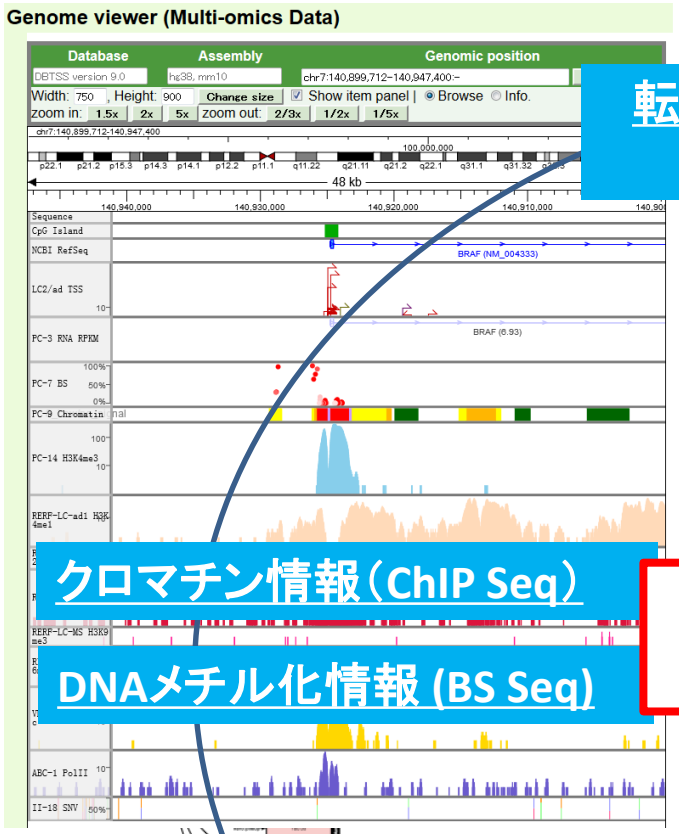
多階層オーミクスデータ DBKERO：(Kashiwa Encyclopedia for Researches of multi-Omics data)

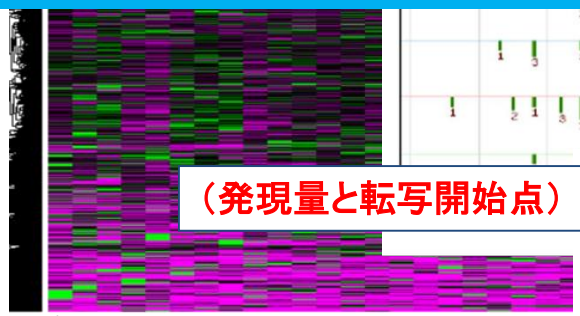
鈴木穣
東京大学新領域創成科学研究科
メディカル情報生命専攻

東京大学
THE UNIVERSITY OF TOKYO

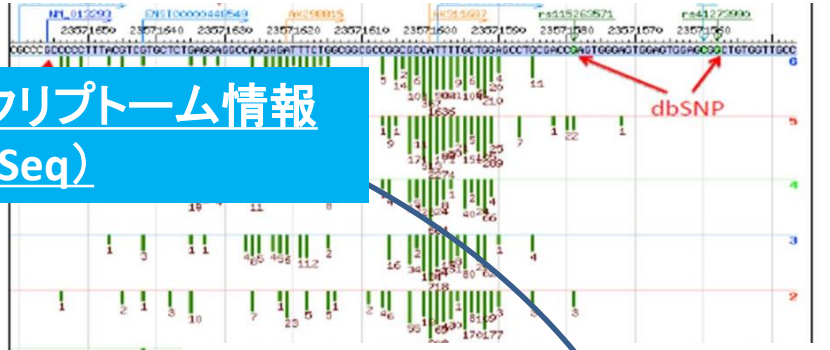
全体のねらい　ヒト応用研究を志向したオミクス情報の統合

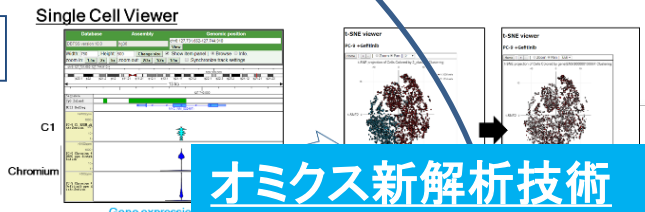


転写開始点/トランスクリプトーム情報
（TSS/RNA Seq）



（発現量と転写開始点）





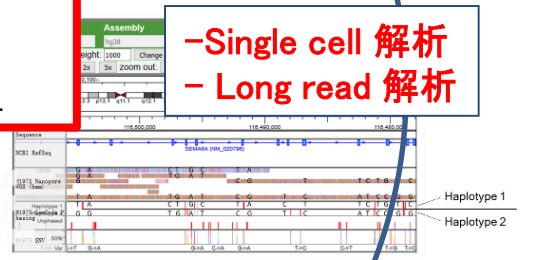
オミクス新解析技術

クロマチン情報（ChIP Seq）

DNAメチル化情報 (BS Seq)

