

2020.12.02

日本分子生物学会

生命科学のデータベース活用法2020

ChIP-Atlas: 公共 ChIP-seq データ を利活用できる



京都大学大学院 医学研究科
創薬医学講座・特定准教授

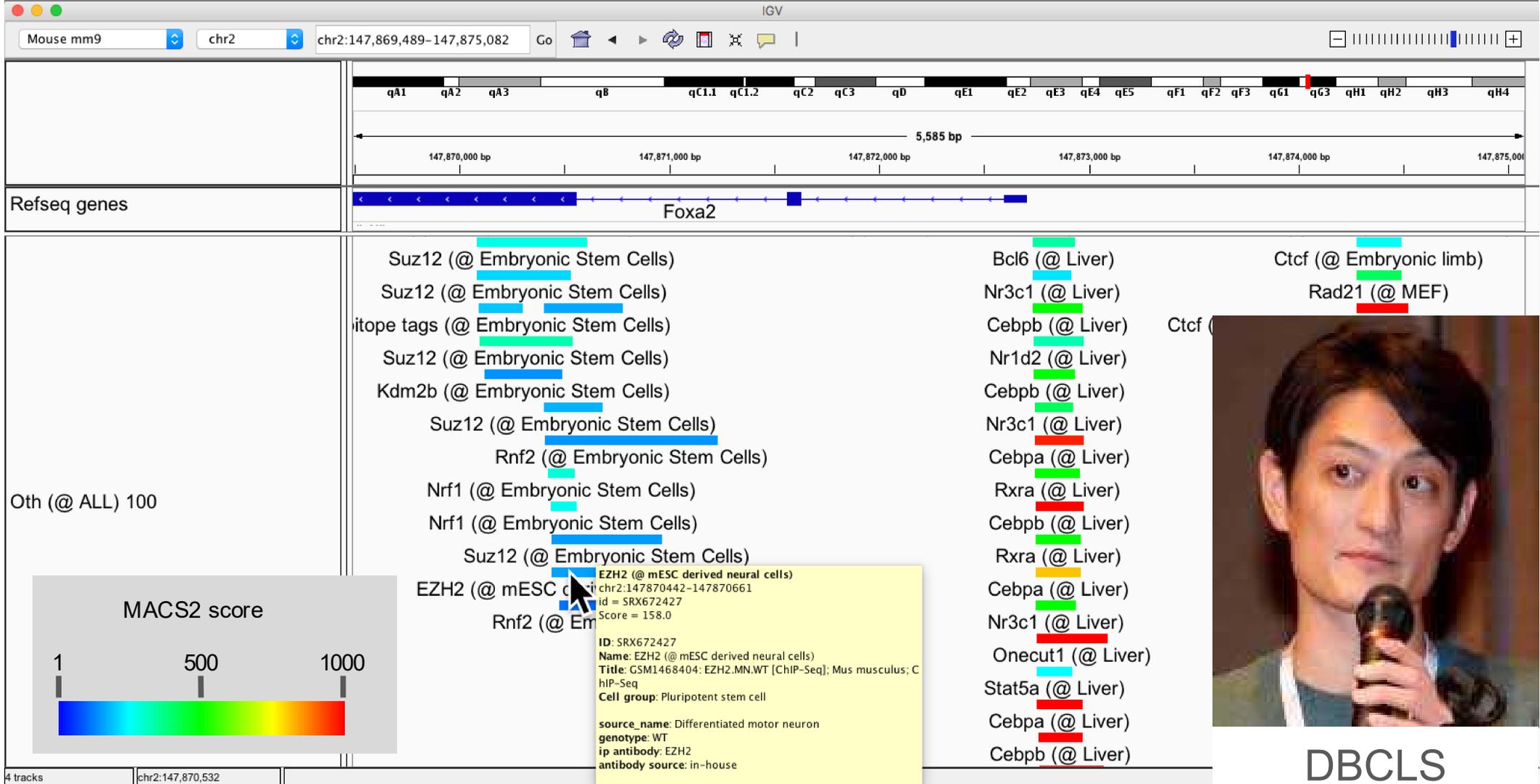
沖 真弥 (Oki, Shinya)

Licensed under a Creative Commons 表示4.0国際ライセンス
©2020沖 真弥(京都大学大学院 医学研究科)



ChIP-Atlas とは

既報 ChIP-seq データを統合的に利活用できる



DBCLS
大田達郎

chip-atlas.org

ChIP-Atlas Peak Browser Target Genes Colocalization *in silico* ChIP Documentation Publications Find an experiment ▾

ChIP-Atlas

ChIP-Atlas is an integrative and comprehensive database for visualizing and making use of public ChIP-seq data. ChIP-Atlas covers almost all public ChIP-seq data submitted to the SRA (Sequence Read Archives) in NCBI, DDBJ, or ENA, and is based on over 69,000 experiments.

[Watch movie introduction](#)

The four main features of ChIP-Atlas are:

<h3>Peak Browser</h3> <p>graphically visualizes protein binding on given genomic loci with genome browser (IGV).</p> <p>Watch Movie</p>	<h3>Target Genes</h3> <p>predicts target genes bound by given transcription factors.</p> <p>Watch Movie</p>	<h3>Colocalization</h3> <p>predicts partner proteins colocalizing with given transcription factors.</p> <p>Watch Movie</p>	<h3>Enrichment Analysis</h3> <p>predicts proteins bound to given genomic loci and genes.</p> <p>Watch Movie</p>
---	---	--	---

 **DBCLS**
Database Center
for Life Science

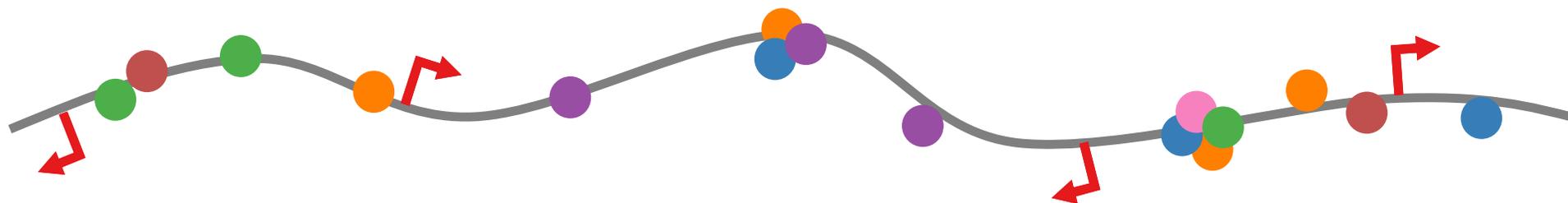
 **KYUSHU UNIVERSITY**

THIS WORK IS SUPPORTED BY NIG SUPERCOMPUTER SYSTEM AND NATIONAL BIOSCIENCE DATABASE CENTER.

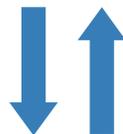
NEED HELP? CREATE AN ISSUE ON GITHUB OR CONTACT US

ChIP-Atlas でできること

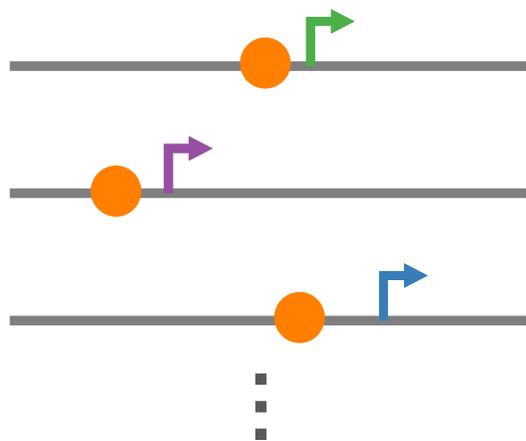
Peak Browser



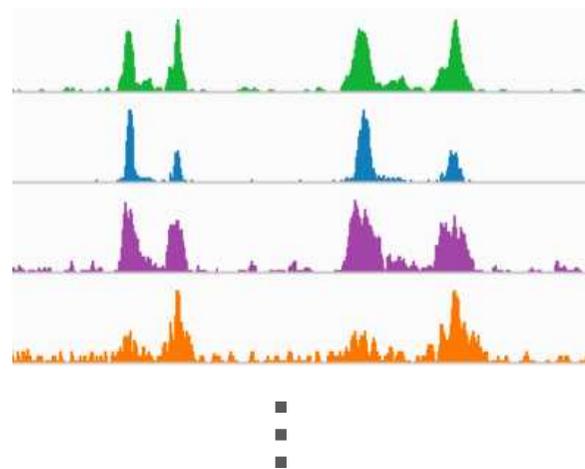
Target Genes



Enrichment Analysis



Colocalization



すべてのピークコールデータをまとめて閲覧できる

Mouse mm9 | chr2 | chr2:147,869,489-147,875,082 | Go

qA1 qA2 qA3 qB qC1.1 qC1.2 qC2 qC3 qD qE1 qE2 qE3 qE4 qE5 qF1 qF2 qF3 qG1 qG3 qH1 qH2 qH3 qH4

