

第41回日本分子生物学会年会
生命科学のデータベース活用法2018

公共 ChIP-seq データを フル活用できるChIP-Atlas



九州大学大学院・医学研究院・発生再生医学分野

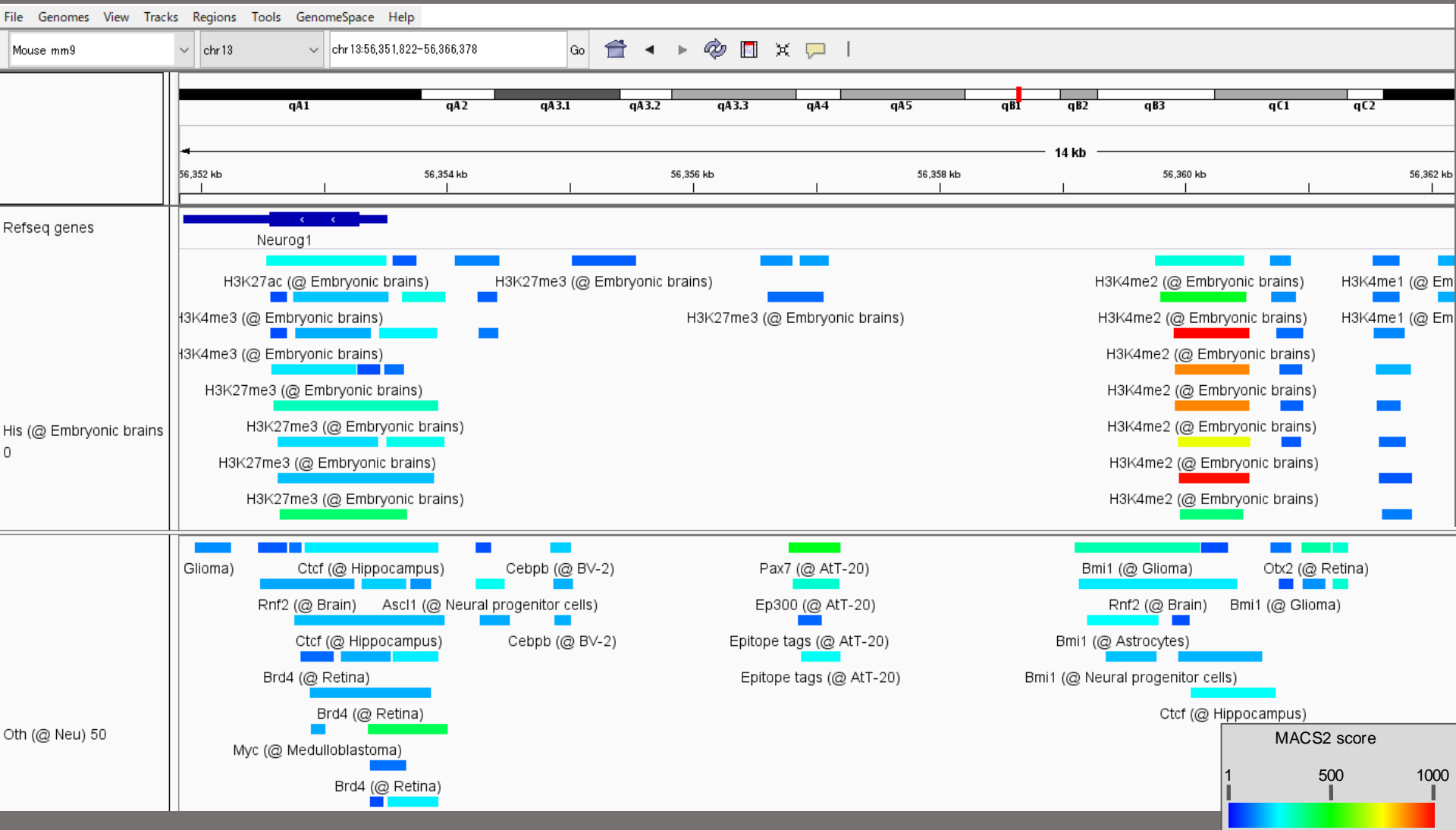
○ 本田 瑞季 , 沖 真弥



