

## 2022 年度 研究開発実施報告

□概要

研究開発課題名	統合的な転写制御データ基盤の構築
開発対象データベースの名称(URL)	INTRARED (INtegrated TRAnscriptional REgulation Data platform) ( <a href="https://www.intrared.org/">https://www.intrared.org/</a> )
研究代表者氏名	粕川雄也
所属・役職	理化学研究所 生命医科学研究センター・チームリーダー(2023年3月時点)



## 目次

§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) 本年度の研究開発計画と達成目標 .....	5
(2) 進捗状況 .....	7
§4. 成果発表等 .....	10
(1) 原著論文発表 .....	10
① 論文数概要 .....	10
② 論文詳細情報 .....	10
(2) その他の著作物(総説、書籍など) .....	11
(3) 国際学会および国内学会発表 .....	11
① 概要 .....	11
② 招待講演 .....	11
③ 口頭講演 .....	12
④ ポスター発表 .....	13
(4) 知的財産権の出願 (国内の出願件数のみ公開) .....	13
① 出願件数 .....	13
(5) 受賞・報道等 .....	13
① 受賞 .....	13
② メディア報道 .....	14
③ その他の成果発表 .....	14
§5. 研究開発期間中に主催した活動(ワークショップ等) .....	14
(1) 進捗ミーティング .....	14
(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 .....	14

## §1. 研究実施体制

グループ名	研究代表者または主たる共同研究者氏名	所属機関・役職名	研究題目
粕川グループ	粕川 雄也	理研IMS・チームリーダー	研究開発の取りまとめ、ヒト・マウスシスエレメントデータの作成、インターフェース開発
川路グループ	川路 英哉	東京都医学総合研究所・副センター長	シスエレメント同定パイプラインの構築と適用
栴屋グループ	栴屋 啓志	理研BRC・室長	ゲノム変異とシスエレメントの対応づけ
沖グループ	沖 真弥	京都大学・特定准教授	トランス因子・エピゲノミクスデータの収集

## §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	INTRARED		<a href="https://www.intrared.org/">https://www.intrared.org/</a>

#### 【その他のデータベース】

No.	名称	別称・略称	URL
1	fanta.bio		<a href="https://fanta.bio/">https://fanta.bio/</a>
2	ChIP-Atlas		<a href="https://chip-atlas.org/">https://chip-atlas.org/</a>
3	SCPortalen		<a href="https://single-cell.riken.jp/">https://single-cell.riken.jp/</a>
4	MoG+		<a href="https://molossinus.brc.riken.jp/mogplus/">https://molossinus.brc.riken.jp/mogplus/</a>

### (2) ツール等一覧

No.	名称	別称・略称	URL
1			

### §3. 実施内容

#### (1) 本年度の研究開発計画と達成目標

##### **【研究開発項目(1) シスエレメントデータの構築と維持】**

##### **研究開発項目(1)-1: シスエレメント領域データの作成**

シスエレメント領域データの作成のため、以下の2つを行う。

##### 1. 新たなシスエレメント同定パイプラインの構築・評価

バルクならびに一細胞レベルの 5'端 RNA 配列データを対象とした新たな同定法によるシスエレメント同定パイプラインを構築・評価を行う。シスエレメント同定法には、現在研究分担グループで開発を進めている新規手法を含めて複数あるため、これら进行评估し同定法を決定する。

##### 2. シスエレメント領域データの $\alpha$ 版データを作成

限定されたデータセットについて、上記パイプラインを適用しシスエレメント領域データの  $\alpha$  版データを作成する。

##### **研究開発項目(1)-2: 細胞の種類・状態毎のシスエレメント活性データの作成**

シスエレメント領域データの作成のため、以下の2つを行う。

##### 1. シスエレメント活性量を産出するための手法検討・評価

バルクの5'端 RNA 配列ならびに一細胞由来の 5'端 RNA 配列の RNA-Seq データからシスエレメント活性量を産出するための手法の検討を行う。検討すべき項目として、使用するデータセットの評価法の策定、一細胞データから細胞種や状態ごとに分類する方法の検討、正規化の方法などがあり、これらを検討する。

##### 2. シスエレメント活性データの $\alpha$ 版データを作成

限定されたデータセットについて、上記パイプラインを適用しシスエレメント領域データの  $\alpha$  版データを作成する。活性データは研究開発項目(1)-1 で作成するシスエレメント領域データを基準に活性量データを作成する。