147,870,000 bp 147,871,000 bp 147,872,000 bp 147,873,000 bp 147,874,000 bp 147,875,000 bp

5,585 bp

Refseq genes

Foxa2

Oth (@ ALL) 100

MACS2 score

1 500 1000

EZH2 (@ mESC derived neural cells)
chr2:147870442-147870661
id = SRX672427
Score = 158.0

ID: SRX672427
Name: EZH2 (@ mESC derived neural cells)
Title: GSM1468404: EZH2.MN.WT [ChIP-Seq]; Mus musculus; ChIP-Seq
Cell group: Pluripotent stem cell

source_name: Differentiated motor neuron
genotype: WT
ip antibody: EZH2
antibody source: in-house

Suz12 (@ Embryonic Stem Cells)
Suz12 (@ Embryonic Stem Cells)
Nfya1 (@ Embryonic Stem Cells)
Suz12 (@ Embryonic Stem Cells)
Kdm2b (@ Embryonic Stem Cells)
Suz12 (@ Embryonic Stem Cells)
Rnf2 (@ Embryonic Stem Cells)
Nrf1 (@ Embryonic Stem Cells)
Nrf1 (@ Embryonic Stem Cells)
Suz12 (@ Embryonic Stem Cells)
EZH2 (@ mESC derived neural cells)
Rnf2 (@ Embryonic Stem Cells)

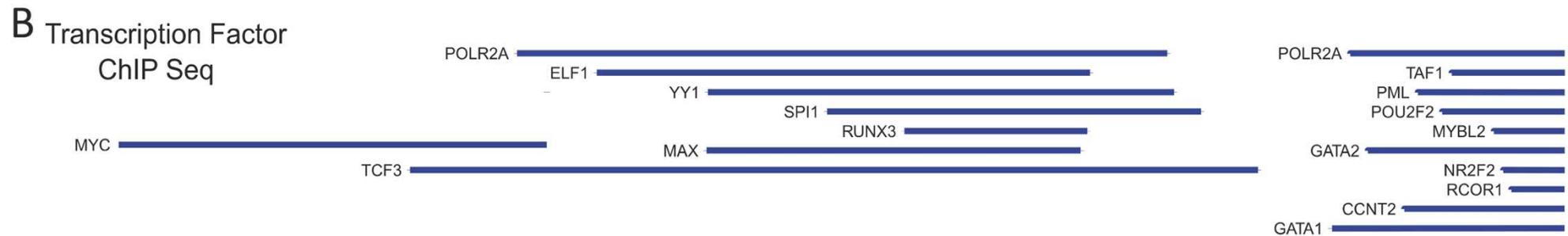
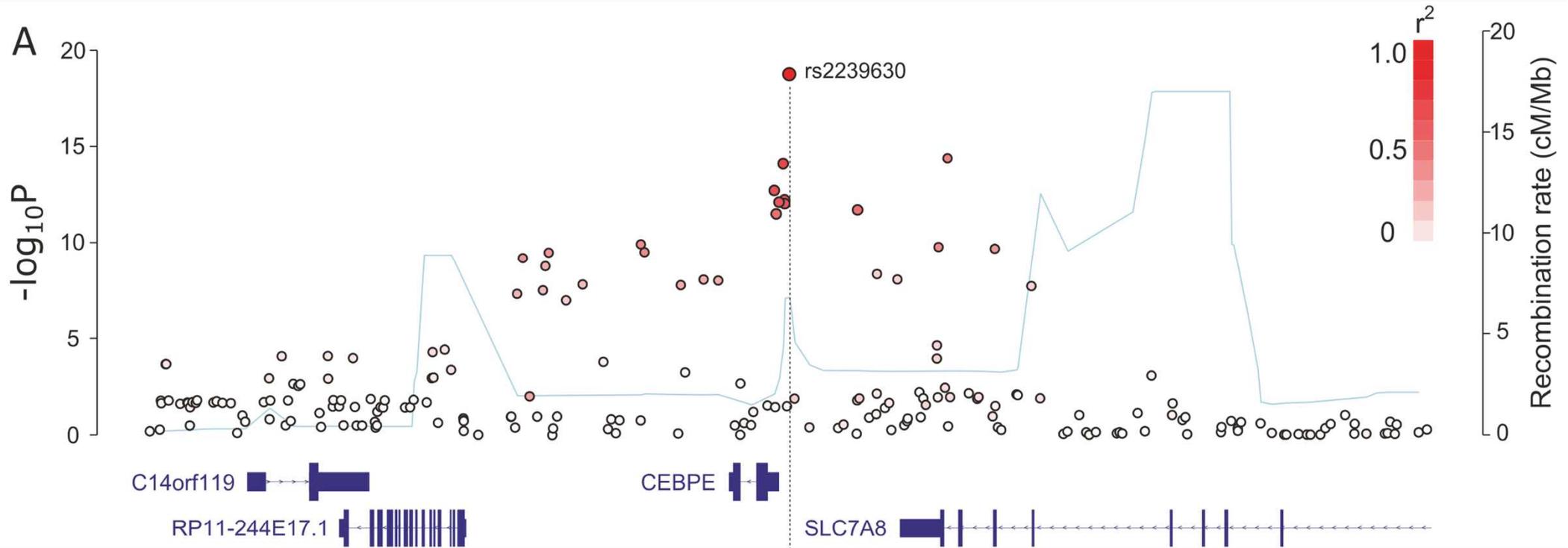
Bcl6 (@ Liver)
Nr3c1 (@ Liver)
Cebpb (@ Liver)
Nr1d2 (@ Liver)
Cebpb (@ Liver)
Nr3c1 (@ Liver)
Cebpa (@ Liver)
Rxra (@ Liver)
Cebpb (@ Liver)
Rxra (@ Liver)
Cebpa (@ Liver)
Nr3c1 (@ Liver)
Onecut1 (@ Liver)
Stat5a (@ Liver)
Cebpa (@ Liver)
Cebpb (@ Liver)

Ctcf (@ Embryonic limb)
Rad21 (@ MEF)
Ctcf (@ mESC derived haematopoietic pro
Gata2 (@ Uterus)
Ctcf (@ MEF)
Ctcf (@ MEF)
Ctcf (@ Liver)
Ctcf (@ Heart)
Ctcf (@ Heart)
Ctcf (@ Embryonic Stem Cells)
Ctcf (@ MEF)
Smc1a (@ MEF)
Ctcf (@ Embryonic Stem Cells)
Stag2 (@ Embryonic Stem Cells)
Smc1a (@ mESC derived neural cell
Srf (@ C3H/10T1/2)

