向けた取り組み

国内外のいずれにおいても手元の PC での解析を可能とするツール群の充実が著しいように思う。実際、シングルセルも含また多様な多層オーミクスデータにしても、可視化さらには個別の目的の解析に対しては、共通のデータベースで実装するよりも、手元での環境構築を行った方が効率が良い。実際、多くのデータベースがデータポータルのみを実装しており、その主眼をデータの品質管理、またデータ形式の標準化に移しているように思う。本 DBKERO においては、データ産生者と一体になった運営をおこなっているため、現在、旧来型のデータベースに固有なこれらの問題は存在しない。ただし逆にそれが汎用的なデータの収集(さらに広い範

困からのデータ取得)の制約になっている。これらの両立は本計画終了後の課題であると考えている。

#### ⑤ 人材の育成

反省点として、データベースを主たる業務/成果とする人材育成を実施することはできなかった。DBKERO では、データの品質管理はデータ産生技術職員が、データ加工/ブラウザの開発は専属プログラマーが、その総合的評価/学会・総説等での発表はそれぞれの分野での研究者が行う、という形をとってきた。換言すると、これらを俯瞰するような研究者については、そのキャリアパスを策定することができなかったために、その人員の確保と育成は行うことができなかった。本分野が内包する本質的な問題点と認識しつつも現在も問題の解決の方策は立っていない。この部分についてはある程度、公的機関の枠組みを利用するか、民間でのポストの確保を優先する必要があるように思われる。

#### ⑥ 国際連携・国際貢献

NAR DB issue (Suzuki A et al. 2018)を含め、いくつかの国際誌において DBKERO の紹介論文を掲載した。また国際学会口頭発表においても周知につとめ、利用者増大に努めた(成果論文の項を参照)。

#### ⑦ その他

私見ではあるが、今後、ヒト関連データの産生は国主導のものだけでなく民間主導の割合がますます増加していくと考えている。これは疾患関連にとどまらず多様な分野を含む。産業利用を当初から想定するこれらのデータについては、少なくともコミュニティ共有、あるいは1対1の双方個性共有が重要になると思われる。そのためには、部分的に情報を切り分けつつ、それでも参照データベース、共通解析ツール等の部分は共有化できるような新しい計算機環境の構築自体が必須である。我々はこのような目的に利用するための計算機環境を試行的に整備、実用化する目的で社団法人「柏の葉オーミクスゲート」を設立した(<https://www.kog.or.jp/>)。民間の資本も導入しつつ多面的なデータベースの構築の枠組みを模索しているものである。

## §4. 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(月間平均)

名称	種別	2017年度	2018年度	2019年度	2020年度	2021年度 (10月時点)
DBKERO	訪問者数	171	243	223	1,230	1,423
	訪問数	283	410	363	1,721	1,942
	ページ数	2,135	6,733	10,823	21,644	13,452

#### (2) 分析

- DBKERO は 2016 年度から DBTSS から分割する形で公開を開始した。2018 年の 1 月には N AR Database Issue に論文が掲載され、ユーザー数が訪問者数、訪問数、ページビュー数とも増加した。
- 2020 年には、DBTSS への来訪者への DBKERO への誘導や、その他ユーザビリティの改善、コンテンツの見直し等ユーザー目線で利用しやすいようにサイトの改良を行った。また、セミナーの開催や学会での宣伝等で普及に努めた結果、大幅にユーザー数が増加した。
- 2021 年度はページ数が減少している。我々のデータベースではゲノムブラウザ等のビューワも Web 上のデータファイルにアクセスするため、これらのアクセスがページビューの多くを占める。サイトの軽量化のため、Webブラウザがサーバー上の DB を json 経由でルックアップする方式を見直し、bam, bigBed 等のファイルに直接アクセスするように改善を行っている。データの切り替え作業を継続的に行っているため、アクセス軽量化の効果でページビュー数が減少したのかもしれない。訪問数、訪問者数は引き続き上昇しており、利用ユーザー数もさらに増加していると言える。こちらは新規ユーザー獲得のための、短いチュートリアルビデオを配置するなど入門者向けのコンテンツを充実させた効果の表れであると思う。今後も、ロングリードや、シングルセル解析等、データをいち早く取り入れることで、引き続きユーザー数が増加に推移するような形に持っていきたい。