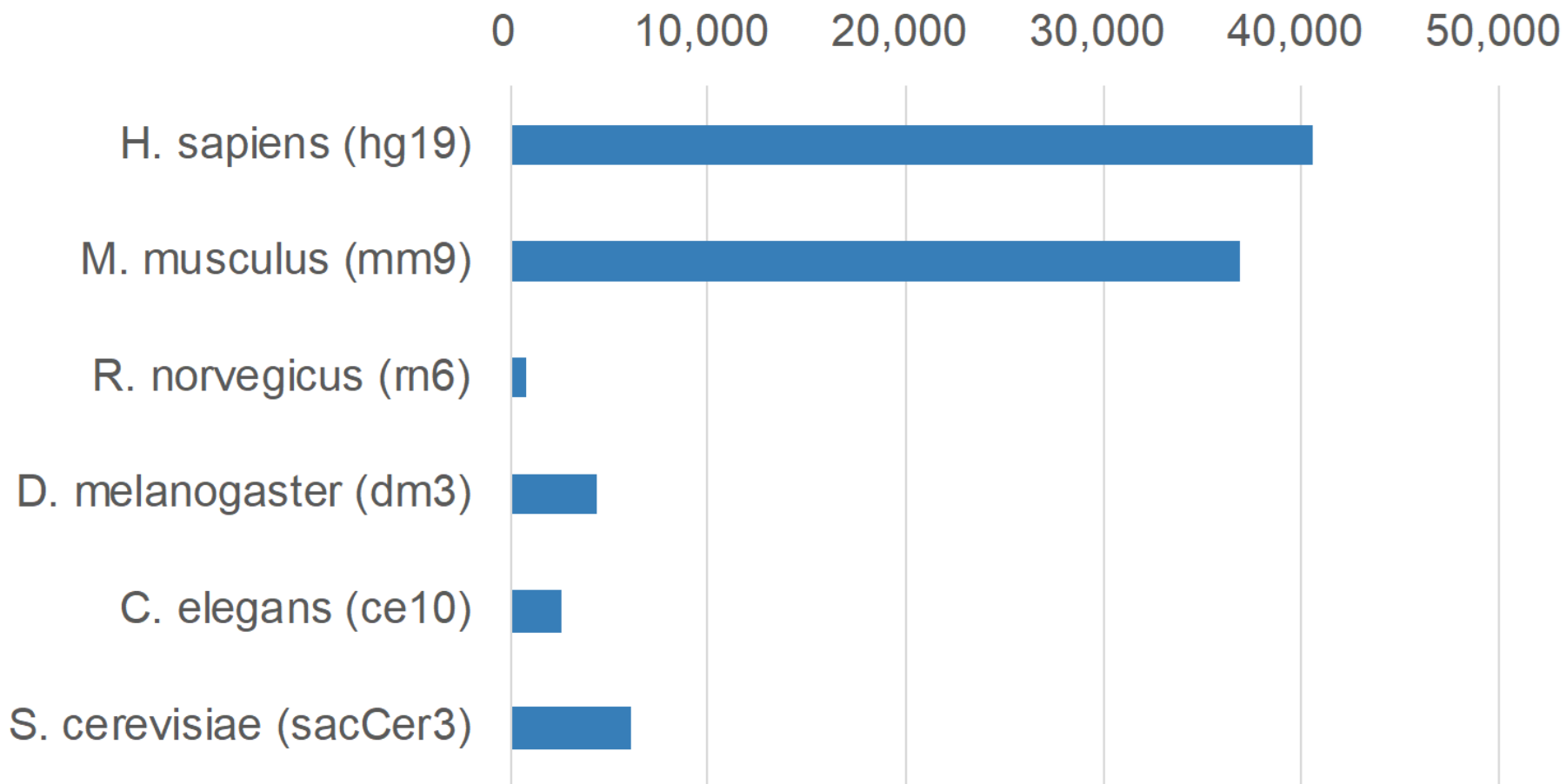
Licensed under a Creative Commons表示4.0国際ライセンス
(c)2018本田 瑞季 (九州大学大学院医学研究院)



なにがどこにひっつく？がすべてわかる！！



ChIP-seqデータの報告件数



As of Nov. 2018.