##### **研究開発項目(1)-3 シスエレメント近傍のゲノム変異情報の収集と維持**

本課題で収集・整理するシスエレメント領域について、疾患モデル動物として用いられる実験動物マウスの各系統における配列の多様性(ゲノムバリエーション)を、各シスエレメントのメタ情報としてデータ化する

##### 1. シスエレメントとゲノム変異との対応づけ方法の開発

シスエレメントデータに対して、ゲノムの短鎖解読由来、および長鎖解読由来、それぞれの解析により得られたゲノムバリエーション情報を統合的に可視化するための準備開発を行う。VCF 形式を基盤として、由来する組織や細胞株の表現型情報をいかに記述するかを策定する。

##### 2. 対応データ、メタデータの作成

マウス短鎖解読由来の約 50 系統の情報について、上記の方法に基づき、VCF データとして収集、および作成する。ヒトについても公開データベースから VCF を収集し、ゲノムブラウザにおいてシスエレメントとの位置関係を可視化する。

## 【研究開発項目(2) ChIP-Atlas 高度化によるエピゲノミクス・トランス因子データの構築と維持】

### 研究開発項目(2)-1 エピゲノミクス・トランス因子データの新規コンテンツの追加

下記の機能注釈データを準備する。

#### 発現制御に関わる機能性領域 (ChromHMM, FANTOM enhancer)

最近報告された 100 種類の細胞タイプの ChromHMM データ(H. Vu, & J. Ernst. Genome Biol. 2022)を入手し、Strong Enhancer, Heterochromatin など 15 種類の分類群をゲノムブラウザ (IGV) で色分け表示できるような BED ファイルを作成する。FANTOM enhancer データは、FANTOM5 プロジェクトの Web サイトから入手し、IGV 上では細胞名とエンハンサーの強度に応じて濃淡表示できるような BED ファイルを作成する。

#### 疾患関連 SNPs (GWAS Catalog, ClinVar)

GWAS Catalog と ClinVar より種々の形質や疾患と関連する SNP 座位を収集し、IGV 上では形質名と P-value に応じて濃淡表示できるような BED ファイルを作成する。

#### 量的形質 SNPs (eQTL)

GTEx より遺伝子の発現変動と関連する SNP 座位を収集し、IGV 上では P-value に応じて濃淡表示できるような BED ファイルを細胞名ごとに作成する。

#### その他レファレンス情報 (RepeatMasker, phastCons, LD-block)

RepeatMasker, phastCons, LD-block などの情報は全て UCSC Genome Browser において BigWig または BED 形式のファイル公開されているため、それらの IGV での閲覧やダウンロードを可能にする。

### 研究開発項目(2)-2 ChIP-Atlas データ更新

ChIP-Atlas は 2021 年度までの大規模アップデート(Zou, et al., 2022 Nucleic Acids Res)で Bisulfite-seq と ATAC-seq データを追加したが、その計算に多大な時間とリソースを費やしたため、2020 年 8 月以降のデータに対応できていない。更新のための自動化パイプライン(下記)を通して解析し、公開をめざす。

- 1) SRA メタデータのダウンロードと更新データのチェック
- 2) 一次解析(シーケンスデータのダウンロード、マッピング、ピークコール)
- 3) サンプルメタデータのキュレーション
- 4) データ統合、データマイニング結果の作成
- 5) 公開サーバへのデータ転送

なお、上記の 3) サンプルメタデータのキュレーションでは、これまで同様、SRA メタデータに記載された細胞・組織名や ChIP の抗原名などを controlled vocabulary またはそれに準ずる表記法に変換する(例:細胞名 MDA231 → MDA-MB-231、転写因子 Oct4 → Pou5f1)。前期プログラムにおいて開発した機械学習による半自動化システムを活用し、最終的には複数の研究参加者によるダブルチェックをおこない、キュレーションの品質管理に努める。

### 【研究開発項目(3) データベース連携】

今後のシスエレメント・トランス因子・エピゲノミクスデータの連携を視野に入れて、まずはシスエレメントデータの  $\alpha$  版データについて、UCSC の TrackHub 用サイトを構築し、UCSC Genome Browser で参照できるようにする。また、シスエレメントデータと ChIP-Atlas のトランス因子データについての一括表示の方法について検討ならびにテスト実装を行う。