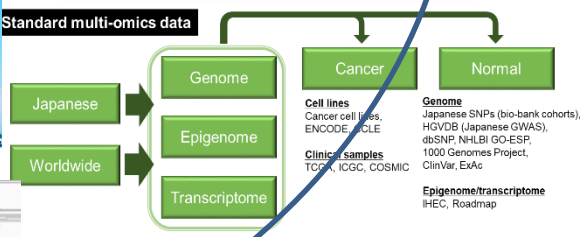
多層オーミクス情報の
ヒトゲノム変異情報への統合

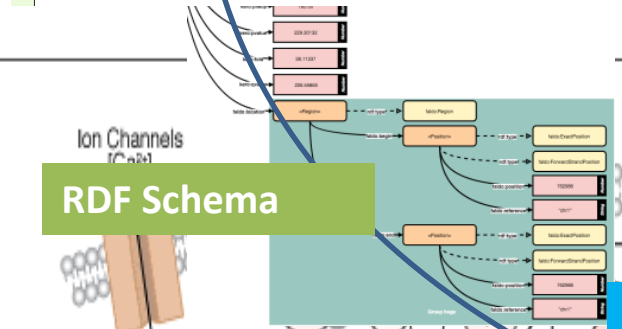
-Single cell 解析
- Long read 解析



統合DB（DBKERO）

（それぞれの検体での変異部位）

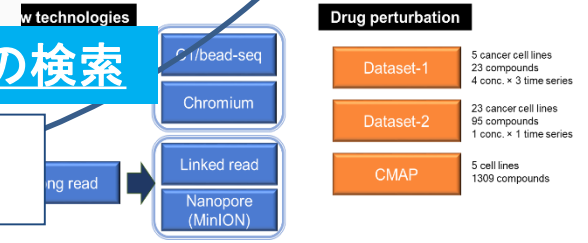




RDF Schema

パスウェイマップ（文献情報）からの検索

（該当集団中の遺伝子変異頻度を
赤の濃さで示す）



http://kero.hgc.jp/

DBKERO
Data Base of Encyclopedia of Regulatory Omics

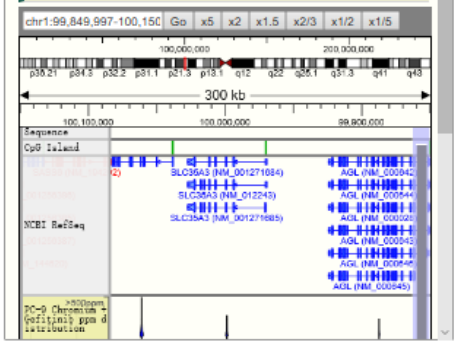
We recommended to use Edge (V40 above), Google Chrome (V61 above) or Firefox (V56 above) for the DBKERO browsing. Internet Explorer has not been supported.

Tools

- Genome Browser [GitHub]
  - Search from Keyword or Genomic Position (human and mouse)
  - Search from SNV-enriched Gene in Cancers
- TF Binding Site Search
  - Transcription Factor Binding Site Search [New]
- Pathway Map
  - Human Pathway Map
- RDF for NGS Analysis Results (Trial) (Lung adenocarcinoma 26 cell lines: RNA-seq, ChIP-seq, SNV, BS-seq, TSS-seq)
  - RDF Schema
  - RDF Browser
  - SPARQL Endpoint
- Chromatin Features (for Lung adenocarcinoma 26 cell lines)
  - Search from Genomic Position
  - Search from SNP (dbSNP rsID)
  - Search from SNV (COSMIC: somatic mutation)
- Overview of mutation frequency in patients
  - SNV Summary in Cancers
- Documents
  - Experimental Procedures
  - Data Contents
  - Help
  - Download
  - Links
  - References

Genome browser

- Display tracks
  - Single cell
  - Long read
  - Multi-omics
  - Cancer genome
  - Disease genome
- Demo



Movie

Overview movie in Japanese (53 min.)