### 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

#### • 研究開発期間全体で顕著な成果を表す論文の被引用件数

顕著な成果を表す論文(Suzuki A, et al. Nucleic Acids Res. 2018)の引用件数は 30 回(Google Scholar)である。

#### • データベース講習会の参加人数

DBKERO の利用ニーズは広く大きいと見込まれるが、シングルセルからロングリード、がんや疾患の GWAS データ等、扱うデータの範囲が広いため、潜在的なユーザーを取り込めきれていないと考えている。そのことを踏まえ、データコンテンツの利用に関心がある人に使ってもらうために講習会を行っている。例えば、2020 年度の医科研ヒトゲノム解析センター主催の講習会では定員(30名)を超える参加希望があり大変盛況であった(最終的な参加者は約 100名)。このように利用ニ

ーズは高いと思われ、引き続き講習会等を積極的に行っていく必要がある。

• **データベース利用を契機とした博士課程入学者**

データベースの利用が契機になり、複数の民間p企業より、東京大学に社会人学生(博士課程)が入学し、2021年9月に学位を取得した。

• **DBKERO に掲載された新規データの解析方法・利活用に関する問い合わせ**

まだ全面更改を開始して日が浅いため特に外国第三者研究者から論文成果として発表されている例は限られているが、DBKERO に掲載されているデータは、その規模および質の面から貴重であり、特に解析ツールを開発しているバイオインフォマティクス研究者にモデルデータとして活用されている。実際に今年度に入り、いくつかの実務的な問い合わせも受けている。

多層オーミクスデータの問い合わせ: Assistant Professor, Rutgers - Robert Wood Johnson Medical School, USA 等

scRNA-seq/scATAC-seq 解析に関する問い合わせ: PhD student, UW-Madison, USA 等

scRNA-seq/scATAC-seq データリクエストに関する問い合わせ: Research Associate, Georgia Institute of Technology, USA 等

**3. 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. データ産生、情報解析を遂行する次世代の生命データサイエンス人材を育成する教育プログラム DSTEP を開始し、本データベースを教育用教材としても活用している。この枠組みにおいて本データベースは民間企業が大規模データ利活用を行う情報インフラの構築に向けて重要なモデルになっていると考えている。
2. MuNAGe(Mutated Nucleotide and Amino-acid sequence Generator)は、NGS 結果から質量分析同定用のタンパク質配列を生成するソフトウェア(論文投稿中)である。このツールの検証に DBKERO に掲載されている3種の肺腺癌細胞株の WGS 及び RNA-Seq のデータが活用された。