## (2) 進捗状況

### 【研究開発項目(1) シスエレメントデータの構築と維持】

#### 研究開発項目(1)-1: シスエレメント領域データの作成、研究開発項目(1)-2: 細胞の種類・状態毎のシスエレメント活性データの作成

NET-CAGE 法についての論文 (Hirabayashi S et al., 2019. <https://doi.org/10.1038/s41588-019-0485-9>) で提供されているヒト5細胞種 (MCF-7, GM12878, HeLa S3, Hep G2, K-562) ならびにマウス2臓器 (kidney, brain) を対象とした CAGE 法および NET-CAGE 法による 5'端 RNA 配列データを用いて、シスエレメントの同定を行った。シスエレメント領域の同定ならびにゲノム座標ごとの活性量データの作成には川路グループで新たに開発した CREate 法を用いた。作成したシスエレメント領域ならびにゲノム座標ごとの活性データについては UCSC TrackHub 用のデータセットを作成し、UCSC Genome Browser で表示できるようにした。

シスエレメントデータを参照するためのウェブサイトを作成し 2023 年 3 月末に公開した (<https://fanta.bio/>)。fanta.bio サイトでは UCSC TrackHub データの提供ならびに UCSC Genome Browser で表示するためのリンクの提供、バルクダウンロード用のページ、簡単な検索インターフェースを提供している。

#### 研究開発項目(1)-3 シスエレメント近傍のゲノム変異情報の収集と維持

##### シスエレメントとゲノム変異との対応づけ方法の開発

マウス 46 系統のゲノムバリエーション情報を VCF 形式で作成し代表機関と共有した。加えて、短鎖解読由来のゲノム解析により得られた 10 種類のマウス系統のゲノムバリエーションデータを、シスエレメントデータと同時表記するためのアノテーション作業を行なった。この作業において、データの由来する生物、組織、細胞株およびその表現型情報をいかに記述するかを代表機関と連携して策定した。さらに、実験的に確定されたマウスのシスエレメントおよびその変異ともたらされた表現型について、MGI、IMPC 等のマウスのゲノムおよび表現型データベースや公開論文情報から、ゲノム座標およびゲノム配列データを抽出するための調査を行い、これに基づいて、一部のデータ抽出を開始した。

##### 対応データ、メタデータの作成

シスエレメントデータとゲノムバリエーション情報との対応データの作成を開始した。研究分担者グループ(1)および(3)が公開するデータ形式等への対応の検討をおこなった。短鎖解読由来の 46 種のゲノムバリエーション情報については、2022 年度中に VCF に加え、DBCLS、NBDC、遺伝研が策定を開始している VCF の RDF 版である形式への変換作業を完了した。

以上、2022 年度は、短鎖解読由来の 46 種のゲノムバリエーション情報を中心に作業を進める事ができた。長鎖配列解読由来ゲノム多型データの整備については、別プロジェクトで進行中であり 2023 年度からの作業を計画している。その他、他のデータベースとの連携として、日本人の疾患関連ゲノムバリエーション情報を公開している TogoVar との連携を開始した。

### 【研究開発項目(2) ChIP-Atlas 高度化によるエピゲノミクス・トランス因子データの構築と維持】

2022 年度のキックオフ時においては、2020 年 8 月までの ChIP-seq, WGBS, ATAC-seq, DNase-seq データ合わせて 185,978 件が ChIP-Atlas に収録されていたが、最終的には 2023 年 3 月までのデータの収録を達成した (86,305 件の追加で、272,283 件を公開)。おもな実施項目は下記の通り。

## 1) 一次解析

- ・ シーケンスデータのダウンロード
- ・ レファレンスゲノムへのアライメント
- ・ ピークコールデータの作成 (BED 形式; WGBS は高/低メチル化領域)
- ・ カバレッジデータの作成 (BigWig 形式; WGBS はメチル化率データ)

## 2) サンプルメタデータのキュレーション

- ・ 既存データを学習セットとした機械学習モデルを作成
- ・ 新規 SRA metadata の入手、機械学習による予測
- ・ 人員によるチェック
- ・ 予測できなかったデータは文献などを参照し、人員によるキュレーション