4 tracks | chr2:147,870,532 | 141M of 494M

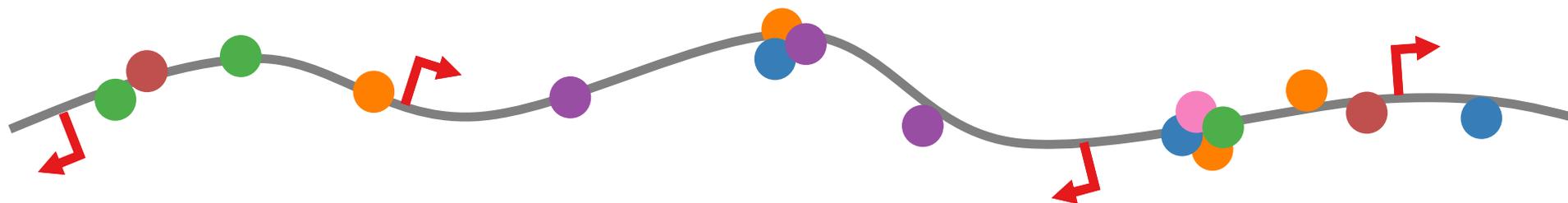
Peak Browser

急性リンパ性白血病の GWAS

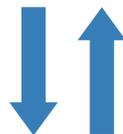


ChIP-Atlas でできること

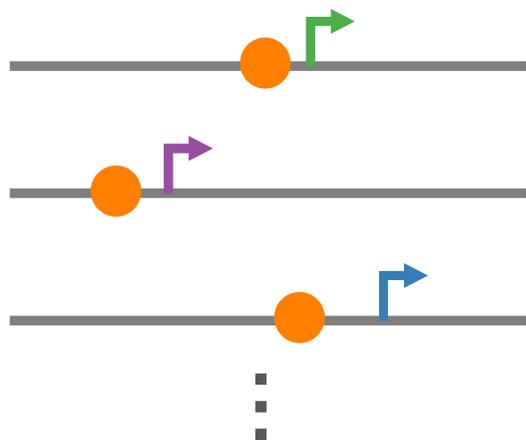
Peak Browser



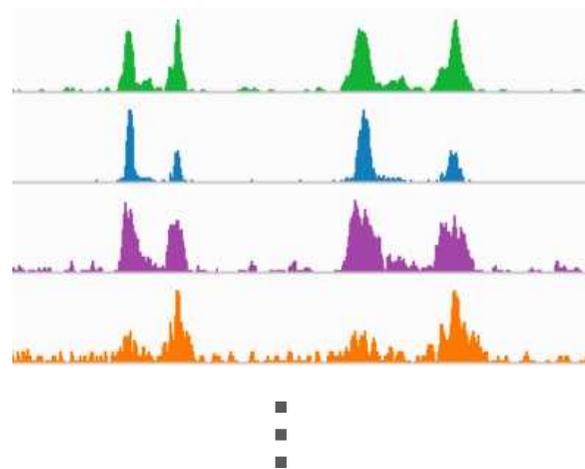
Target Genes



Enrichment Analysis



Colocalization



GATA6 の標的遺伝子は？

ChIP-Atlas

Peak Browser

Target Genes

Colocalization

Enrichment Analysis

Documentation

Publications

ChIP-Atlas - Target Genes

Predict potential target genes of TFs.

H.

GATA6

R. norvegicus

D. melanogaster

C. elegans

S. cerevisiae

1. Choose Antigen

GATA6

FXR1
FXR2
GABPA
GATA1
GATA2
GATA3
GATA4
GATA6

2. Choose Distance from TSS

±1k

±5k

±10k

TSS ± 10 kb

クリック

View Potential Target Genes

Download (TSV)