## §5. 今後の展開

近年、ヒトゲノム関連データの利活用に関する議論を耳にする機会が増えた。ただし、これらはがんを始めとする疾患ゲノムについて、その DNA 配列を1次元的に解析するものであることが多いように思える。実際、DNA 配列の異同と患者臨床情報の相関を、人工知能等の情報学的手法を駆使して、対応付けしようとする試みは数多い。しかし、いかにしてゲノム配列の変異がエピゲノム制御を介して、トランスクリプトームとして読みだされているのか、その生物学的根拠なしに DNA 配列と臨床像を直接結び付けようとするスキームは基礎生物学者の観点からはやや短兵急に過ぎるように思える。実際、依然として多くのがんについて変異情報に基づいた薬剤療法での根治は困難であり、他の疾患、特に多因子疾患は、依然として病態機序が不明である。GWAS で検出されたゲノム領域が遺伝子発現制御に関わるとされる遺伝子間領域に位置するケースが 90%を超える、とする推定もある。これらに関しては、より高次元のオーミクス空間において、より生物学的証拠に依拠した仮設構築を進める必要がある。そのためにも本データベースで試みたような多層オーミクスデータの補完的利用は必須に思える。今期の開発でもデータ収載を実践したように、シングルセル解析、大規模統合解析、長鎖リード解析等の新規の技術の進展も著しい。これらを臨床検体の解析へと直接に利用した例も急速に増えている。本データベース該当部分のさらなる拡充は製薬企業からの関心も高い部分である。さらに高い精度と密度でのデータの蓄積はより密接に分子機序に依拠した定量的なモデリングを可能とすると考えている。本研究開発終了後、“ゲノム配列以降”について、ますます急速に進展する“新規技術”より産出される大規模データの“モデル化”を志向する産官学の研究者に提供するデータベースとして、さらなる発展を企画したいと考えている。

また、全体として、特にデータ産生について我が国を取り巻く状況は必ずしも良好ではない、ように思われる。解析技術の革新はますますその速度を増している。せつかく高価なコストを投じてもその活用法が定かでないのではないかと、いうためらいがデータ産生への遅れを生み、それがさらに新しいデータ解析をした経験者の欠乏を生んで、負のサイクルが回り始めているように思う。實際上、海外のデータの利活用を基盤とした学術的・産業的価値の創造には限界がある。それぞれの思惑を持って産出されたデータが必ずしも、利用者にとって合目的であるとは限らないこと、また何より計画に立案～遂行段階から参画できないことで、新規のデータに速やかにアクセスできないという理由による。実際に日進月歩の解析プラットフォームから得られるデータについてのタイムラグに起因する相対的な技術格差、成果格差は徐々に拡大しているように感じる。小規模であつても我が国でデータ産出と一体となった情報解析、データベース公開を実現する枠組みを維持する必要性はますます大きくなると考えている。

これらが比較的数年以内の現場対応であるのに対して、現在は、さらに長期的視野になった立案が必要な時期と思う。ただし長期的視野と言っても、初めてのヒトゲノム解読が公表されてまだ 20 年に満たない。2001 年に 1 人だったゲノムデータが、2010 年で 1000 人ゲノム、2020 年では 100 万人ゲノムへと拡大している。20 年前には想像もしていなかった個人ゲノム配列解読の実益が、科学的知見でなく、その個人へと直接還元される、という時代が確かに実現しはじめている。また一方で、その解析機器の高出力化を多個体に適用するのではなく、内なる個体、すなわち、限られた個体数の多数の細胞の多点計測に応用する、というシングルセル/空間トランスクリプトーム解析もますます速度をまして大規模データ生産フェーズに入ろうとしている。そう遠くない将来、これらが融合、多個体に対して多点解析を行ったデータの産出が行われ始めるように思う。今後、20 年以内には、モデルとしての“ヒト”でなく、リアルワールドでの個人について、ゲノムだけでなく個々の細胞、時機での遺伝子発現等に規定される生物学的“人”のデータ化が実現されるようになるの

かもしれない。その“個人オーミクスの実益”の時代において、参照としての欧米等の公共データベースの利活用の価値はますます高まるにしても、真に重要なのはその国、地域、集団、あるいは産官学民においてそれぞれに取集、産生、流通されるべき大規模データに機動的に対応できる枠組みであるように思う。今後、企画されるデータベースあるいはそれを支える情報集積基盤は、そのような分散化に対応できる仕組みを念頭に設計されるべきかもしれない。