## 3) 統合解析

- ・ 全てのピークコールデータの結合、ソート
- ・ Peak Browser ツール用に、IGV 閲覧のための加工データとインデクス作成
- ・ Target Genes ツール用に、各転写因子が結合する遺伝子の特定
- ・ Colocalization ツール用に、共起的に結合する転写因子の特定
- ・ Enrichment Analysis ツール用のライブラリファイルの作成

## 4) データ公開

- ・ 公開サーバーへのアップロード

上記の多くは既存のパイプラインを活用したが、2年ほどのギャップの間にデータ増加による遅延や不具合が生じたため、パイプラインの修正を図った。おもな変更点は下記の通り。

### 1) 一次解析

数万件もの実験データを迅速に処理するため、解析サーバ(遺伝研スパコン)の有料計算ノードを使用してスレッド数を引き上げた。

### 2) サンプルメタデータのキュレーション

教師データ(キュレーション前のサンプルメタデータと、キュレーション後の文字列)が増大したため、学習モデルを作成するための時間が二週間を超えるようになった。そのため、教師データを分割してマルチスレッド処理による機械学習ができるようスクリプトを修正した。その結果、どの生物種も2日以内に学習モデルを作成できるよう改善できた。

### 3) 統合解析

全てのピークコールデータを結合してソートすることは、その後の統合解析データの作成に必須であるが、データ量の増加に伴い、全ピークデータを結合してソートする時間が一週間を超えるようになった。そのため、全ピークデータを結合した後にソートするのではなく、分割ソートと結合を繰り返すようにスクリプトを修正した。その結果どの生物種も2日以内にソート済みの統合ピークデータを作成できるよう改善できた。

### 4) データ公開

従来の NBDC 公開サーバーが廃止するとの通達により、DBCLS の公開サーバーを使用することになった。そのため、サーバーの設置と転送コマンドの整備をおもに DBCLS の大田達郎氏のご協力のもと進めた。また全てのデータに付与されている URL を新サーバー用に変更することで、新サーバーからのデータのダウンロードや閲覧が可能となるよう、Web アプリの設定を変更した。新サーバーはクライアント(遺伝研スパコン)



から距離的に近いこともあり、転送速度が大幅に改善し、3日以内に更新データを転送できるようになった。

このように、2022年度は単に更新の遅れを取り戻すだけでなく、データ量増加に伴う更新パイプラインの迅速化や、メタデータのキューレーションの効率化に多くのエフォートを割いた。これはデータベースの持続可能性を維持するための最優先課題として実施したが、そのため、エピゲノミクス・トランス因子データの新規コンテンツの追加は達成できなかった。一方で2023年1月以降は従来通り毎月の更新を継続しており、今後しばらくは持続可能なデータベースとしての継続が可能と思われる。

### 【研究開発項目(3) データベース連携】

今年度は fanta.bio と ChIP-Atlas の一括表示のため、以下を行った

- Fanta.bio の  $\alpha$  版データで使用したサンプルについて、ChIP-Atlas と同じ細胞種名になるよう、名前の統一を行った。具体的には、ChIP-Atlas ですでに使用していた cellosaurus の ID を参照し、そこで使用されている office name を使用することで、同じ名前になるようにすることとした。
- $\alpha$  版データで使用したサンプルと同じサンプルの ChIP-Atlas データについて、fanta.bio 用 UCSC TrackHub データに追加し、UCSC Genome Browser 上で一括参照できるようにした。これによりシスエレメントとトランス因子のデータが同時に参照できるようになった。

また本プロジェクトの統合的転写制御データ基盤のウェブサイト (INTRARED) を作成し、公開した (<https://www.intrared.org/>)。このサイトでは、本データ基盤の説明や、fanta.bio および ChIP-Atlas データベースの説明ならびにウェブサイトへのリンク等の情報を提供している。今後統合検索などのインターフェース設置などを検討している。