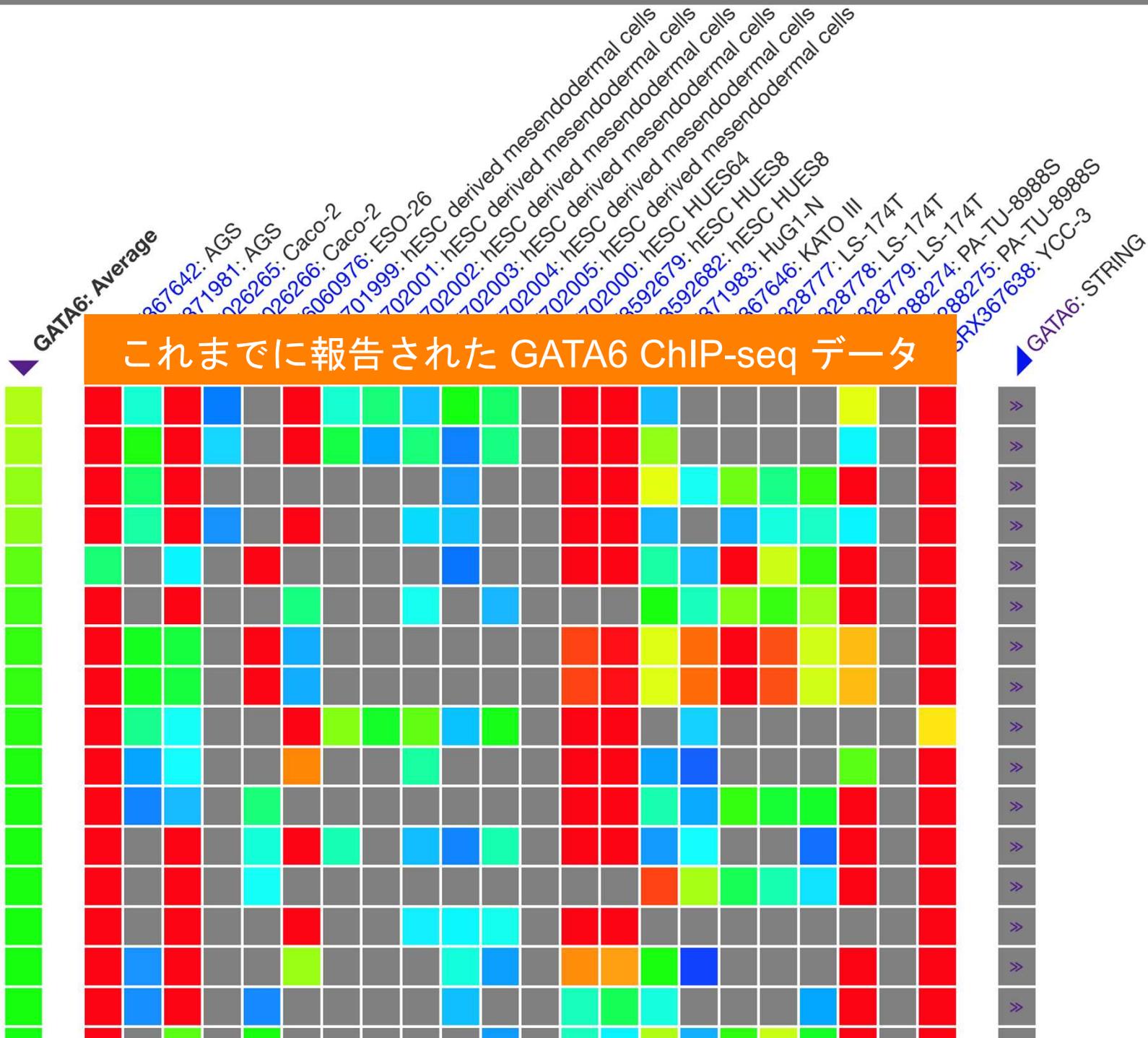
GATA6 の標的遺伝子は？

TSS ± 10 kb に結合する遺伝子

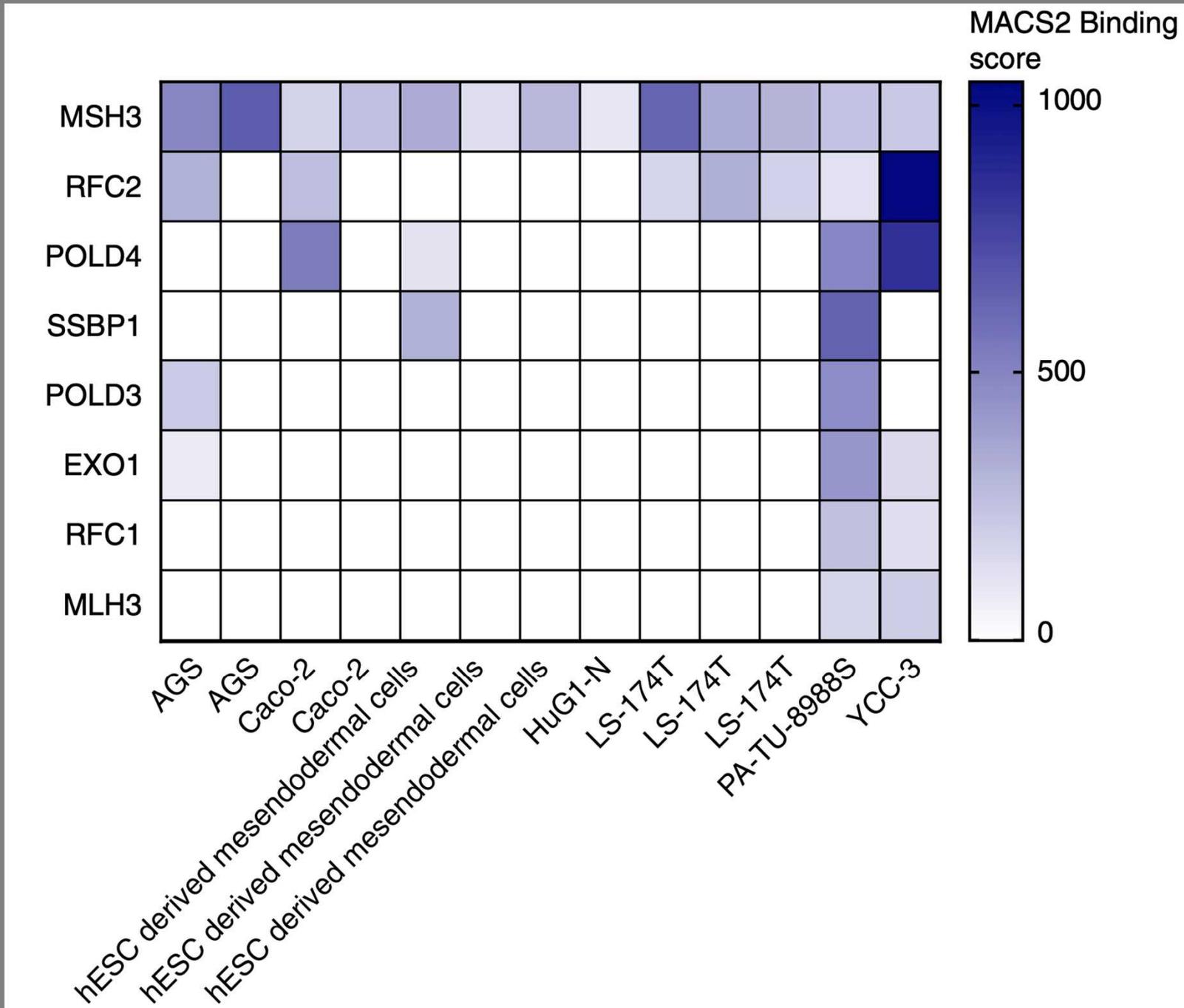
GATA6's
Target genes

RGS5
KDM5B
RBBP8NL
SIX1
BAHCC1
SPINK4
WDR74
STX5
IL23R
PLAUR
CDC42EP1
RASGRP3
BTNL8
BAMBI
TAGLN2
MICAL2

これまでに報告された GATA6 ChIP-seq データ



GATA6 の標的遺伝子は？



Enrichment Analysis

H. sapiens

M. musculus

R. norvegicus

D. melanogaster

C. elegans

S. cerevisiae

1. Antigen Class

All antigens (31081)
DNase-seq (1440)
Histone (7674)
RNA polymerase (1011)
TFs and others (7353)
Input control (3388)
Unclassified (6329)
No description (3886)

2. Cell type Class

All cell types (31081)
Adipocyte (275)
Blood (8600)
Bone (466)
Breast (2889)
Cardiovascular (785)
Digestive tract (2349)
Epidermis (964)

3. Threshold for Significance

50
100
200
500

4. Select your data

- Genomic regions (BED) or sequence motif ⓘ
 Gene list (Gene symbols) ⓘ

ADI1
AGO1
AHCYL2
AJAP1
APOBEC3C
APOBEC3D
ARHGEF38
ATPAF2

HNF4s

FOXAs

肝臓で特異的に
発現する遺伝子

5. Select dataset to be compared

- Refseq coding genes (excluding user data) ⓘ
 Gene list (Gene symbols) ⓘ

AGBL5
ALOXE3
AMZ2
ANG
APBB3
ARID2
ASF1A
ATP5O

その他の
遺伝子

6. Describe datasets

User data title ⓘ

My data

Compared data title ⓘ

Control

Project title ⓘ

My project

Distance range from TSS ⓘ

- 5000 bp ≤ TSS ≤ + 5000 bp

submit

Estimated run time: 4 mins

Enrichment Analysis

H. sapiens M. musculus R. norvegicus D. melanogaster C. elegans S. cerevisiae

1. Antigen Class

- All antigens (31081)
- DNase-seq (1440)
- Histone (7674)
- RNA polymerase (1011)
- TFs and others (7353)
- Input control (3388)
- Unclassified (6329)
- No description (3886)

2. Cell type Class

- All cell types (31081)
- Adipocyte (275)
- Blood (8600)
- Bone (466)
- Breast (2889)
- Cardiovascular (785)
- Digestive tract (2349)
- Epidermis (964)