ChIP-Atlasの4つの機能

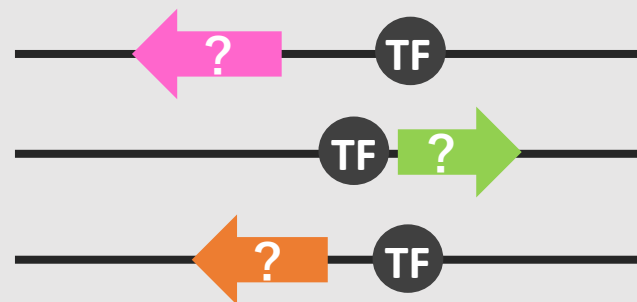
① Peak Browser

何がどこに結合する？



② Target Genes

転写因子の標的遺伝子を探す



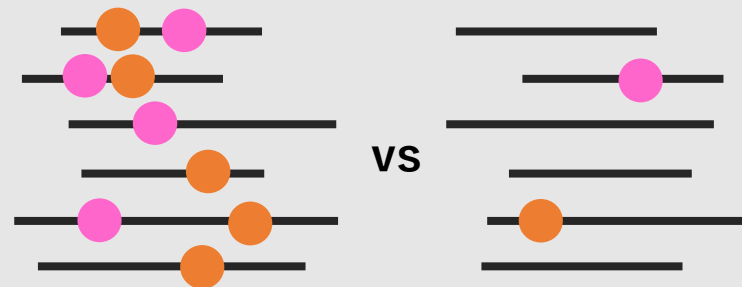
③ Colocalization

共局在する転写因子を探す



④ Enrichment Analysis

ゲノム領域へのEnrichment解析



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 78,000 experiments.

[Watch movie introduction](#)

The four main features of ChIP-Atlas are:

Peak Browser

graphically visualizes peaks at
given genomic loci with
(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 (formerly known as *in
silico* ChIP).

[Watch Movie](#)

CLICK!!

Peak Browser

ChIP-Atlas

Peak Browser

Target Genes

Colocalization

Enrichment Analysis

Documentation

Publications

Find an experiment ▾

ChIP-Atlas - Peak Browser

Tutorial movies ▾

Visualize All Peaks from Published ChIP-Seq data.

H. sapiens

M. musculus

R. norvegicus

D. melanogaster

C. elegans

S. cerevisiae

Antigen Class

All antigens (31775)
DNase-seq (700)
Histone (10011)
RNA polymerase (953)
TFs and others (7879)
Input control (4345)
Unclassified (5876)
No description (2011)

Cell type Class

Cardiovascular (448)
Digestive tract (473)
Embryo (2821)
Embryonic fibroblast (2435)
Epidermis (218)
Gonad (882)
Kidney (170)
Liver (1871)

Threshold for Significance ⓘ

50
100
200
500

Antigen

type to search

All
H2A.XS139ph (3)
H2BK20ac (3)
H3 (1)
H3K27ac (158)
H3K27me3 (192)
H3K36me3 (8)
H3K4me1 (57)

Cell type

type to search

4-cell stage (13)
8-cell stage (6)
Blastocyst Inner Cell Mass (13)
E6.5 embryos (4)
E7.0 embryos (6)
E7.5 embryos (8)
Embryonic brains (125)
Embryonic eye (4)
Embryonic face (3)