## §6. 自己評価

個々の開発項目については、当初の計画通りに概ね順調に推移している、と考えている。ただし、これは、投下予算に比して十分な成果が得られているという自己評価には直結しない。実際、研究期間開始以来の 5 年間で関連分野を取り巻く状況は一変している。日本人ゲノムあるいはがんゲノムカタログについても、今年度に入って国規模での大規模解析計画が始められようとしている。前述の通り、これを真に生物学的な知識に昇華させるには多層オームクス解析データのさらなる充実が必須になる。これらは研究開始時には必ずしも予見できなかった付帯状況の変化ではある。しかし、ますます拡大する大規模データの利活用について、最新のデータを産官学の現場へ届け、その開発基盤として資すべく、より弾力的な開発計画とその迅速な実現が今後も最重要な課題になると考えている。

現時点での反省点として、DBKERO 自体が雑多なデータの寄せ集めであって統一的な使用感、確固たる目的に欠く、という部分がある。特にゲノム支援でのモデル生物種のデータ、さらには新規技術データ等、急速に拡大する部分を当座しのぎとしてでもその収載を優先させた結果として、統一感の形成は困難であった背景もある。これらについては、今後、データポータルのより集中的な活用、がん創薬等のそれぞれの出口を明確に意識した個別の解析系譜への系統収れん等で実装してくべき、と考えている。

## §7. 外部発表等

### 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)	3件
	国際(欧文)	4件
未発行論文	国内(和文)	0件
	国際(欧文)	0件

#### (2) 論文詳細情報

1. Suzuki A, Kawano S, Mitsuyama T, Suyama M, Kanai Y, Shirahige K, Sasaki H, Tokunaga K, Tsuchihara K, Sugano S, Nakai K, Suzuki Y, “DBTSS/DBKERO for integrated analysis of transcriptional regulation.”, *Nucleic Acids Res.*, 4;46(D1):D229-D238, 2018 (DOI:10.1093/nar/gkx1001).  
概要:本論文では、CREST-IHEC 日本チームで収集した全エピゲノムデータを収載したデータベースを公開したものである。ヒトゲノム多様性データベース(HGVDB)のデータについての更新も併せて記載されている。また特に、近年、シングルセル解析、ロングリード解析あるいは薬剤摂動解析について、新規技術開発により得られた新しいデータベースの枠組みについても記載した。
2. Suzuki A, Onodera K, Matsui K, Seki M, Esumi H, Soga T, Sugano S, Kohno T, Suzuki Y, Tsuchihara K. “Characterization of cancer omics and drug perturbations in panels of lung cancer cells.”, *Scientific Reports*, 9, Article number: 19529, 2019 (DOI: 10.1038/s41598-019-55692-9)
3. Sakamoto Y, Xu L, Seki M, Yokoyama TT, Kasahara M, Kashima Y, Ohashi A, Shimada Y, Motoi N, Tsuchihara K, Kobayashi SS, Kohno T, Shiraishi Y, Suzuki A, Suzuki Y. “Long-read sequencing for non-small-cell lung cancer genomes”, *Genome Res.*, 30(9): 1243-1257, 2020 (DOI: 10.1101/gr.261941.120.)
4. Oka M, Xu L, Suzuki T, Yoshikawa T, Sakamoto H, Uemura H, Yoshizawa AC, Suzuki Y, Nakatsura T, Ishihama Y, Suzuki A, Seki M. “Aberrant splicing isoforms detected by full-length transcriptome sequencing as transcripts of potential neoantigens in non-small cell lung cancer” *Genome Biol.* , 22, Article number: 9, 2021 (DOI: 10.1186/s13059-020-02240-8).

#### 2. その他の著作物(総説、書籍など)

1. 鹿島幸恵、鈴木絢子、関真秀、鈴木穰、「シングルセル解析とがんの不均一性」、*実験医学* Vol.36 No.2(増刊)がん不均一性を理解し、治療抵抗性に挑む、2018
2. 関真秀、鈴木穰、「シングルセル解析とナノポアシーケンシング」、*医学のあゆみ*、266 巻 5号、2018
3. 鹿島幸恵、鈴木絢子、関真秀、鈴木穰、「ゲノム医療における一細胞解析」、*実験医学* 増刊 36 巻 15 号、2018