3. Threshold for Significance

50
100
200
500

4. Select your data

Genomic regions (BED) or sequence motif ⓘ
 Gene list (Gene symbols) ⓘ

- ADI1
- AGO1
- AHCYL2
- AJAP1
- APOBEC3C
- APOBEC3D
- ARHGEF38
- ATPAF2

HNF4s
FOXAs

肝臓で特異的に

5. Select dataset to be compared

Refseq coding genes (excluding user data) ⓘ
 Gene list (Gene symbols) ⓘ

- AGBL5
- ALOXE3
- AMZ2
- ANG
- APBB3
- ARID2
- ASF1A
- ATP5O

その他の

6. Describe datasets

User data title ⓘ
My data

Compared data title ⓘ
Control

Project title ⓘ
My project

Distance range from TSS ⓘ
- 5000 bp ≤ TSS ≤ + 5000 bp

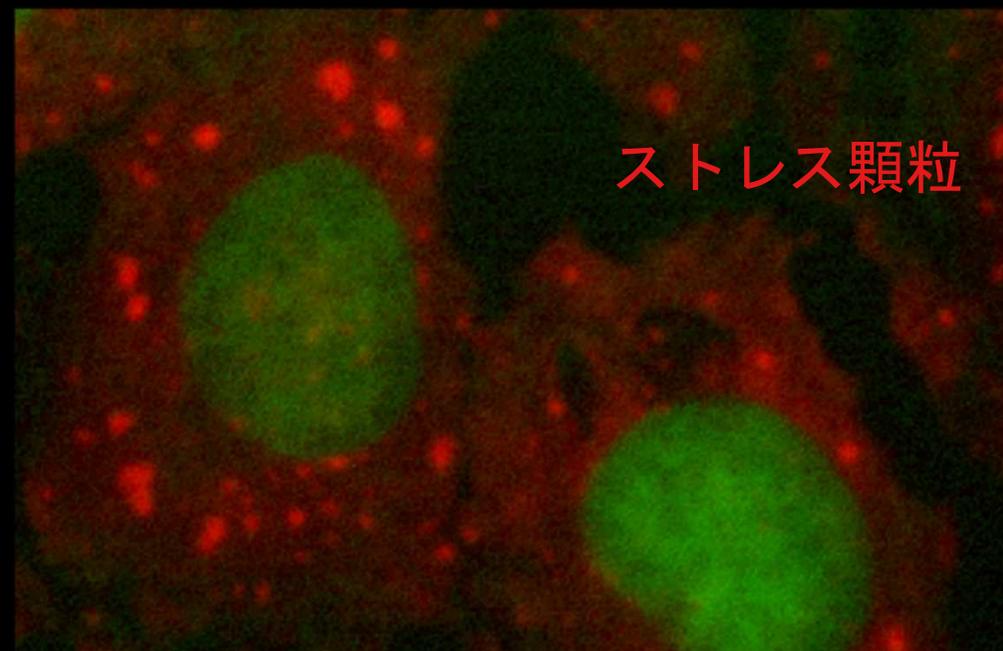
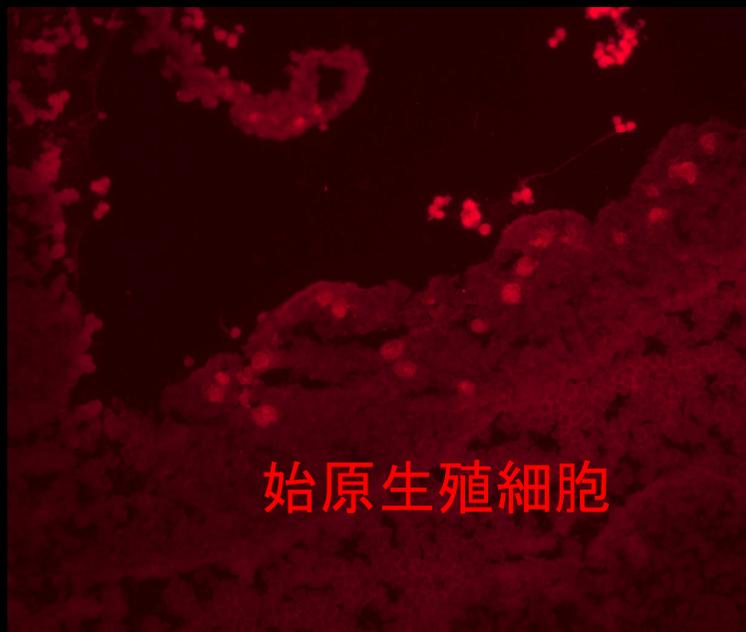
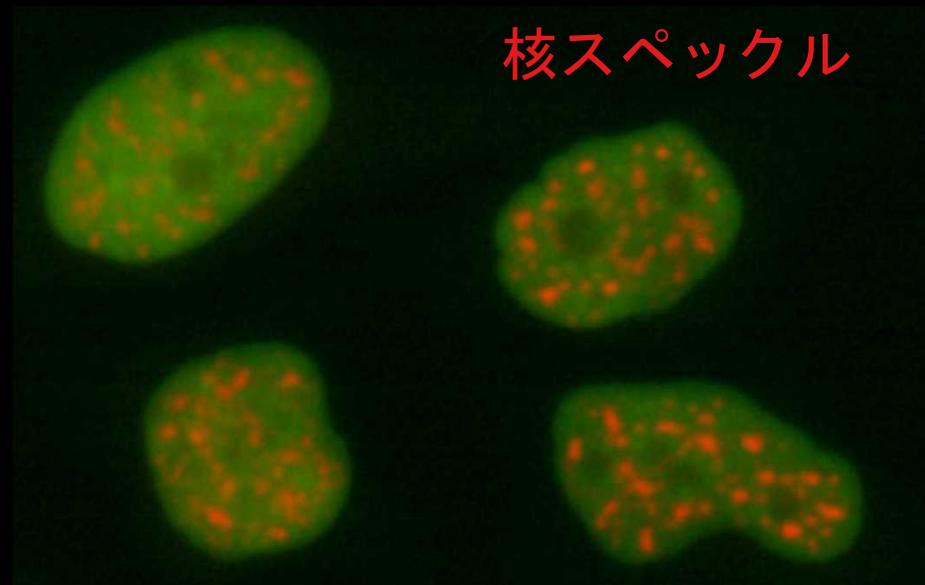
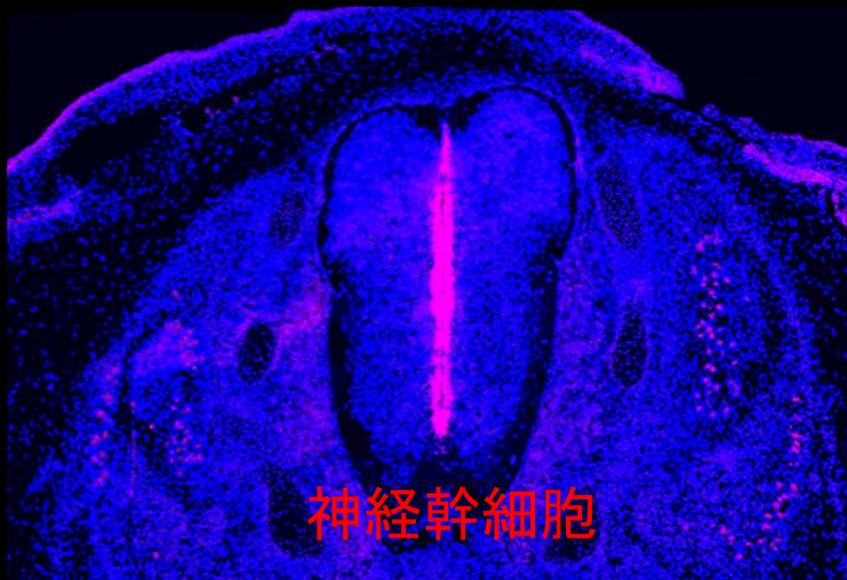
[2P-0248]

パイオニア転写因子を考慮したデータ駆動型ダイレクトリプログラミング
江口 凌平 (九工大・山西芳裕 lab)



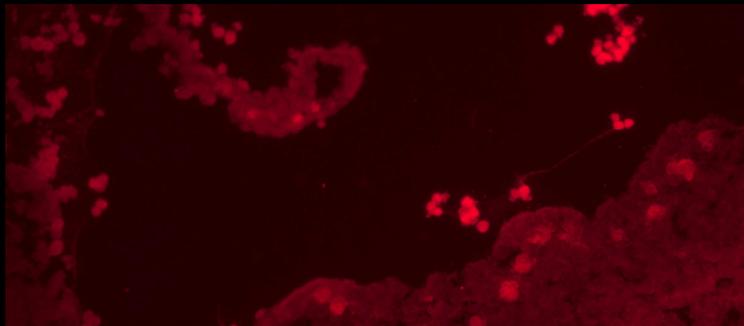
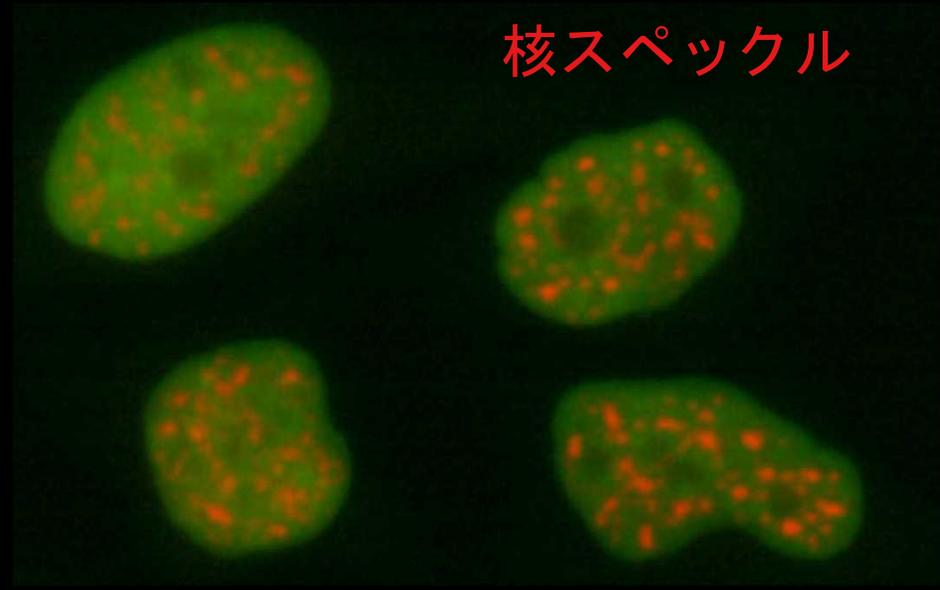
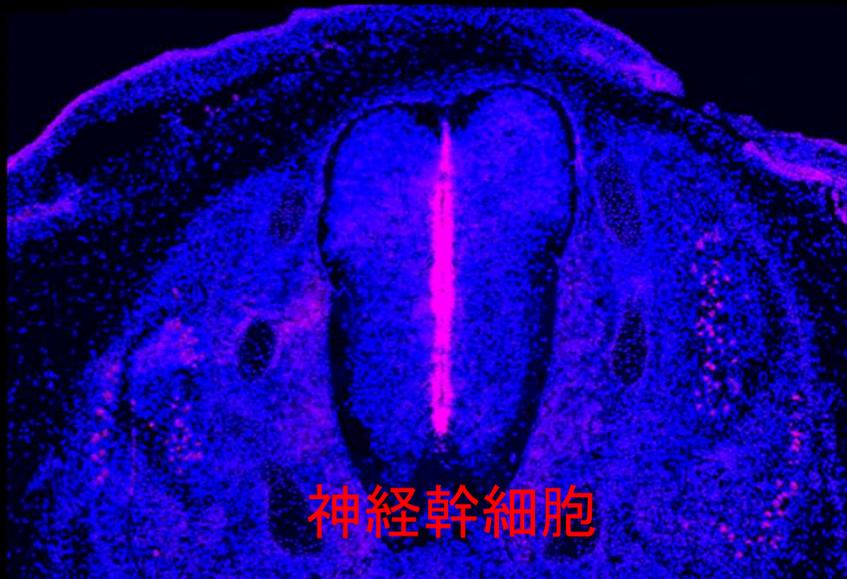
山西 芳裕
九工大

Photo-isolation chemistry



微小領域のトランスクリプトーム解析ができる

Photo-isolation chemistry



[2PS-03-3] 16:24~16:51
空間的な遺伝子発現制御のしくみを探る



本田 瑞季 助教



大川 恭行 九大

[3AW-05-6] 10:45~11:05

Photo-Isolation Chemistryによる高解像度の空間的トランスクリプトーム技術

ATAC-seq の解析

Open chromatin 領域に、何が結合するか？

Analyze your data with public ChIP-seq data.