使い方の説明

Tutorialの動画

何ができるか（1）

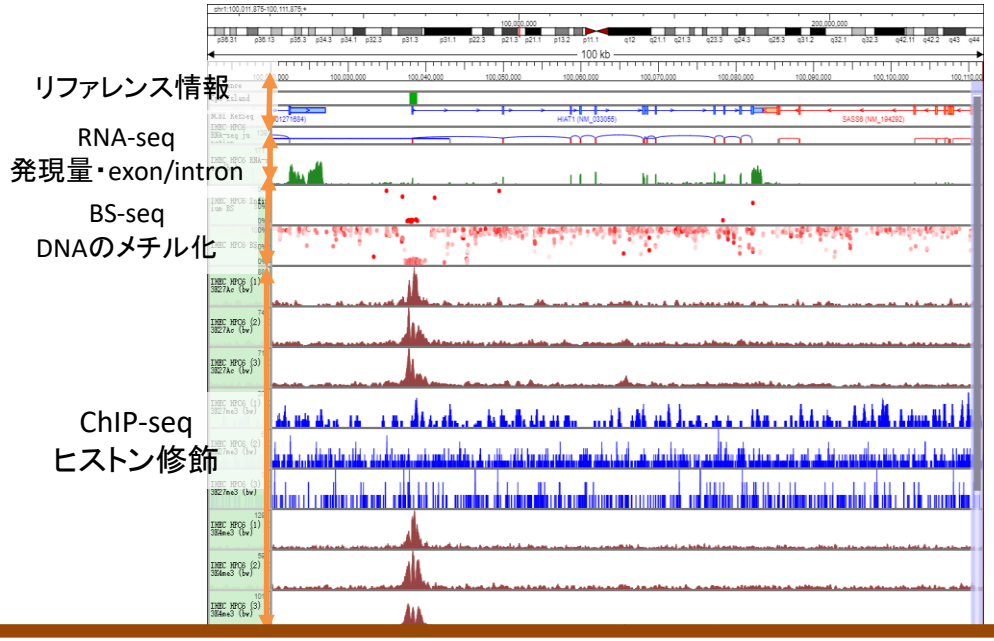
日本人正常オーミクスデータ （IHEC）

国際エピゲノムコンソシアム （IHEC）
で日本CRESTチームが収集した日本人標準エピゲノムが検索できます。

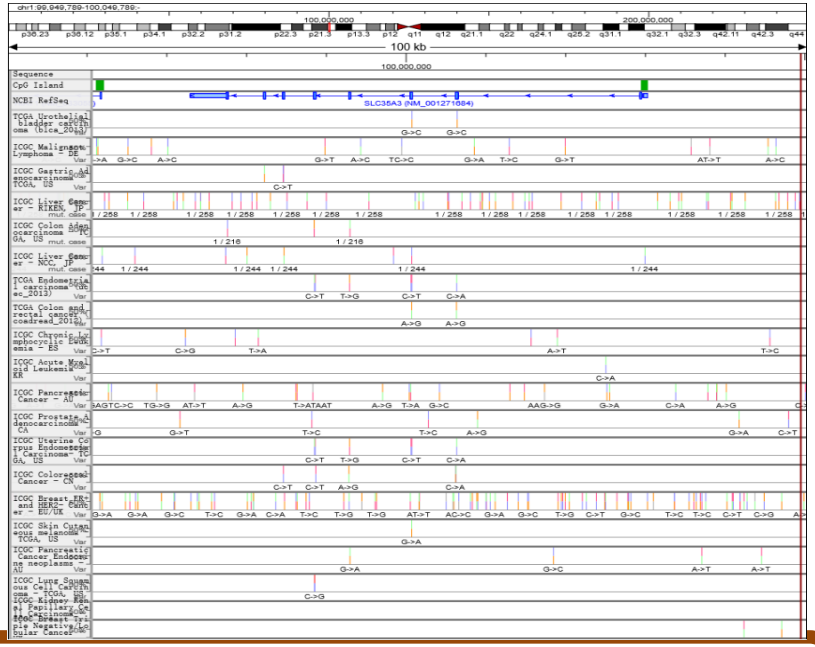
CREST-IHECデータの統合（日本人標準エピゲノム）

IHEC（国際エピゲノムコンソシアム）データの公開

HPC6肝臓細胞のマルチオーミクス情報の表示



TCGA/ICGC変異情報の表示



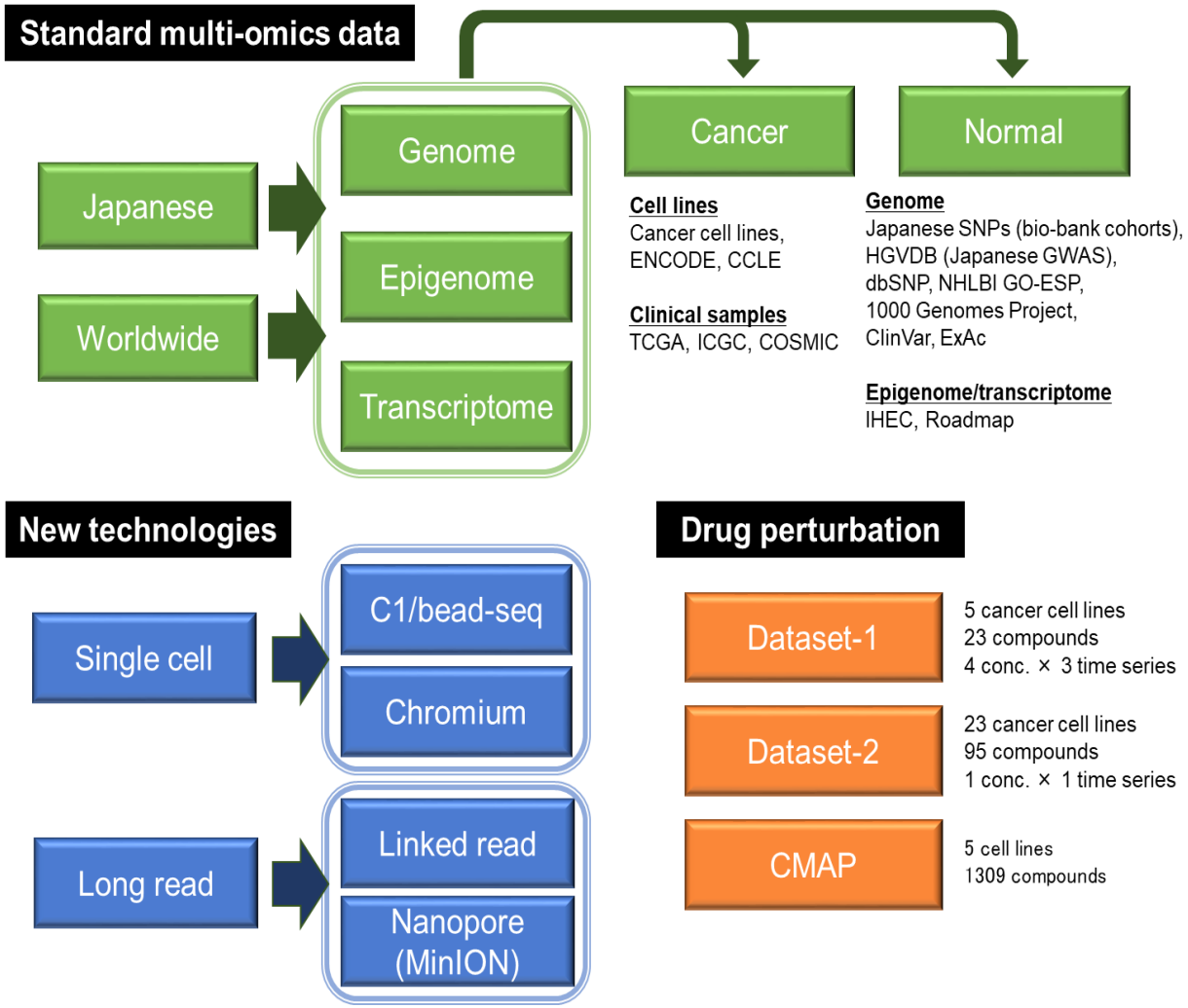
IHECデータとしてゲノムブラウザから閲覧可能なオープンデーター覧

Sample	individuals	Expression		DNA methylation	Histone modifications					
		RNA-seq		BS-seq	H3K4me1	H3K4me3	H3K27ac	H3K27me3	H3K36me3	H3K9me3
		Read depth	Exon/intron junction							
Liver	8	8	8	8	13	13	13	13	13	13
Colon	11	11	11	11	11	11	11	11	11	11
Endometrial	15	30	0	13	14	15	15	15	15	15
Vascular endothelial	33	26	0	0	19	27	27	19	27	26

IHEC日本チームの作成した正常肝臓、血管内皮、胎盤組織におけるエピゲノムカタログを標準エピゲノムデータとして収載した．上の表は既にKEROから閲覧可能なデータセットの数．ゲノムブラウザ下部のTracksにある「CREST – IHEC (The International Human Epigenome Consortium (AMED-CREST, Japan))」の一覧表から、任意のデータ結果をゲノムブラウザ上に追加可能である（上図）．

データコンテンツの選択と表示

KEROデータコンテンツ



DBKEROのデータ構造の全体像．日本人の臨床オミックス情報がモデルシステムからの包括的オミックス情報とどのように関連しているかを示す．

New technologiesとして、近年新たに産出可能となったシングルセルとロングリードのデータやDrug perturbationとして、化合物によるマルチオーミクス摂動の情報を含んでいる．