## §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)	12 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. Masaru Koido, Chung-Chau Hon, Satoshi Koyama, Hideya Kawaji, Yasuhiro Murakawa, Kazuyoshi Ishigaki, Kaoru Ito, Jun Sese, Nicholas F. Parrish, Yoichiro Kamatani, Piero Carninci, Chikashi Terao. Prediction of the cell-type-specific transcription of non-coding RNAs from genome sequences via machine learning. *Nature Biomedical Engineering* 2022. doi: 10.1038/s41551-022-00961-8.
2. Sanna Vuoristo, Shruti Bhagat, Christel Hydén-Granskog, Masahito Yoshihara, Lisa Gawrylski, Eeva-Mari Jouhilahti, Vipin Ranga, Mahlet Tamirat, Mikko Huhtala, Ida Kirjanov, Sonja Nykänen, Kaarel Krjutskov, Anastassios Damdimopoulos, Jere Weltner, Kosuke Hashimoto, Gaëlle Recher, Sini Ezer, Priit Paluoja, Pauliina Paloviita, Yujiro Takegami, Ai Kanemaru, Karolina Lundin, Tomi T Airenne, Timo Otonkoski, Juha S Tapanainen, Hideya Kawaji, Yasuhiro Murakawa, Thomas R Berglin, Markku Varjosalo, Mark S Johnson, Timo Tuuri, Shintaro Katayama, Juha Kere. DUX4 is a multifunctional factor priming human embryonic genome activation. *iScience* 2022. 25(4):104137-104137. doi: 10.1016/j.isci.2022.104137.
3. Ryuichi Nakagawa, Kei Takasawa, Maki Gau, Atsumi Tsuji-Hosokawa, Hideya Kawaji, Yasuhiro Murakawa, Shuji Takada, Masashi Mikami, Satoshi Narumi, Maki Fukami, Rajini Sreenivasan, Tetsuo Maruyama, Elena Tucker, Liang Zhao, Josephine Bowles, Andrew Sinclair, Peter Koopman, Yoshihide Hayashizaki, Tomohiro Morio, Kenichi Kashimada. Two ovarian candidate enhancers, identified by time series enhancer RNA analyses, harbor rare genetic variations identified in ovarian insufficiency. *Human molecular genetics* 2022. doi: 10.1093/hmg/ddac023.
4. Akane Yoshikawa, Itaru Kushima, Mitsuhiro Miyashita, Kazuhiro Suzuki, Kyoka Iino, Kazuya Toriumi, Yasue Horiuchi, Hideya Kawaji, Norio Ozaki, Masanari Itokawa, Makoto Arai. Exonic deletions in *IMMP2L* in schizophrenia with enhanced glycation stress subtype. *PloS one* 2022. 17(7):e0270506. doi: 10.1371/journal.pone.0270506.
5. Kazuo Miyazawa, Kaoru Ito, Masamichi Ito, Zhaonan Zou, Masayuki Kubota, Seitara Nomura, Hiroshi Matsunaga, Satoshi Koyama, Hirotaka Ieki, Masato Akiyama, Yoshinao Koike, Ryo Kurosawa, Hiroki Yoshida, Kouichi Ozaki, Yoshihiro Onouchi, Koichi Matsuda, Yoshinori Murakami, Yoichiro Kamatani, Atsushi Takahashi, Koichi Matsuda, Yoshinori Murakami, Hiroyuki Aburatani, Michiaki Kubo, Yukihide Momozawa, Chikashi Terao, Shinya Oki, Hiroshi Akazawa, Yoichiro Kamatani, Issei Komuro. Cross-ancestry genome-wide analysis of atrial fibrillation unveils disease biology and enables cardioembolic risk prediction. *Nature Genetics* 2023. doi: 10.1038/s41588-022-01284-9.
6. Zhaonan Zou, Michio Iwata, Yoshihiro Yamanishi, Shinya Oki. Epigenetic landscape of drug responses revealed through large-scale ChIP-seq data analyses. *BMC Bioinformatics* 2022. 23(1):51. doi: 10.1186/s12859-022-04571-8.