H. sapiens

M. musculus

D. melanogaster

C. elegans

S. cerevisiae

1. Antigen Class

All antigens (16138)
DNase-seq (1024)
Histone (3824)
RNA polymerase (629)
TFs and others (5088)
Input control (1956)
Unclassified (596)
No description (3021)

2. Cell type Class

All cell types (16138)
Adipocyte (120)
Blood (4559)
Bone (200)
Breast (1712)
Cardiovascular (498)
Digestive tract (1205)
Epidermis (431)

3. Threshold for Significance

50
100
200
500

4. Select your data

- Genomic regions (BED) or sequence motif ⓘ
 Gene list (Gene symbols) ⓘ

```
chr1 100128315 100128440
chr1 103190456 103190612
chr1 107541234 107541357
chr1 108325177 108325403
chr1 110412483 110412583
chr1 111120868 111121001
chr1 111693652 111693799
chr1 112421367 112421483
```

ATAC-seq
(処理した細胞)

5. Select dataset to be compared

- Random permutation of user data ⓘ
 BED or sequence motif ⓘ

```
chr1 10488202 10488365
chr1 107973343 107973565
chr1 108293100 108293277
chr1 109370825 109371023
chr1 110319562 110319739
chr1 112298232 112298429
chr1 113351913 113352118
chr1 116524095 116524271
```

ATAC-seq
(Control 細胞)

6. Describe datasets

User data title ⓘ

My data

Compared data title ⓘ

Control

Project title ⓘ

My project

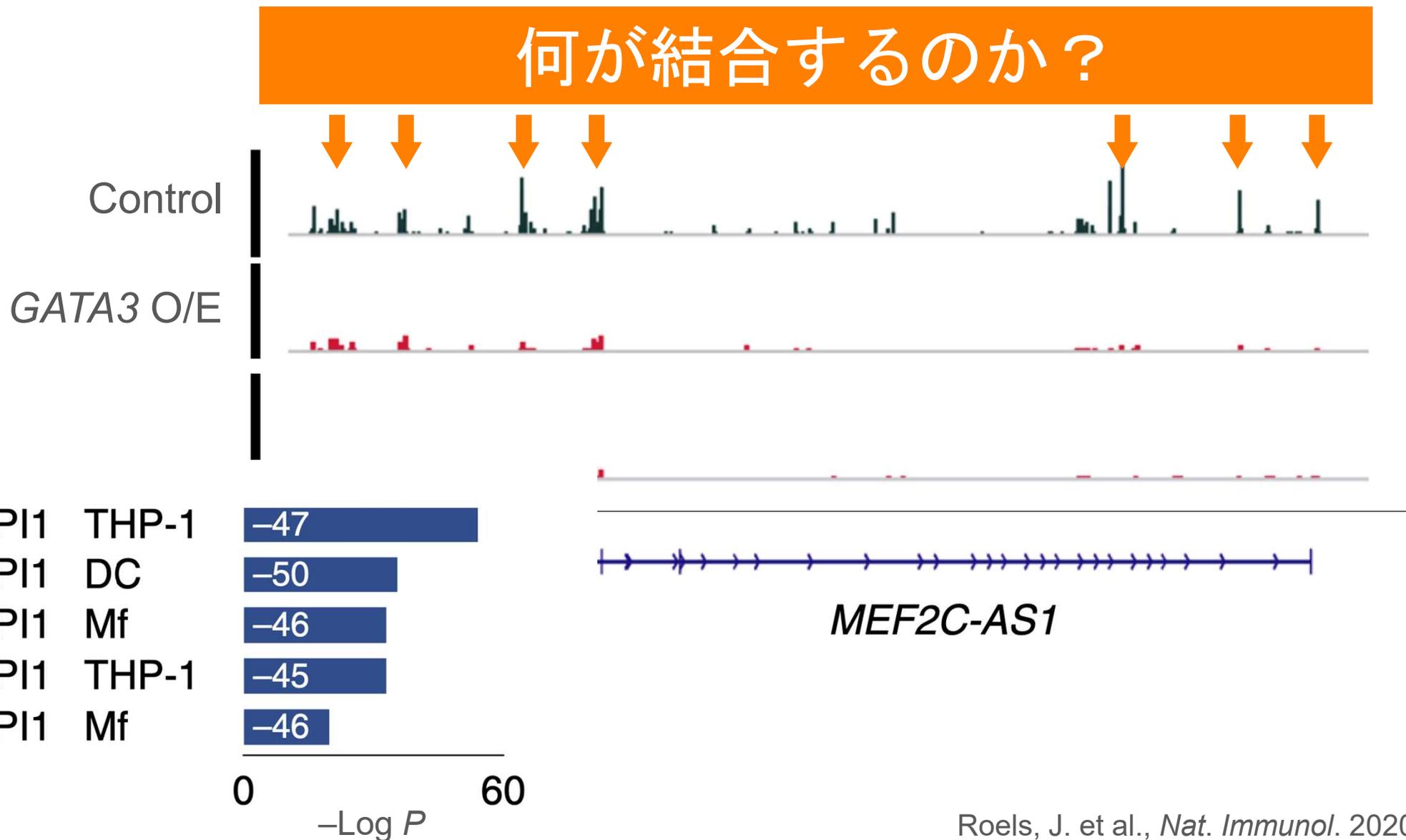
submit

Estimated run time: 2 mins

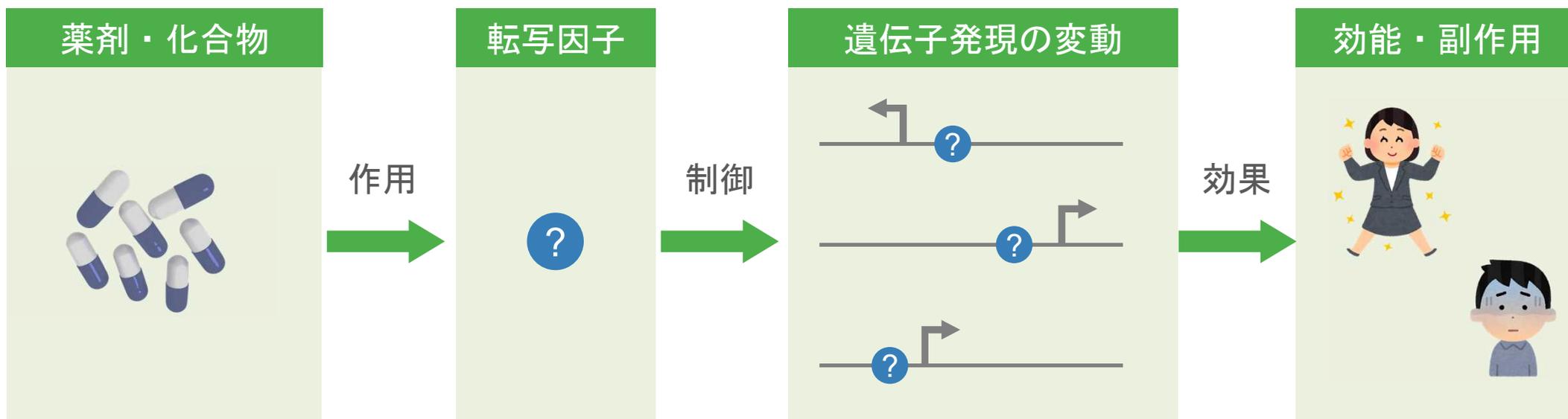
ATAC-seq の解析

血球前駆細胞 → GATA3 強制発現 → ATAC-seq

何が結合するのか？



薬剤標的の予測



Estradiol



ESR1

Testosterone



AR

Leflunomide



RELA TLE3

効能：リウマチ・がん

※ 炎症 ※ 増殖抑制

Valproic acid



CHD8

副作用：自閉症

※ 自閉症関連因子

[2PS-09-5] 岩田 通夫、沖 真弥、山西 芳裕

多層オミクス解析による遺伝子発現機構の
ディジーズーム解析と治療薬探索



山西 芳裕
九工大

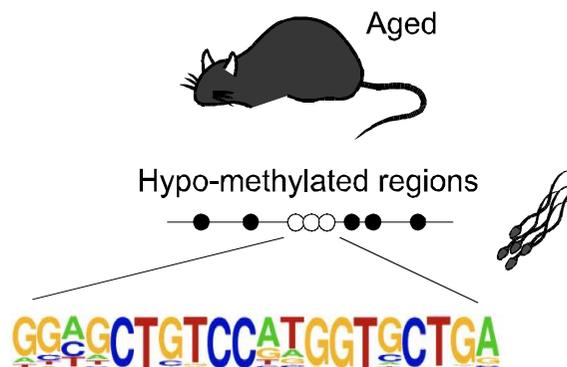
その他の共同研究



Miyamoto, Yoichi
宮本洋一
(医薬基盤研)