異なるカテゴリーのデータセットは、異なる色のボックスとして示した．

Demo

Human Genome Variation DB（旧徳永グループ）との統合

日本人ゲノム多型

DBKERO

Search

HGVDBへのリンク機能

Human Variation DB

Genome browserへGWAS情報の可視化

Whole genome viewer, Manhattan viewer

NBDC ヒトDBとの連携へ

何ができるか（2）

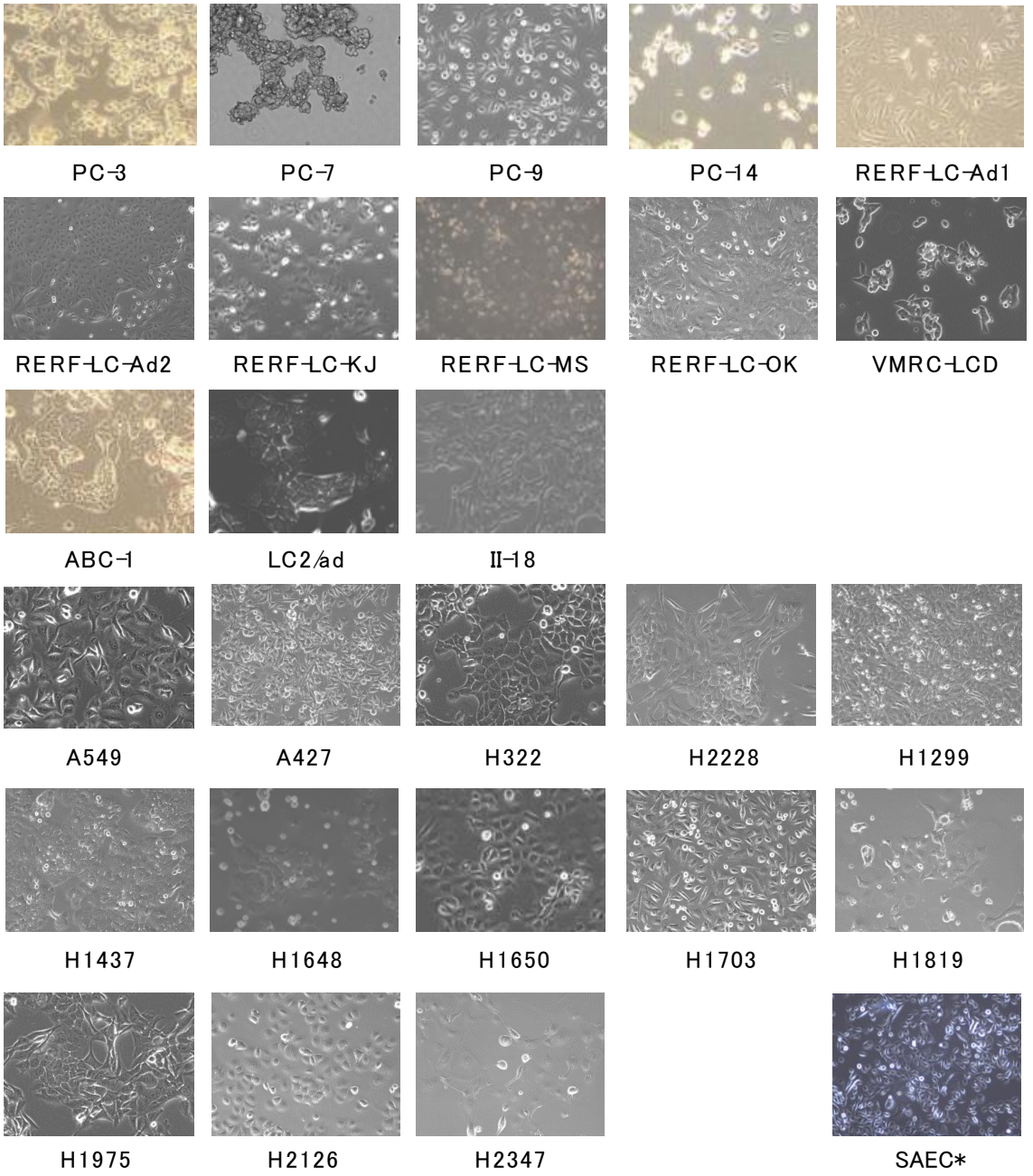
どの培養細胞をモデルにするのか？を選ぶオーミクスデータ

肺腺癌細胞株等で日本人由来のオーミクスデータが充実しています。

Materials

26 lung adenocarcinoma cell lines

Cell line	Origin
PC-3	Japanese
PC-7	
PC-9	
PC-14	
RERF-LC-Ad1	
RERF-LC-Ad2	
RERF-LC-KJ	
RERF-LC-MS	
RERF-LC-OK	
VMRC-LCD	
ABC-1	
LC2/ad	
II-18	
A427	Non-Japanese
A549	
H322	
H2228	
H1299	
H1437	
H1648	
H1650	
H1703	
H1819	
H1975	
H2126	
H2347	



*Normal small airway epithelial cells

Multi-omics sequencing data of lung cancer cell lines

26 NSCLC cell lines

13 Japanese
13 non-Asian
+ normal (SAEC)

Genome

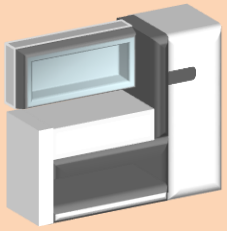
- Whole-genome sequencing
- Long-read sequencing (10X/MinION)