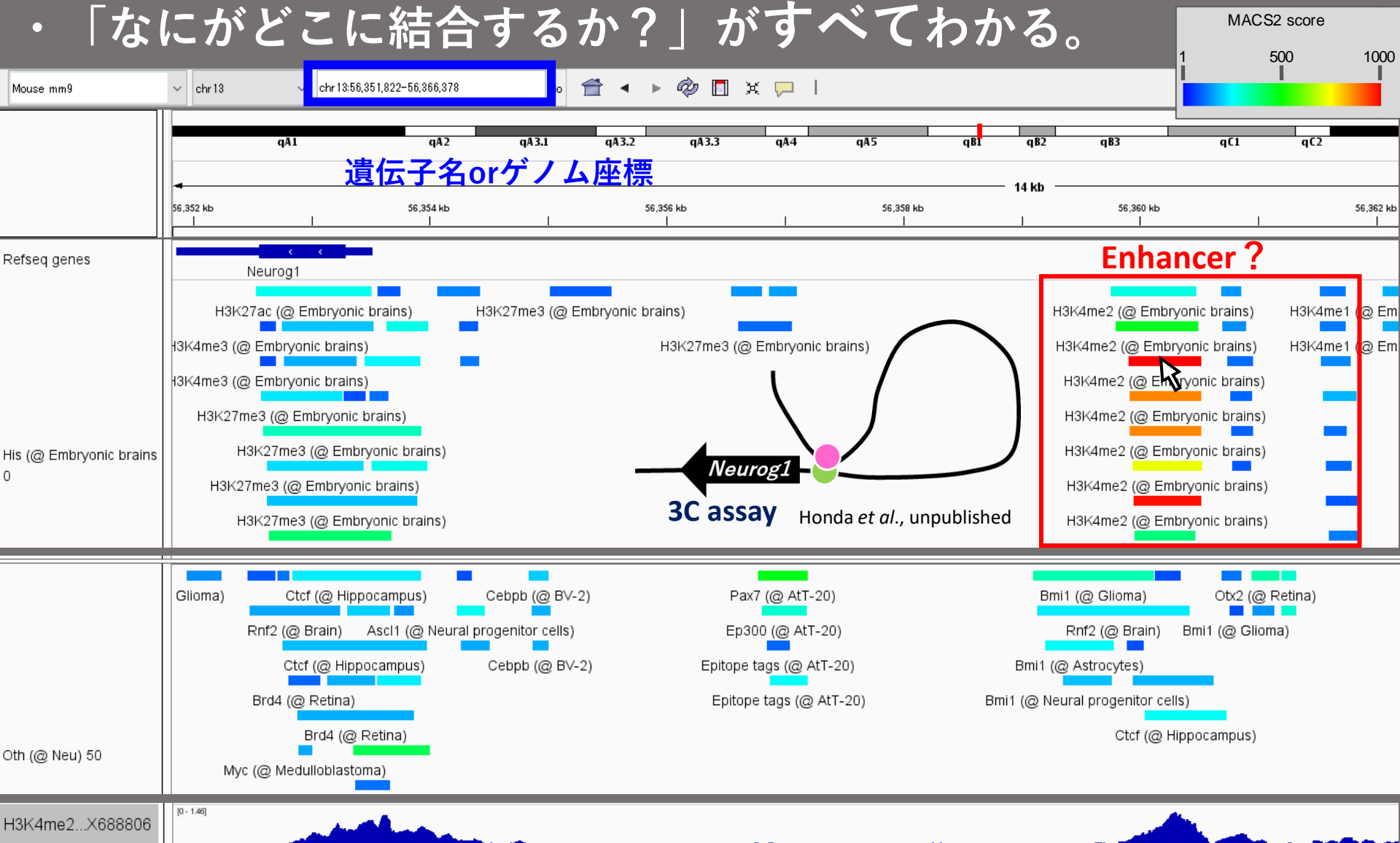
View on IGV

Download BED

CLICK!!