### 3. 国際学会発表及び主要な国内学会発表

#### (1) 概要

種別	国内外	件数
招待講演	国内	7件
	国際	9件
口頭発表	国内	6件
	国際	4件
ポスター発表	国内	0件
	国際	3件

#### (2) 招待講演

##### 〈国内〉

1. 鈴木穰、「シングルセル解析」、日本癌学会、パシフィコ横浜(神奈川)、2017年9月28日-30日
2. 鈴木穰、「ロングリード解析とシングルセル解析」、日本内分泌学会、フェニックス・シーガイア・リゾート(宮崎)、2018年4月26-28日
3. 鈴木穰、「がん細胞株のシングルセル解析:複数プラットフォーム間の横断解析」、日本分子腫瘍マーカー研究会、大阪国際会議場(大阪)、2018年9月26日
4. 鈴木穰、「ゲノム解析技術の進展、共同利用・共同研究拠点「マルチオミクスによる遺伝子発現制御の先端的医学共同研究拠点」シンポジウム、横浜市立大学 福浦キャンパスへボンホール、2020年2月13日
5. 鈴木穰、「がんのシングルセルおよび空間トランスクリプトーム解析、第79回日本癌学会学術総会、リーガロイヤルホテル広島/メルパルク広島(オンラインとのハイブリッド開催)、2020年10月1-3日
6. 鈴木穰、「シングルセル解析技術の進展、第19回日本再生医療学会総会、2020年5月18-29日、2020年8月27日-9月10日、国内(オンライン)、口頭
7. Yutaka Suzuki, Introduction to Single Cell Analysis、第84回日本循環器学会学術総会、オンライン開催、2020年7月27日-8月2日

##### 〈国際〉

1. Yutaka Suzuki, "Multi-omics analysis of cancer cells"、第12回国際ゲノム会議、学術総合センター 一橋講堂(東京)、2017年6月27日-29日
2. Yutaka Suzuki, "Single cell analytical platform"、Human Cell Atlas - Asia Meeting、沖縄科学技術大学院大学(沖縄)、2017年11月30日・12月1日
3. Yutaka Suzuki, "Multi-omics analysis of cancer cells"、Human Genome Meeting 2018、パシフィコ横浜(神奈川)、2018年3月12日-15日
4. Yutaka Suzuki, "Combinatory use of distinct single-cell RNA-seq analytical platforms reveals the heterogeneous transcriptome response"、IUBMB SEOUL, COEX, Seoul, Korea, June 4-8, 2018

5. Yutaka Suzuki, "Single Cell Analysis—from Technical Evaluation to Clinical Application: Pipeline Construction and Data Analysis", 10X Genomics UGM, Matrix Building, Singapore, October 26, 2018
6. Yutaka Suzuki, "Single Cell and Long Read Sequencing of Cancer Cells", TBC/BI OINFO, Hotel Samjung, Seoul, 韓国, October 31 - November 2, 2018
7. Yutaka Suzuki, "Single Cell Analysis Reveals the Heterogeneity of Cancer Cells: Lessons from the Analyses of Adult T Cell Leukemia and Lung Cancer Cells", The 28th KOGO Annual Conference 2019, Sejong Convention Center, Seoul, Korea, September 4-6, 2019
8. Yutaka Suzuki, "Identification and characterization of a new class of cancerous structural variations using PromethION", Nanopore Tech Tour, Beijing, China, September 23, 2019
9. Yutaka Suzuki, "Characterization of complicated structural variants with local copy-number changes in cancer genomes using PromethION", the 19th International Symposium on Pediatric Neuro-Oncology 2020, 軽井沢プリンスホテルウエスト, December 13-16, 2020