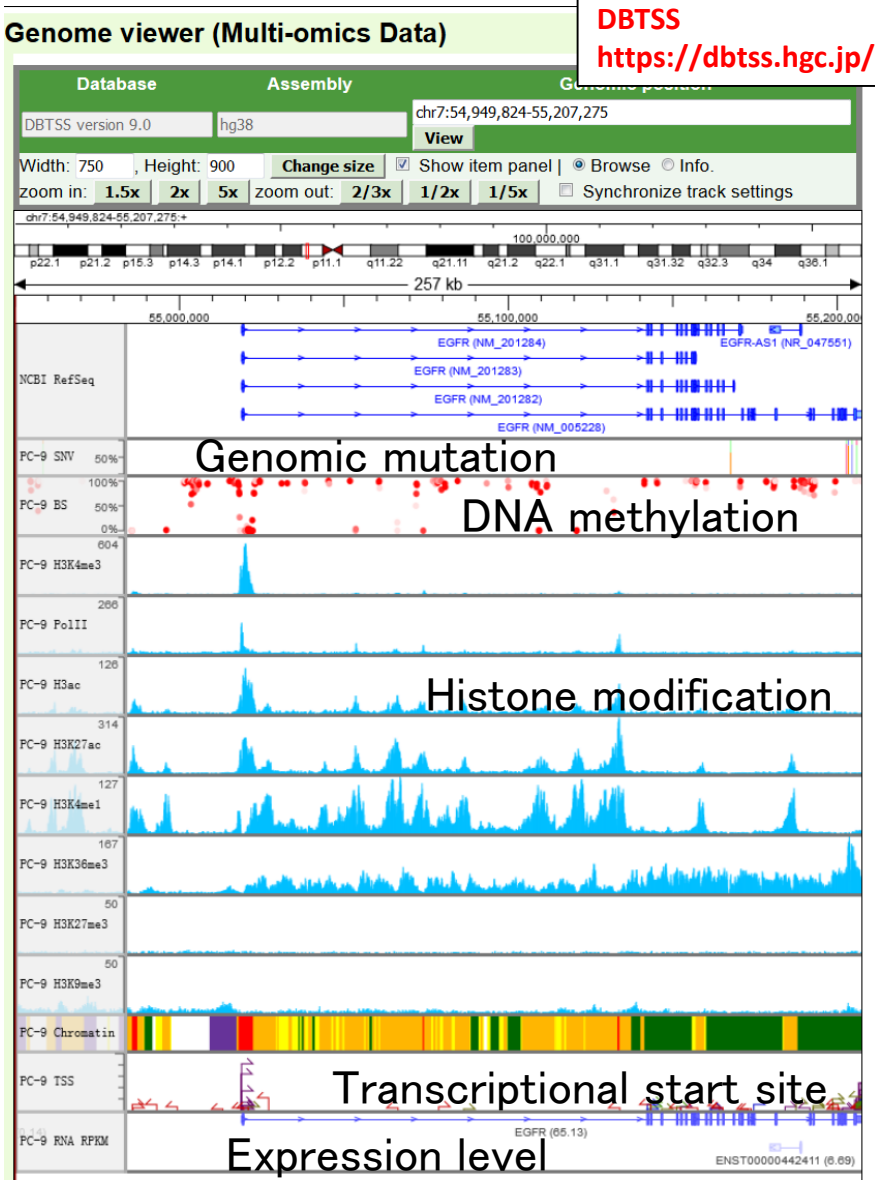
Epigenome

- Bisulfite sequencing
- ChIP-Seq
- ATAC-Seq

Transcriptome

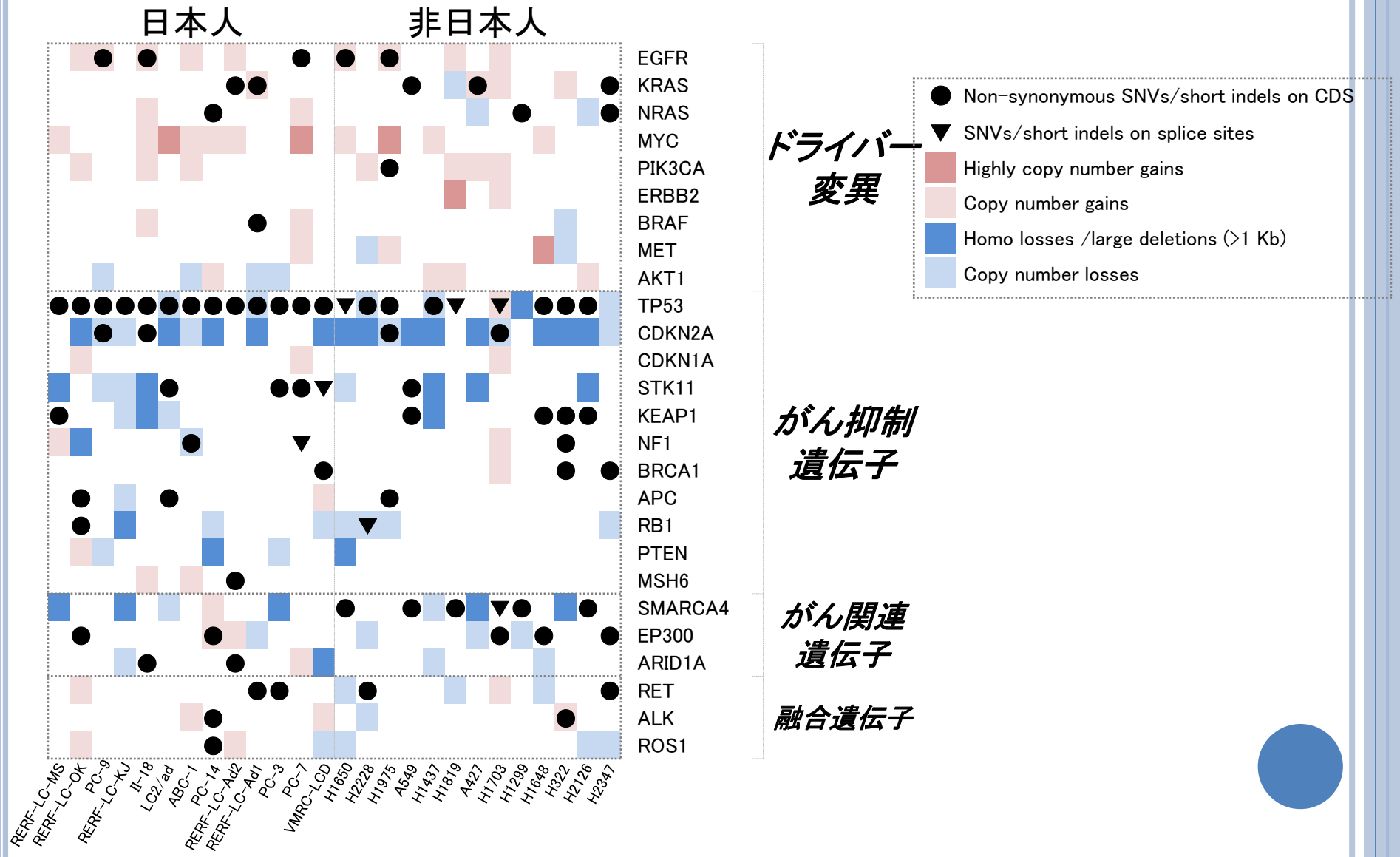
- TSS-Seq
- RNA-Seq
- miRNA-seq



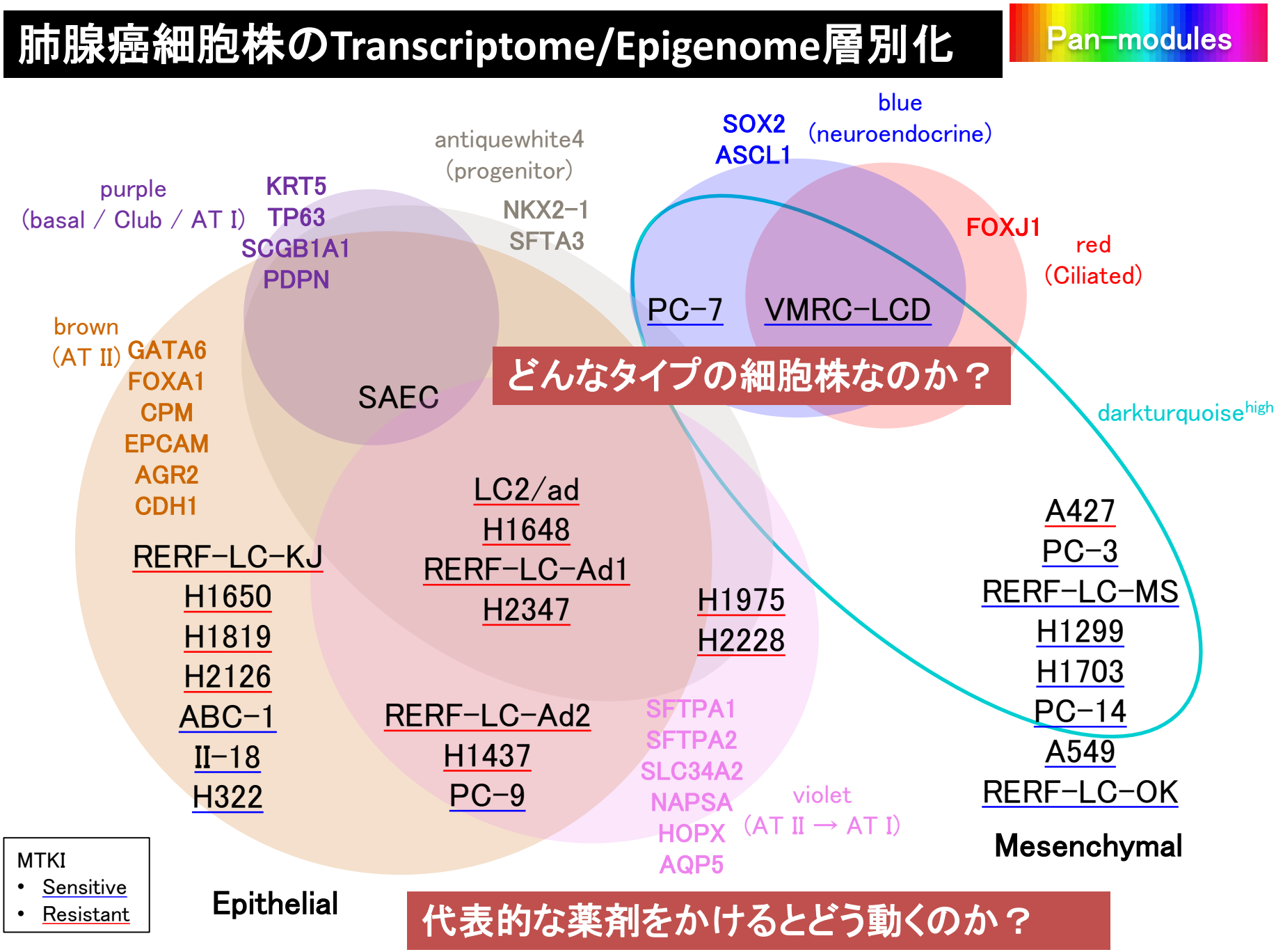


Suzuki *et al.* 2014 *Nucleic Acids Research*
Suzuki *et al.* 2015 *Nucleic Acids Research DB issue*

どんなゲノム変異をもっている細胞株なのか



肺腺癌細胞株のTranscriptome/Epigenome層別化