Peak Browser

- 「なにがどこに結合するか？」がすべてわかる。



- 個々の実験の alignment データも閲覧できる

ChIP-Atlasの4つの機能

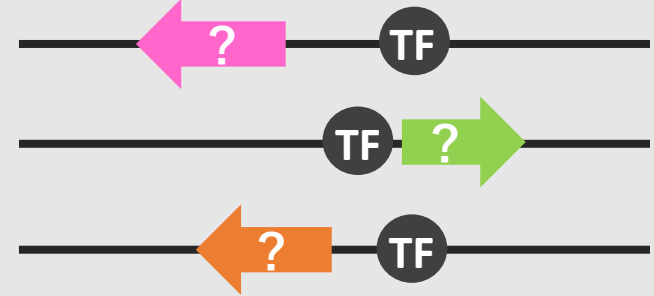
① Peak Browser

何がどこに結合する？



② Target Genes

転写因子の標的遺伝子を探す



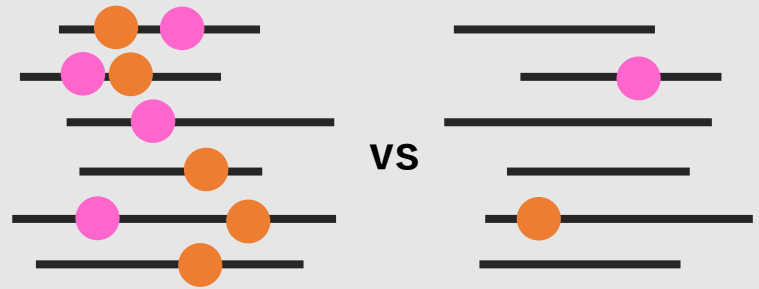
③ Colocalization

共局在する転写因子を探す



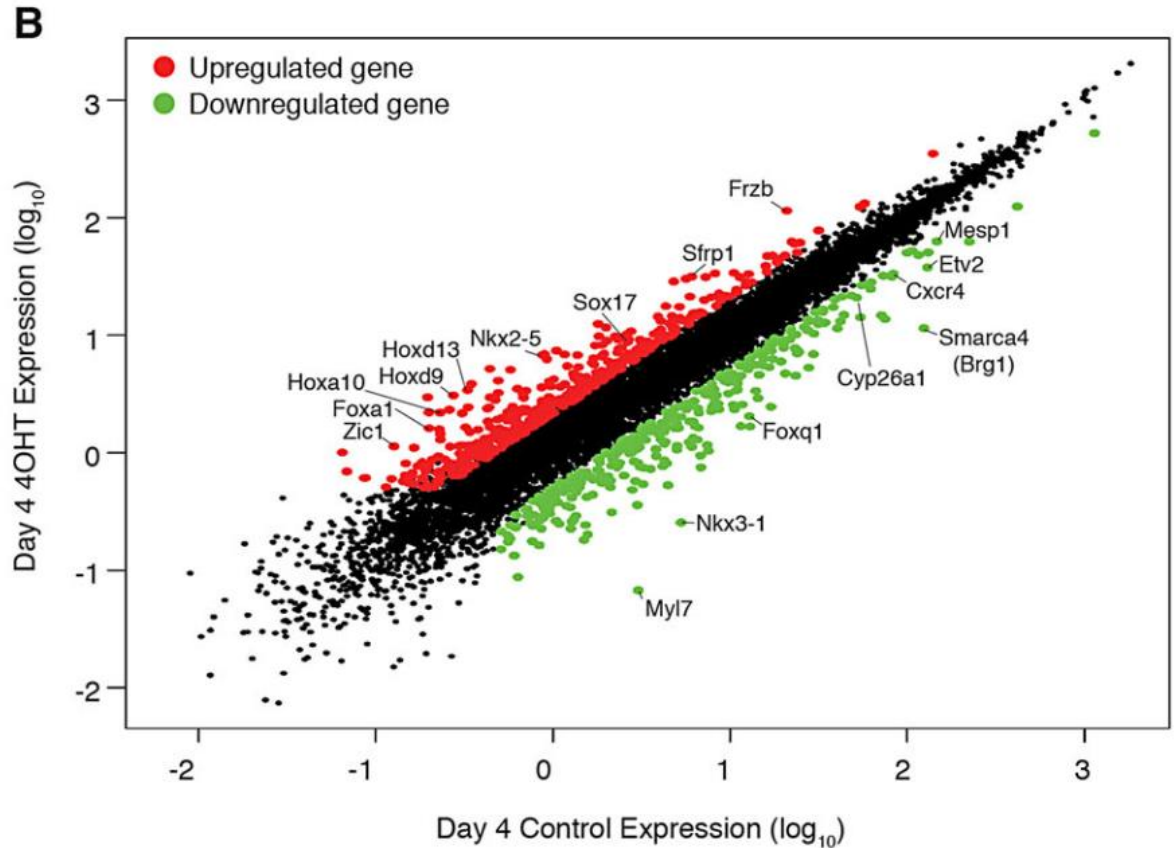
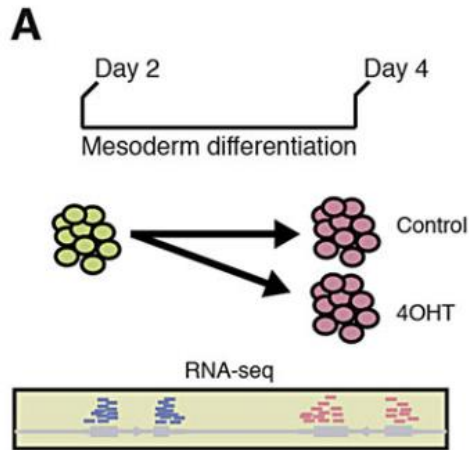
④ Enrichment Analysis

ゲノム領域へのEnrichment解析



こんな方におすすめ

Alexander JM *et al.*, *Development* 2015



遺伝子群の上流を知りたい！

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 78,000 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 (formerly known as *in silico* ChIP).

[Watch Movie](#)

CLICK!!

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

発現が**上昇**する
遺伝子群

5. Select dataset to be compared

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

AGBL5
ALOXE3
AMZ2
ANG
APBB3
ARID2
ASF1A
ATP5O

発現が**低下**する
遺伝子群

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

CLICK!!