7. Michio Iwata, Keisuke Kosai, Yuya Ono, Shinya Oki, Koshi Mimori, Yoshihiro Yamanishi. Regulome-based characterization of drug activity across the human diseasome. *npj Systems Biology and Applications* 2022. 8(1):44. doi: 10.1038/s41540-022-00255-4.
8. Hirohito Hirata, Asana Kamohara, Masatoshi Murayama, Kenichi Nishioka, Hiroaki Honda, Yasuteru Urano, Hidenobu Soejima, Shinya Oki, Toshio Kukita, Shunsuke Kawano, Masaaki Mawatari, Akiko Kukita. A novel role of helix-loop-helix transcriptional factor Bhlhe40 in osteoclast activation. *Journal of Cellular Physiology* 2022. 237(10):3912-3926. doi: 10.1002/jcp.30844.
9. Yuko Hino, Katsuya Nagaoka, Shinya Oki, Kan Etoh, Shinjiro Hino, Mitsuyoshi Nakao. Mitochondrial stress induces AREG expression and epigenomic remodeling through c-JUN and YAP-mediated enhancer activation. *Nucleic Acids Research* 2022. 50(17):9765-9779. doi: 10.1093/nar/gkac735.
10. Ryohei Eguchi, Momoko Hamano, Michio Iwata, Toru Nakamura, Shinya Oki, Yoshihiro Yamanishi. TRANSDIRE: data-driven direct reprogramming by a pioneer factor-guided trans-omics approach. *Bioinformatics* 2022. 38(10):2839-2846. doi: 10.1093/bioinformatics/btac209.
11. Zhaonan Zou, Tazro Ohta, Fumihito Miura, Shinya Oki. ChIP-Atlas 2021 update: a data-mining suite for exploring epigenomic landscapes by fully integrating ChIP-seq, ATAC-seq and Bisulfite-seq data. *Nucleic Acids Research* 2022. 50(W1):W175-W182. doi: 10.1093/nar/gkac199.
12. Hitomi Yamamoto-Imoto, Satoshi Minami, Tatsuya Shioda, Yurina Yamashita, Shinsuke Sakai, Shihomi Maeda, Takeshi Yamamoto, Shinya Oki, Mizuki Takashima, Tadashi Yamamuro, Kyosuke Yanagawa, Ryuya Eda, Miki Iwatani, Mizue So, Ayaka Tokumura, Toyofumi Abe, Ryoichi Imamura, Norio Nonomura, Yukinori Okada, Donald E. Ayer, Hidesato Ogawa, Eiji Hara, Yoshitsugu Takabatake, Yoshitaka Isaka, Shuhei Nakamura, Tamotsu Yoshimori. Age-associated decline of MondoA drives cellular senescence through impaired autophagy and mitochondrial homeostasis. *Cell Reports* 2022. 38(9):110444. doi: 10.1016/j.celrep.2022.110444.

**(2) その他の著作物(総説、書籍など)**

該当なし

**(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	22 件
	国際	1 件
口頭発表	国内	5 件
	国際	1 件
ポスター発表	国内	3 件
	国際	1 件

② 招待講演

〈国内〉

1. 沖 真弥. 光単離化学(PIC)による局所的トランスクリプトーム解析. 第 100 回日本生理学会大会. 2023/3/3/16.
2. 本田 瑞季. ミクロ組織を踏まえた空間的トランスクリプトーム. 第 39 回日本毒性病理学会. 2023/1/25.