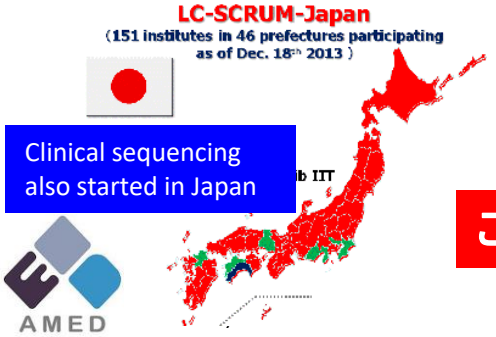
何ができるか（3）

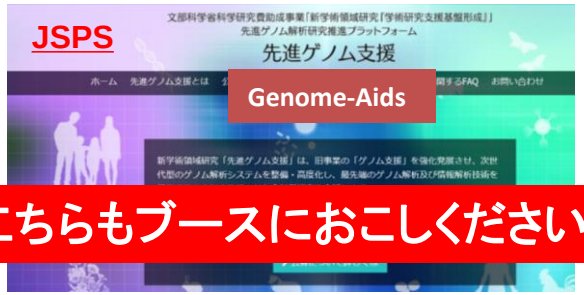
新しい手法でどんなデータが出てくるのか？

シングルセルデータ、ナノポアデータ等がおちています。

東大・柏拠点

On-going contributions as a sequence center



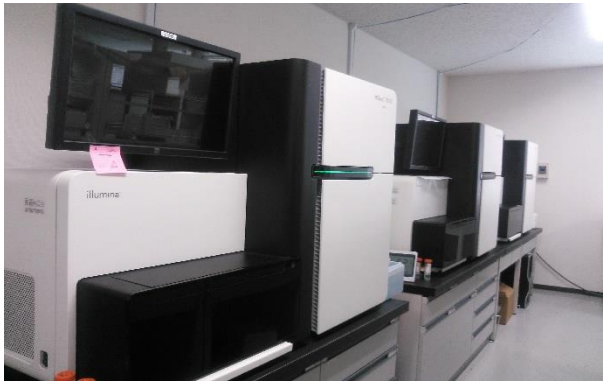


こちらもブースにおこしください。

Hiseq2500 x 6+ Hiseq3000 x 1
NovaSeq x 1

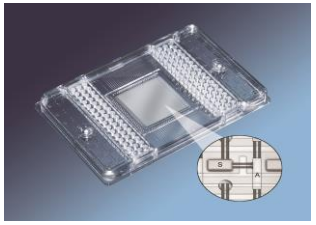
Approximate amount of data produced so far