Enrichment Analysis

ゲノム座標を入れることもできる

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

炎症性腸疾患 SNP
のゲノム座標

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

その他の SNP の
ゲノム座標

6. Describe datasets

User data title ⓘ

My data

Compared data title ⓘ

Control

Project title ⓘ

My project

submit

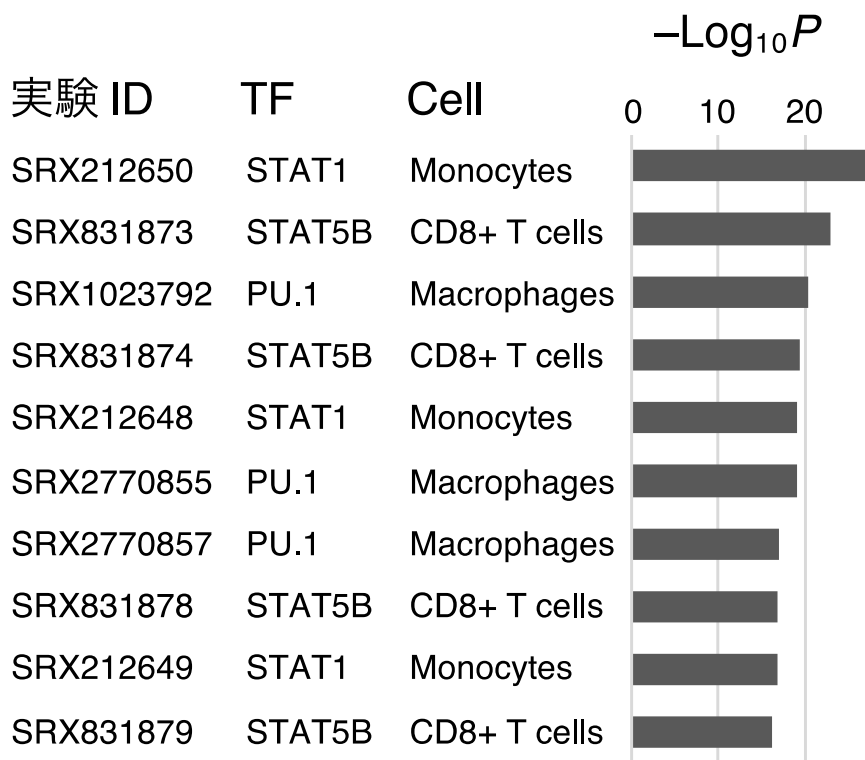
Estimated run time: 2 mins

CLICK!!

Enrichment Analysisの利用例

候補因子 Top10

炎症性腸疾患 SNPs の結果



- 疾患感受性 SNP に enrich する転写因子を同定できる。
- 発症やその進行に関わる分子基盤の解明につながる。

Enrichment Analysisの利用例

解析された**遺伝子**群

- 薬剤投与で変動する遺伝子群 (Anan et al, 2018, *NAR*)
- 若い vs 老化細胞 (Onodera et al, 2018, *Aging Cell*)
- がん化のステージ特異的遺伝子群 (Chatterjee et al, 2018, *BCRT*)

解析された**ゲノム領域**群

- 疾患と関連する SNPs (Oki et al, *in preparation*)
- 遺伝子発現に影響する SNPs (Ishigaki et al, 2018, *Nat Genet*)
- ユーザの ChIP-seq データ (Anan et al, 2018, *NAR*)
- 進化的保存領域 (Ferris et al, 2018, *Cell Rep*)

BioDB9

ChIP-Atlas: 既報のChIP-seqデータをフル活用できる

沖 真弥



第41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 —

日時

2018年11月28日(水) ~ 11月30日(金)

10:00 ~ 17:00
(最終日は15:30まで)

会場

パシフィコ横浜展示ホール A・B・C
バイオデータベース (BioDB) コーナー



謝辞

Web UI の作製, 様々な提言

沖 真弥 (九大)

大田 達郎 (DBCLS)

WABI の作製

小笠原 理

奥田 喜広 (DDBJ)

研究費



構想, 提案

塩井 剛 (RIKEN)

サーバ提供

畠中 秀樹 (NBDC)



基盤研究 B (特設分野)

CoLo の提供

仲木 竜 (東大)

統計解析

瀬々 潤 (産総研)



データ可視化

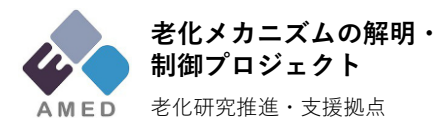
川路 英哉 (RIKEN)

データ考察

目野 主税 (九大)

計算機

NIG supercomputer



その他、ご協力いただいた方々

公益信託
加藤記念難病研究助成基金