Spermatogenesis

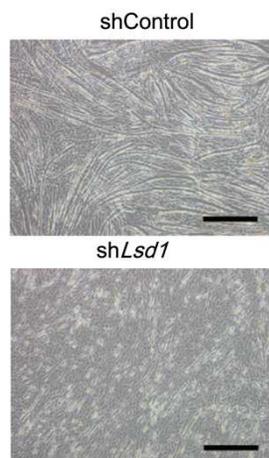
Miyamoto, et al., *FASEB J.* 2020



Osumi, Noriko
大隅典子
(東北大学)

ASD by aged sperm

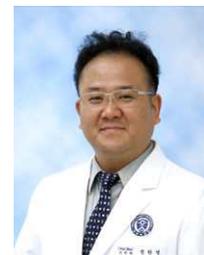
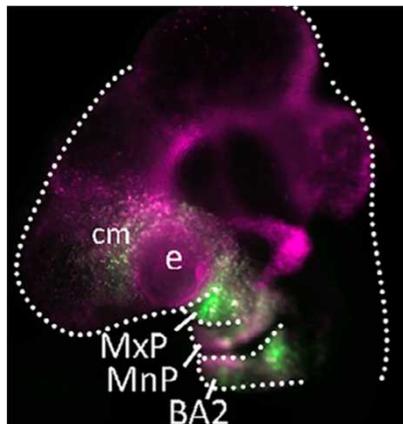
Yoshizaki, et al., *EMBO Rep.* in press



Hino, Shinjiro
日野信次朗
(熊本大学)

Myogenesis

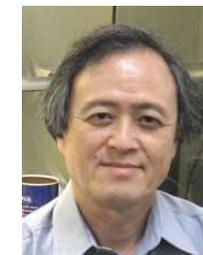
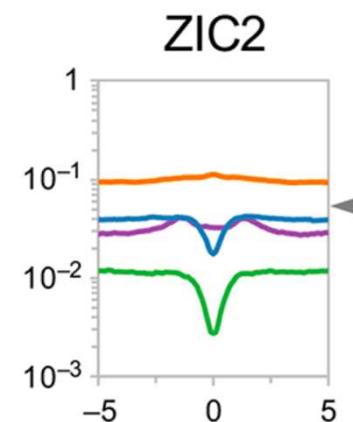
Anan, et al., *NAR.* 2018



Han-Sung Jung
(Yonsei Univ)

Neural crest

Wu, et al., *Dev Dyn.* 2019

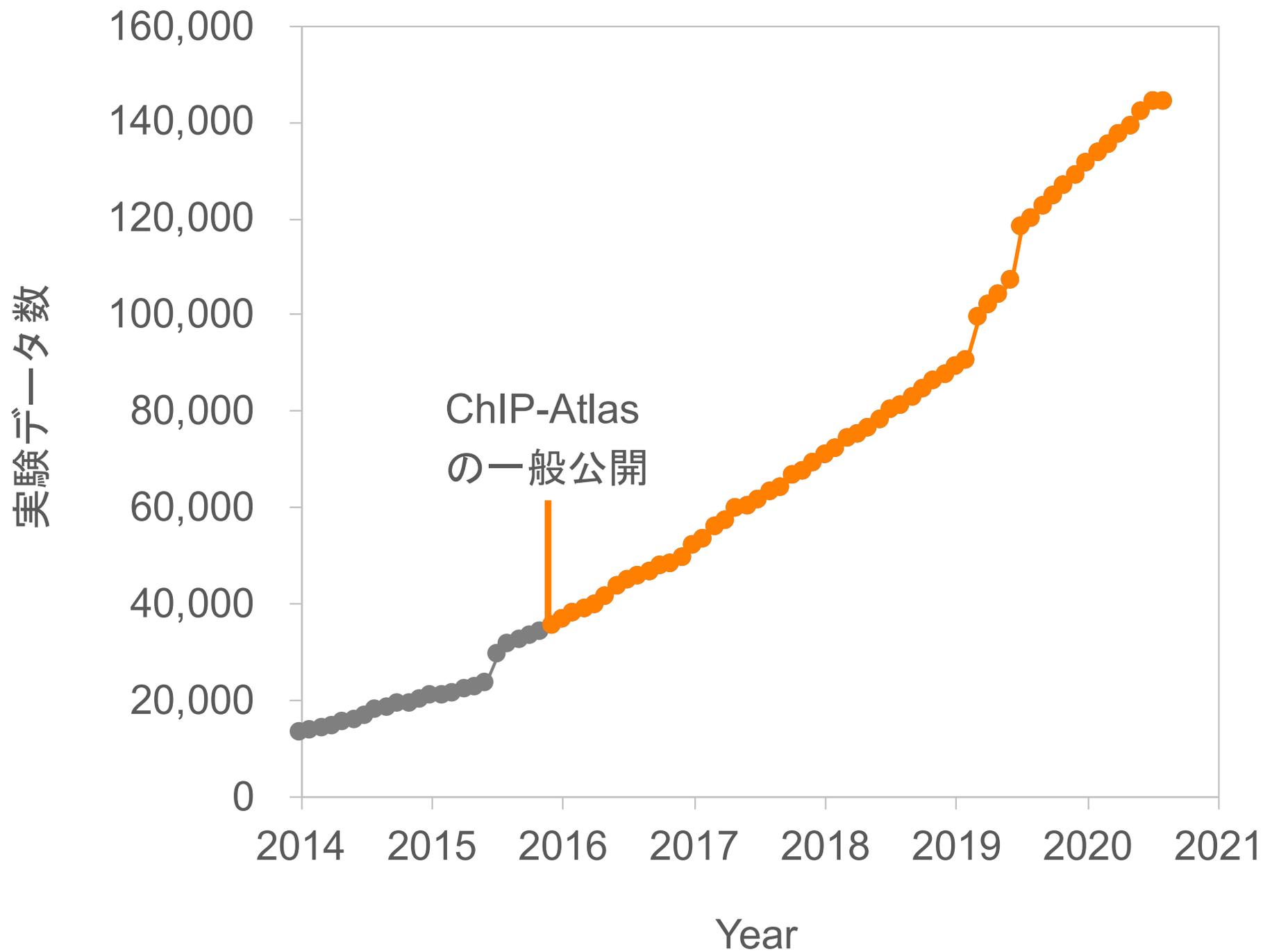


Kondoh, Hisato
近藤寿人
(京産大)