3. 沖 真弥. 空間的な遺伝子発現制御のしくみを探る. 東北大学分子代謝病態学 Web セミナー. 2023/1/17.
4. 沖 真弥. Photo-Isolation Chemistry を用いた局所的なトランスクリプトーム解析. 第 67 回日本人類遺伝学会. 2022/12/15.
5. 沖 真弥. 光単離化学による局所的かつ高深度のトランスクリプトーム解析. 第 45 回日本分子生物学会. 2022/12/2.
6. 粕川 雄也、川路 英哉、榎屋 啓志、転写制御理解のためのゲノムワイドなシスエレメントデータリソースの構築、第 45 回日本分子生物学会年会、幕張、千葉、2022/11/30.
7. 鄒 兆南. 統合的な転写制御データ基盤の構築 ChIP-Atlas update: Bisulfite-seq と ATAC-seq データを統合しエピゲノム制御の全貌に迫る. 第 45 回日本分子生物学会年会. 2022/11/30.
8. 沖 真弥. Photo-isolation chemistry による局所的高深度トランスクリプトーム解析. ExCELLS セミナー. 2022/11/15.
9. 沖 真弥. 空間的な遺伝子発現制御のしくみを探る. NGS EXPO 2022. 2022/10/19.
10. 沖 真弥. 空間的な遺伝子発現制御のしくみを探る. 東京慈恵会医科大学 医学研究の基礎を語り合う集い. 2022/10/11.
11. 粕川 雄也、統合的な転写制御データ基盤の構築、トーゴの日シンポジウム、オンライン、2022/10/5.
12. 本田 瑞季, 沖 真弥. 光単離化学 (PIC) による 高解像度かつ高深度トランスクリプトーム解析. 第 81 回 日本癌学会学術総会. 2022/10/1.
13. 川路 英哉、FANTOM5 における UCSC Genome Browser Track Hub の利用とその展開, Biopackathon #9, オンライン, 2022/9/2
14. 粕川 雄也、転写制御データ基盤のためのシスエレメント・データベースの開発、Biopackathon #9、オンライン、2022/9/2
15. 沖 真弥. Photo-isolation chemistry による局所的高深度トランスクリプトーム解析. シングルセルゲノミクス研究会 2022. 2022/8/30.
16. 沖 真弥. データ駆動型解析による noncoding 病モデルマウスの作製. 株式会社セツロテック コラボ Web セミナー. 2022/8/4.
17. 沖 真弥. 空間的な遺伝子発現制御のしくみを探る. 第 25 回小児心血管分子医学研究会. 2022/7/21.
18. 沖 真弥. 空間的な遺伝子発現制御のしくみを探る. 徳島大学 医学研究科セミナー. 2022/7/6.
19. 沖 真弥. Photo-Isolation-Chemistry (PIC): 局所領域に対する高深度トランスクリプトーム解析. 京都大学 先端バイオメディシン解析技術室シンポジウム. 2022/6/30.
20. 沖 真弥. Photo-isolation chemistry for high-resolution and deep spatial transcriptome with tissue sections. 第 74 回日本細胞生物学会大会. 2022/6/29.
21. 沖 真弥. Data-driven and technical approaches to understand spatial gene regulation. 京都大学医学研究科・神経科学教育コース. 2022/6/27.
22. 木村 龍一. 光を用いた空間トランスクリプトーム解析. 第 4 回タタバイオ分子クラブ. 2022/5/16.

〈国際〉

1. Shinya Oki. Photo-isolation chemistry for high-resolution and deep spatial transcriptome with tissue sections. IMGC 2023 RIKEN Symposium. 2023/3/30.

### ③ 口頭講演

〈国内〉

1. 本田 瑞季, 木村 龍一, 原田 哲仁, 前原 一満, 田中 かおり, 大川 恭行, 沖 真弥. Photo-Isolation Chemistry による高解像度かつ高深度空間トランスクリプトーム解析. 第 45 回日本分子生物学会. 2022/12/1.
2. 鄒 兆南, 沖 真弥. 大規模 NGS データを横断し薬剤と疾患のしくみを理解する. NGS EXPO 2022. 2022/10/18.
3. 本田 瑞季, 木村 龍一, 原田 哲仁, 前原 一満, 田中 かおり, 大川 恭行, 沖 真弥. Photo-Isolation Chemistry による高解像度かつ高深度空間トランスクリプトーム解析. NGS EXPO2022. 2022/10/18.
4. 高田豊行、臼田大輝、榎田達矢、城石俊彦、榎屋啓志. マウスゲノム多型データベース MoG+の機能向上. 日本遺伝学会第 94 回大会. 札幌. 2022/9/14.
5. 鄒 兆南、大田達郎、三浦史仁、沖 真弥. ChIP-Atlas 2.0: ATAC-seq と WGBS データを統合しエピゲノム制御の全貌に迫る. 第 4 回日本メディカル AI 学会学術集会. 2022/6/11.

〈国際〉

1. Toyoyuki Takada, Hideyuki Miyazawa, Shinya Ayabe, Kentaro Fukuta, Shinji Kondo, Atsushi Toyoda, Masaru Tamura, Atsushi Yoshiki, Kuniya Abe, Yuichi Obata, Toshihiko Shiroishi, Takanori Amano, Hideki Noguchi, Hiroshi Masuya. Exploring genomic variants for unique phenotypes of inbred mouse strains developed in Japan. The 36th International Mammalian Genome Conference. Tsukuba, Japan. 2023/3/31.