### (3) 口頭講演

〈国内〉

1. 菅野純夫、「疾患ヒトゲノム変異の生物学的機能注釈を目指した多階層オミクスデータの統合」、トーゴの日シンポジウム 2017、東京大学弥生講堂一条ホールアネックス(東京)、2017年10月4日・5日
2. 鈴木穰、「生命科学のデータベース活用法(フォーラム)」、日本分子生物学会、神戸ポートアイランド(兵庫)、2017年12月6日-9日
3. 菅野純夫、「疾患ヒトゲノム変異の生物学的機能注釈を目指した多階層オミクスデータの統合」、トーゴの日シンポジウム、日本科学未来館(東京)、2018年10月5日
4. Ayako Suzuki, Sarun Sereewattanawoot, Yutaka Suzuki, Katsuya Tsuchihara, "Multi-omics characterization of lung cancer cells based on gene co-expression modules", 日本癌学会、大阪国際会議場(大阪)、2018年9月27-29日
5. 鈴木 絢子, Sereewattanawoot Sarun, 土原一哉, 鈴木穰「トランスクリプトームモジュールによるがんの新規層別化」(口頭発表+ポスター)第42回日本分子生物学会年会、福岡国際会議場・マリンメッセ福岡、2019年12月3-6日
6. 鈴木穰、疾患ヒトゲノム変異の生物学的機能注釈を目指した多階層オミクスデータの統合、第43回日本分子生物学会年会、オンライン開催、2020年12月2-4日

〈国際〉

1. Yutaka Suzuki, "Multi-omics analysis of cancer cells", The 1st International Symposium for Trans-Omics、東京大学小柴ホール(東京)、2017年11月21日・22日
2. Yutaka Suzuki "Phasing regulatory SNVs to their regulating transcript regions in cancers by using MinION", London Calling 2018, May 24-25, 2018, Old Billingsgate, London, UK, May 24, 2018
3. Suzuki A, Sakamoto Y, Suzuki Y. "Identification of a new class of structural aberrations in lung cancer genomes using a long read sequencer". (selected short t

alk), The 13th international Workshop on Advanced Genomics (13AGW), 東京・学術総合センター 一橋講堂, June 26-27, 2019

4. Suzuki A, Sakamoto Y, Suzuki Y. "Characterization of complicated structural variants with local copy-number changes in lung cancer genomes using PromethION". (selected flash talk + poster), Advances in Genome Biology and Technology (AGBT) 2020 General Meeting. 米・Marco Islands, February, 23-26, 2020

#### (4) ポスター発表

〈国内〉

該当なし。

〈国際〉

1. Ayako Suzuki, Sarun Sereewattanawoot, Yutaka Suzuki, Katsuya Tsuchihara, "Characterization of aberrant transcriptional regulations in lung cancers", Advances in Genome Biology and Technology 2018, Hilton Orlando Bonnet Creek, Orland, Florida, USA, 2018年2月12日-15日
2. Ayako Suzuki, Masahide Seki and Yutaka Suzuki, "Identification and Characterization of a new class of cancerous structural variations using PromethION", Advances in Genome Biology and Technology (AGBT) 2019 General Meeting, Marco Islands, FL, USA, February 27 - March 2, 2019
3. Yutaka Suzuki, "Integration of the Panel Sequencing and In-Depth Single Cell Analyses Reveals the Clonal Evolution of Adult T Cell Leukemia Cells", Advances in Genome Biology and Technology (AGBT) 2020 General Meeting, 米・Marco Islands, February 23-26, 2020