	Genome	Transcriptome	Epigenome
#samples	2000	3000	5000



And serving as an Incubation center for new genome technologies)

Single cell analyzer;
C1 (Fluidigm)



“Chip” of C1

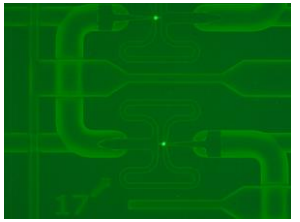
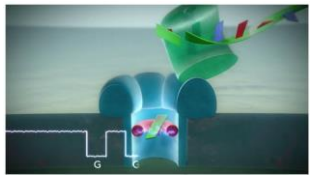


Image of single-cell capture on C1

PI: Yutaka Suzuki.
Univ. Tokyo

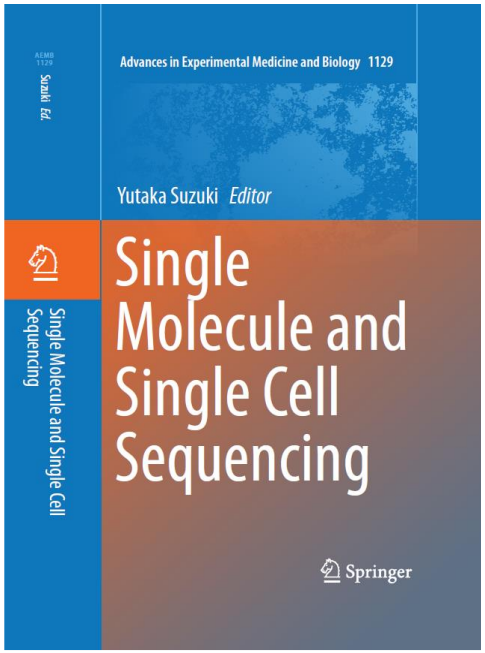
Operators:
Technicians 5
Programmers 4



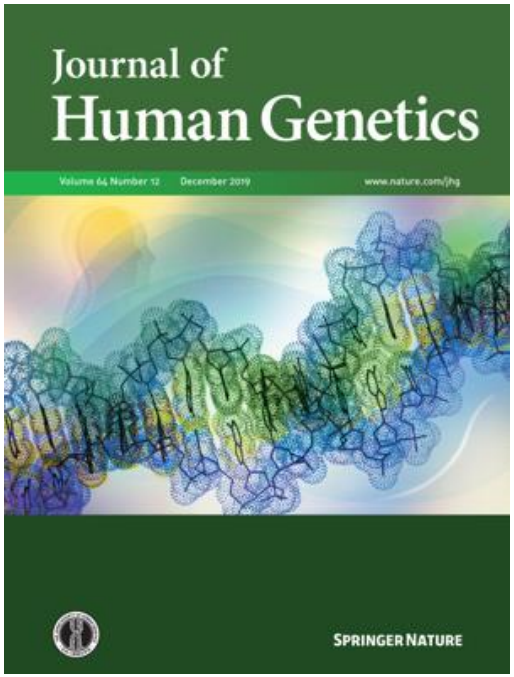
Beta-test of nanopore sequencer

ysuzuki@hgc.jp

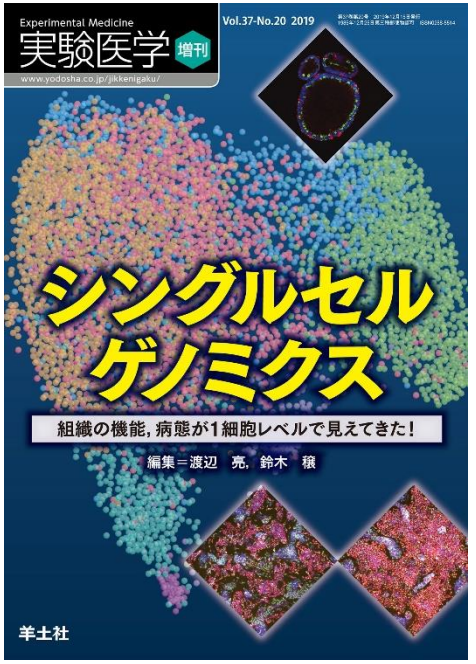
参考文献（シングルセル＋ロングリード）



Nature Springer Book
Single Molecule and Single Cell Sequencing.
(Yutaka Suzuki ed.)