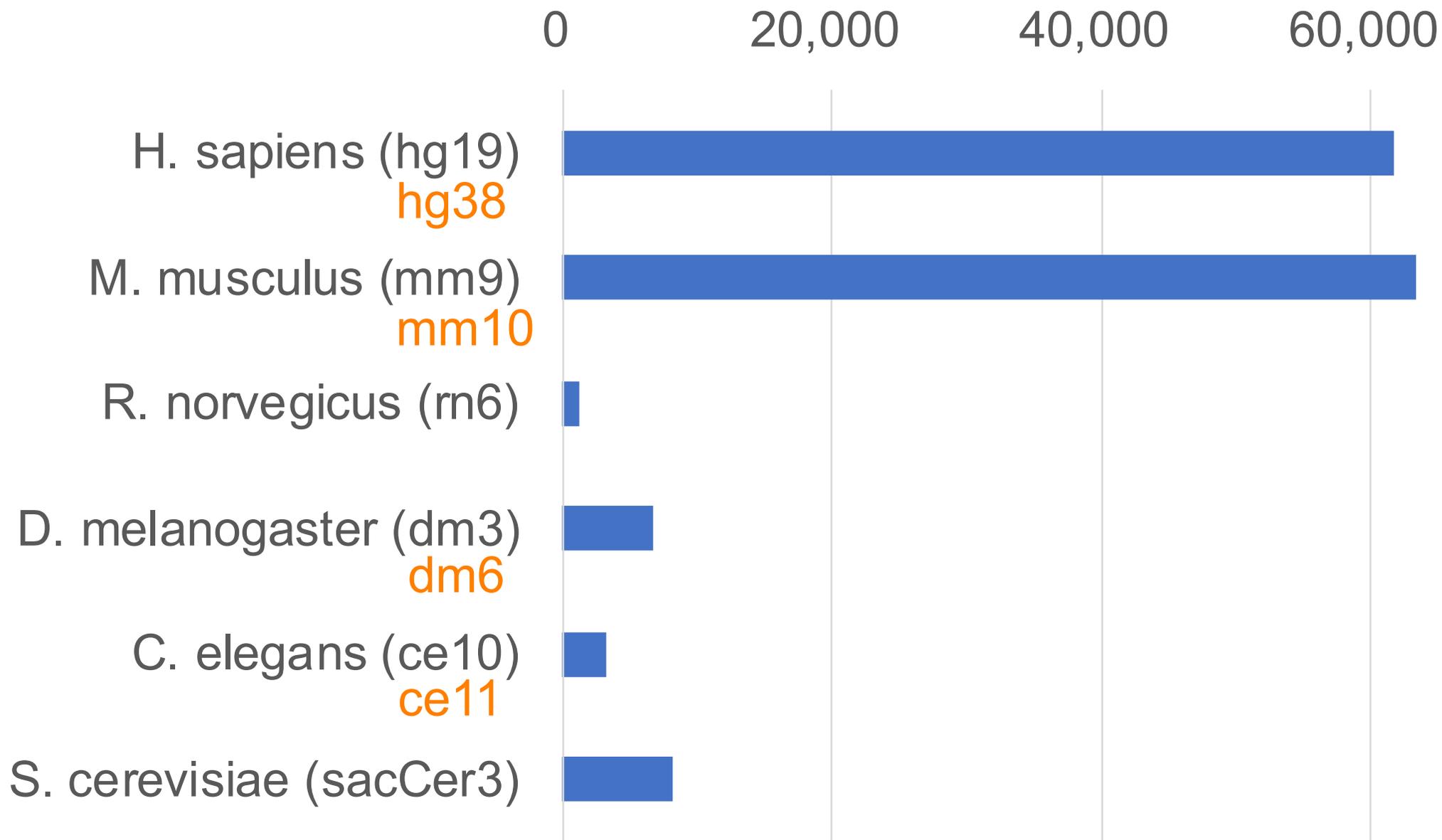
Epiblast stemness

Matsuda, et al., *Development.* 2017

データ更新



データ数



chip-atlas.org

ChIP-Atlas Peak Browser Target Genes Colocalization *in silico* ChIP Documentation Publications Find an experiment ▾

ChIP-Atlas

ChIP-Atlas is an integrative and comprehensive database for visualizing and making use of public ChIP-seq data. ChIP-Atlas covers almost all public ChIP-seq data submitted to the SRA (Sequence Read Archives) in NCBI, DDBJ, or ENA, and is based on over 69,000 experiments.

[Watch movie introduction](#)

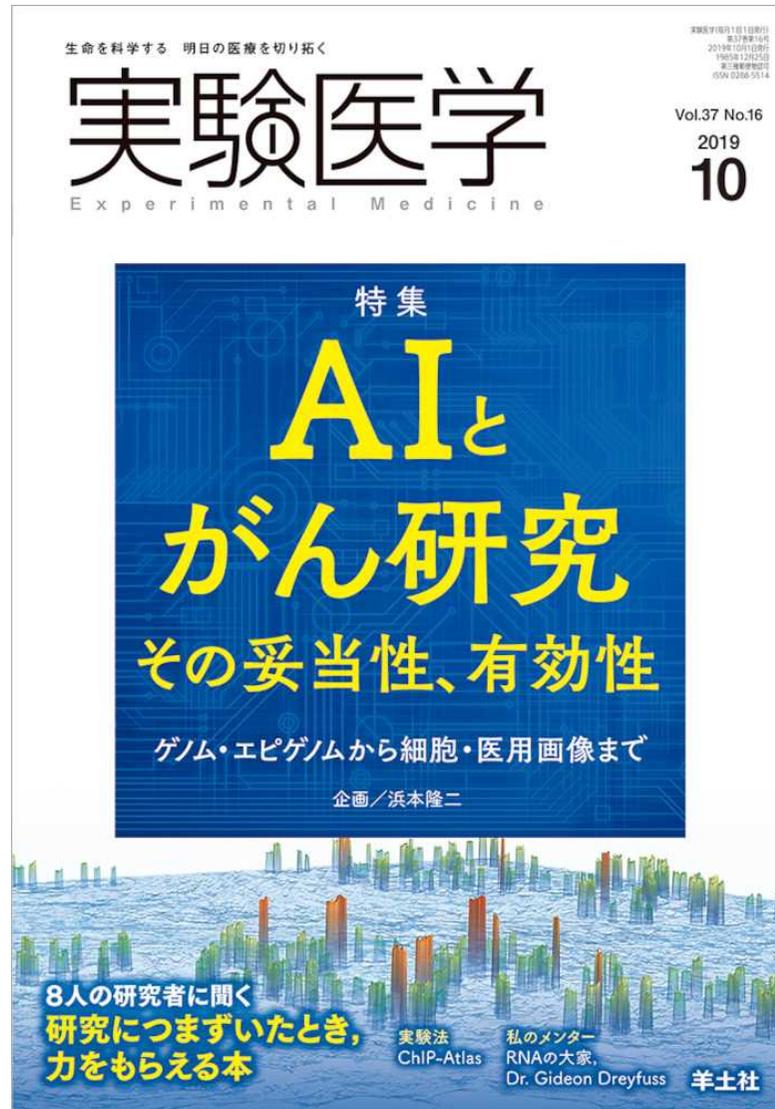
The four main features of ChIP-Atlas are:

- Peak Browser**
graphically visualizes protein binding on given genomic loci with genome browser (IGV).
[Watch Movie](#)
- Target Genes**
predicts target genes bound by given transcription factors.
[Watch Movie](#)
- Colocalization**
predicts partner proteins colocalizing with given transcription factors.
[Watch Movie](#)
- Enrichment Analysis**
predicts proteins bound to given genomic loci and genes.
[Watch Movie](#)

 **DBCLS**
Database Center
for Life Science

 **KYUSHU UNIVERSITY**

THIS WORK IS SUPPORTED BY NIG SUPERCOMPUTER SYSTEM AND NATIONAL BIOSCIENCE DATABASE CENTER.
NEED HELP? CREATE AN ISSUE ON GITHUB OR CONTACT US



2019年10月号 Vol.37 No.16

ChIP-Atlas : 既報のChIP-seqデータをフル活用する
ためのウェブサービス

沖 真弥, 大田達郎

謝辞

ChIP-Atlas

DBCLS 大田達郎

RIKEN 塩井 剛

東大 仲木 竜

RIKEN 川路 英哉

DDBJ 小笠原 理

DDBJ 奥田 喜広

NBDC 畠中 秀樹

産総研 瀬々 潤

九大 目野 主税

ChIP-Atlas 共同研究

医薬基盤研 宮本 洋一

東北大 大隅 典子

熊本大 大隅 典子

東北大 日野信次郎

Yonsei Univ HS Jung

PIC

京大 本田 瑞季

京大 木村 龍一

京大 菅原 由美

九大 大川 恭行

九大 原田 哲仁

九大 前原 一満

九大 田中 かおり

薬効解析

京大 鄒 兆南

九工大 山西 芳裕

九工大 岩田 通夫

疾患 SNP

徳島大 竹本 龍也

徳島大 沢津橋 俊

Grant-in-Aid



H29～ 統合化推進プログラム 代表
H29～ さきがけ「多細胞」 代表



H30～ 基盤研究 B (特設分野) 代表
H31～ 基盤研究 B (医化学) 代表



H29～
分担



H29～ 老化メカニズムの解明・
制御プロジェクト

老化研究推進支援拠点 分担

Special Thanks



京都大学大学院医学研究科
創薬医学講座
Department of Drug Discovery Medicine

大学院生を募集しています



鄒 兆南
DC1

木村 龍一
特定助教

沖 真弥
特定准教授

本田 瑞季
特定助教

菅原 由美
教務補佐員

Wet/Dry どちらも歓迎します！



京都大学大学院医学研究科
創薬医学講座
Department of Drug Discovery Medicine