#### 4. 知財出願

##### (1) 出願件数

種別	件数	
特許出願	国内	0件
	国外	0件
その他の知的財産出願		0件

##### (2) 一覧

###### ① 国内出願

該当なし。

###### ② 海外出願

該当なし。

###### ③ その他の知的財産権

該当なし。

## **5. 受賞・報道等**

### **(1) 受賞**

該当なし。

### **(2) メディア報道**

該当なし。

### **(3) その他**

該当なし。

## §8. 研究開発期間中の活動

### 1. 進捗ミーティング

年月日	名称	場所	参加人数	目的・概要
2017年 4月24日	キックオフミーティング (非公開)	東京大学医学部	15人	本研究開発費で行う研究の役割分担の調整
2017年 4月28日	チーム内ミーティング(非公開)	東京大学フューチャーセンター	5人	同上
2017年 5月15日	チーム内ミーティング(非公開)	東京大学医学部	3人	進捗状況確認のためのミーティング
2017年 6月19日	チーム内ミーティング(非公開)	東京大学医学部	3人	同上
2017年 7月14日	チーム内ミーティング(非公開)	東京大学柏キャンパス	3人	同上
2017年 7月18日	沖真弥チーム、NBDCとのミーティング(非公開)	JST 東京本部	10人	追加実施策に関する打ち合わせ
2017年 7月19日	チーム内ミーティング(非公開)	東京大学医学部	3人	進捗状況確認のためのミーティング
2018年 3月28日	チーム内ミーティング(非公開)	東京大学医学部	3人	同上
2018年 8月30日	チーム内ミーティング(非公開)	東京大学柏キャンパス	5人	研究進捗報告のためのミーティング
2018年 11月22日	チーム内ミーティング(非公開)	東京大学柏キャンパス	5人	研究進捗報告のためのミーティング
2019年 4月26日	チーム内ミーティング(非公開)	東京大学柏キャンパス	4人	進捗状況確認のためのミーティング
2019年 5月29日	チーム内ミーティング(非公開)	東京大学柏キャンパス	4人	同上
2019年 6月20日	チーム内ミーティング(非公開)	東京大学柏キャンパス	4人	同上
2019年 7月26日	チーム内ミーティング(非公開)	東京大学柏キャンパス	4人	同上
2019年 8月23日	チーム内ミーティング(非公開)	東京大学柏キャンパス	4人	同上
2019年 9月25日	チーム内ミーティング(非公開)	東京大学柏キャンパス	4人	同上
2019年 10月2日	チーム内ミーティング(非公開)	東京大学柏キャンパス	4人	同上
2019年	チーム内ミーティング(非公開)	東京大学柏	4人	進捗状況確認のためのミーテ

年月日	名称	場所	参加人数	目的・概要
4月26日	公開)	キャンパス		イング
2019年 5月29日	チーム内ミーティング(非公開)	東京大学柏キャンパス	4人	同上
2019年 6月20日	チーム内ミーティング(非公開)	東京大学柏キャンパス	4人	同上
2019年 7月26日	チーム内ミーティング(非公開)	東京大学柏キャンパス	4人	同上
2019年 8月1日	チーム内ミーティング	国際医療センター(外山)	5人	投入データのデータ処理等についての打ち合わせ
2019年 8月23日	チーム内ミーティング(非公開)	東京大学柏キャンパス	4人	同上
2019年 9月25日	チーム内ミーティング(非公開)	東京大学柏キャンパス	4人	同上
2019年 10月2日	チーム内ミーティング(非公開)	東京大学柏キャンパス	4人	同上
2019年 12月20日	チーム内ミーティング	東京大学 柏キャンパス	3人	DBKEROの機能・ユーザインタフェース改良について打ち合わせ
2020年 1月8日	チーム内ミーティング	東京大学 柏キャンパス	3人	DBKEROの機能・ユーザインタフェース改良について打ち合わせ
2020年 1月23日 24日	SPARQLthon88にてRDF化に係わるスキーマ設計に関する打ち合わせ	DBCLS 柏ラボ	3人	ChIP-Seqピーク情報のRDF化に関する打ち合わせ。スキーマの設計について
2020年 2月19日	RDF化に係わるスキーマ設計に関する打ち合わせ	DBCLS 柏ラボ	3人	上記に加え、抗原・抗体の表現方法について意見の取り交わし
2020年 2月20日 21日	SPARQLthon89にてRDF化に係わるスキーマ設計に関する打ち合わせ	DBCLS 柏ラボ	3人	・ChIP-Seqピーク情報のRDF化に抗体・抗原についての表現方法について合意 ・IHECのRNA-Seq結果のRDF化について相談
2020年 12月4日	チーム内ミーティング(非公開)	オンライン	5人	DBTSS/DBKERO 講習会・シングルセル解析入門のためのミーティング
2020年 12月18日	チーム内ミーティング(非公開)	オンライン	5人	DBTSS/DBKERO 講習会・シングルセル解析入門のためのミーティング
2021年	チーム内ミーティング(非公開)	東京大学柏	3人	進捗確認のためのミーティング