#### ④ ポスター発表

〈国内〉

1. 高田豊行、臼田大輝、榎田達矢、城石俊彦、榎屋啓志. リサーチとリソースをつなぐマウスゲノム多型データベース, MoG+. 第 45 回日本分子生物学会年会. 幕張. 2022/12/2.
2. 長谷川哲、粕川雄也、MOIRAI2 - データ処理の自動化の可能性、第 11 回生命医薬情報学連合大会 (IIBMP2022)、大阪、2022/9/13
3. Masaki Morioka, Akira Hasegawa, Atsushi Kondo, Taichi Akase, Nishad Thalath, Imad Abugessaisa, Takeya Kasukawa、医科学研究応用に向けた次期転写開始点情報データベース(refTSS4.0)の開発、第 11 回生命医薬情報学連合大会(IIBMP2022)、大阪、2022/9/13

〈国際〉

1. Takeya Kasukawa, Scott Walker, Akira Hasegawa, Nishad Thalath, Toyoyuki Takada, Hiroshi Masuya, Hideya Kawaji, Shinya Oki, INTRARED: a new integrated transcriptional regulation data platform with cis regulatory elements, trans factors and epigenomic states, the 36th International Mammalian Genome Conference. Tsukuba, Japan. 2023/3/31.

#### (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

#### (5) 受賞・報道等

##### ① 受賞

該当なし

## ② メディア報道

1. プレスリリース「心房細動の遺伝的基盤を解明」理化学研究所 ([https://www.riken.jp/press/2023/20230119\\_1/index.html](https://www.riken.jp/press/2023/20230119_1/index.html))、2023/01/19
2. プレスリリース「ミトコンドリアから細胞核への逆行性シグナリングと標的遺伝子の解明」熊本大学 (<https://www.imeg.kumamoto-u.ac.jp/np136/>)、2022/09/14
3. プレスリリース「細胞の直接変換を誘導する転写因子を高精度で予測する AI 手法を開発」九州工業大学 (<https://www.kyutech.ac.jp/whats-new/press/entry-9065.html>)、2022/04/26
4. プレスリリース「オートファジーを介した細胞老化の制御メカニズムを解明」大阪大学 ([https://www.fbs.osaka-u.ac.jp/ja/research\\_results/papers/detail/1039](https://www.fbs.osaka-u.ac.jp/ja/research_results/papers/detail/1039))、2022/03/02

## ③ その他の成果発表

該当なし

## §5. 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
2022/9/5 ～9/9	チーム内ミーティング(非公開)	高知	5人	DBCLS 主催第13回 国内版バイオハッカソンにおいて、進捗報告と、表現型情報の記述など、今後の進め方について確認した
2022年10 月24日	チーム内ミーティング(非公開)	理化学研究所横浜研究所	7人	研究進捗報告のためのミーティング

### (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
2022年4 月18～20 日	脱 Dry!脱写経! ゲノムインフォマティクス演習	京都大学	22人	企業研究者向けのゲノムインフォマティクス演習
2022年7 月27～29 日	脱 Dry!脱写経! ゲノムインフォマティクス演習	京都大学	21人	企業研究者向けのゲノムインフォマティクス演習
2022年11 月30日	分子生物学会ワークショップ「ゲノム機能・構造データを起点とするゲノム科学の展開」	幕張メッセ	100人	研究交流

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	INTRARED		転写制御の総合的理解の推進を目的に、シスエレメント・トランス因子・エピゲノミクスデータを統合したデータベース。本データベースは fanta.bio と ChIP-Atlas の2つのデータベースで構成される。	<a href="https://www.intrared.org/">https://www.intrared.org/</a>	2022/11/29	新規	データベース等	
2	ChIP-Atlas		SRAとして公開されている全てのChIP-seq, DNase-seq, ATAC-seq, Bisulfite-seqデータを統合したデータベースであり、またそれらを簡便に活用するためのWebサービスである。	<a href="https://chip-atlas.org/">https://chip-atlas.org/</a>	2015/12/1	維持・発展	データベース等	Oki, et al., (2018) EMBO Rep Zou, et al., (2022) Nucleic Acids Res
3	fanta.bio		ヒト・マウス・非ヒト霊長類を対象に、ゲノム中の転写調節に関わるシスエレメント (cis-regulatory element / CRE) について、ゲノム中の位置、様々な細胞種や細胞状態における活性、シスエレメントと関連する変異についてまとめたデータベース	<a href="https://fanta.bio/">https://fanta.bio/</a>	2023/3/27	新規	データベース等	