年月日	名称	場所	参加人数	目的・概要
1月8日	公開)	キャンパス		グ

## 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
2017年 12月6日 ～9日	生命科学系学会合同年次大会(ConBio2017)特別企画「使ってみようバイオデータベース—つながるデータ、広がる世界(BioDB)」	神戸ポートアイランド	30人	目的:データベースの一般への利用と普及の促進 概要:DBの紹介
2018年 11月28日 ～30日	日本分子生物学会特別企画「使ってみようバイオデータベース—つながるデータ、広がる世界(BioDB)」	パシフィコ横浜	30人	目的:データベースの一般への利用と普及の促進 概要:DBの紹介
2018年 12月5日	統合データベース講習会:AJACS名古屋2	愛知県がんセンター 国際医学交流センター	100人	目的:生命科学系のデータベースやツールの使い方を紹介する(初心者向け) 概要:DBKERO およびその他がんゲノム解析関連データベースの紹介
2019年 12月3日	第42回日本分子生物学会年会 フォーラム(チュートリアル企画)「生命科学のデータベース活用法2019」	福岡国際会議場	30人	目的:データベースの一般への利用と普及の促進 概要:DBの紹介
2019年 12月3日 ～6日	第42回日本分子生物学会年会 特別企画「使ってみようバイオデータベース—つながるデータ、広がる世界(BioDB)」	福岡国際会議場	30人	目的:データベースの一般への利用と普及の促進 概要:DBの紹介
2020年 12月9日	DBTSS/DBKERO 講習会・シングルセル解析入門	オンライン	100人	転写開始点および疾患多層オーミクスデータベース DBTSS/DBKERO の利用説明会と、DBKERO データを用いたシングルセル解析の入門講習
2021年 3月10日	空間トランスクリプトーム解析【Visium】講習会	オンライン	100人	空間トランスクリプトーム Visium データ解析の入門講習

以上

別紙 研究開発対象のデータベース等

No.	正式名称	別称	概要	URL	公開日	状態	分類	生命科学系データベースアーカイブ	NBDCヒトデータベース	NBDC RDFポータル	関連文献 (論文リストに記載があれば、その番号でも可)
1	ヒト疾患多層オミクスデータベース	DBKERO	本データベースはヒトゲノム多型・変異に生物学的機能注釈を与えるべく、ゲノム変異位置、近傍のエピゲノム(ヒストン修飾、DNAメチル化パターン)、トランスクリプトーム情報(発現量、スプライスパターン)をヒトゲノム情報に統合したものである。特に、培養細胞系あるいは生物種を超えてマウスをはじめとする動物モデル系から得られたオミクスデータに焦点を当てている。	<a href="https://kero.hgc.jp/">https://kero.hgc.jp/</a>	2015年7月9日	継続・発展	データベース等	対象外	対象外	公開済	Ayako Suzuki, Shin Kawano, Toutai Mitsuyama, Mikita Suyama, Yae Kanai, Katsuhiko Shirahige, Hiroyuki Sasaki, Katsushi Tokunaga, Katsuya Tsuchihara, Sumio Sugano, Kenta Nakai, Yutaka Suzuki. "DBTSS/DBKERO for integrated analysis of transcriptional regulation." Nucleic Acids Res. 2018 (Database issue).