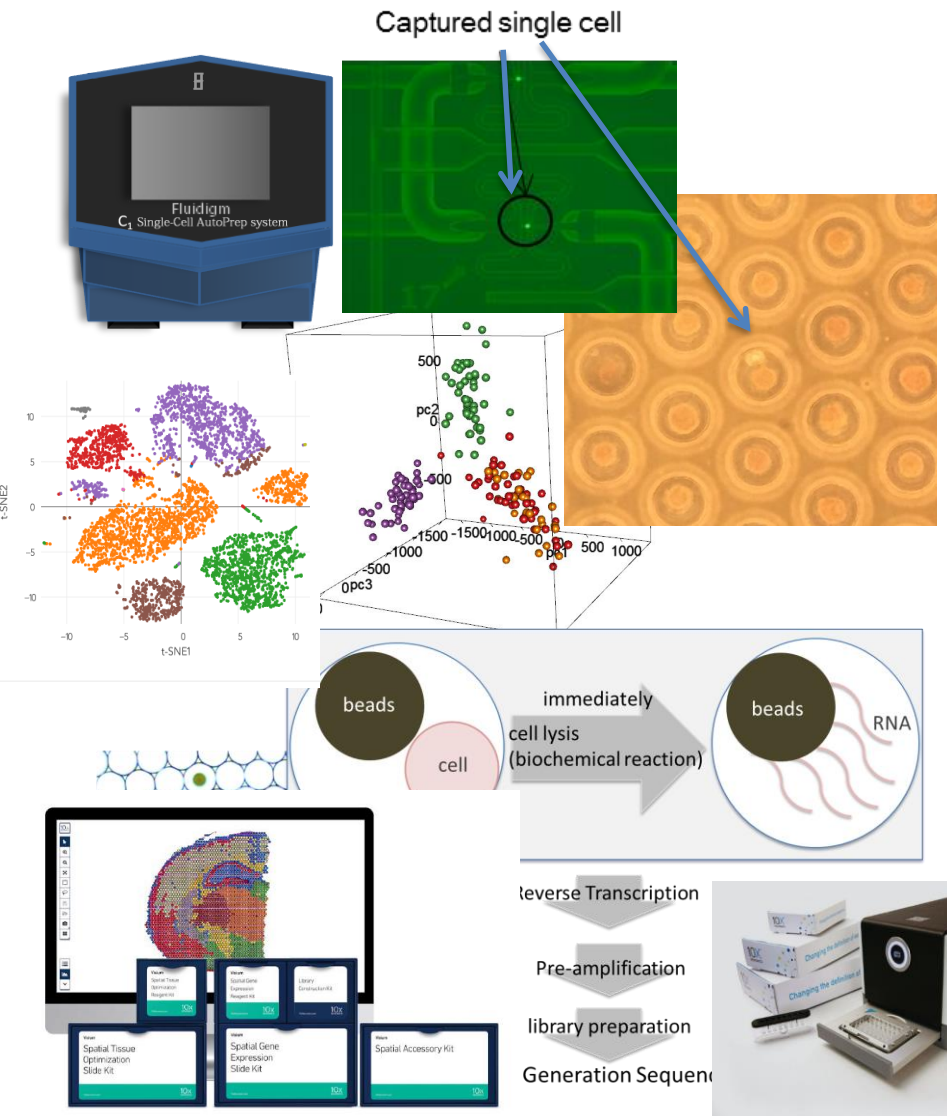
Journal of Human Genetics
Special issue for “New type Sequencer”
Jan (2020)
(Yutaka Suzuki ed.)

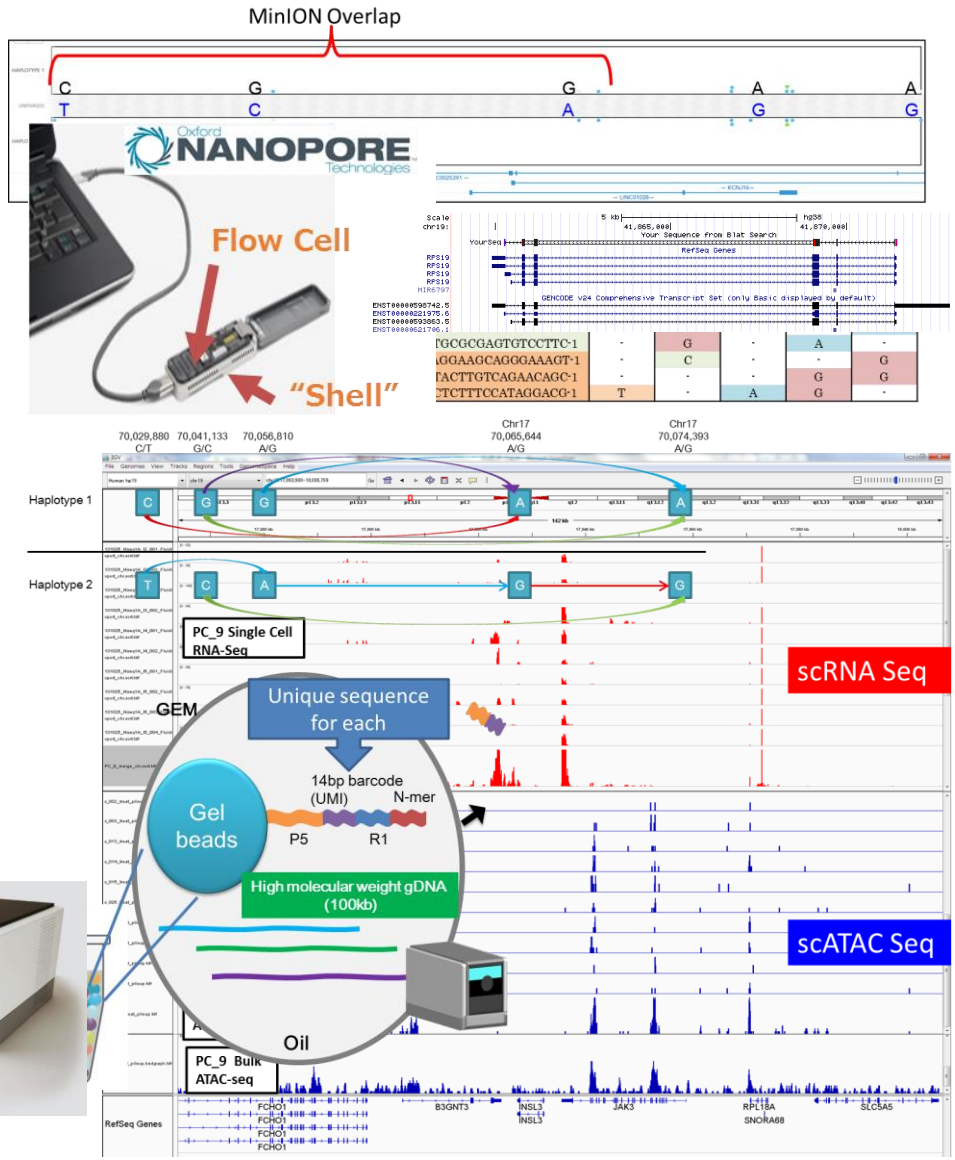


本日先行発売！

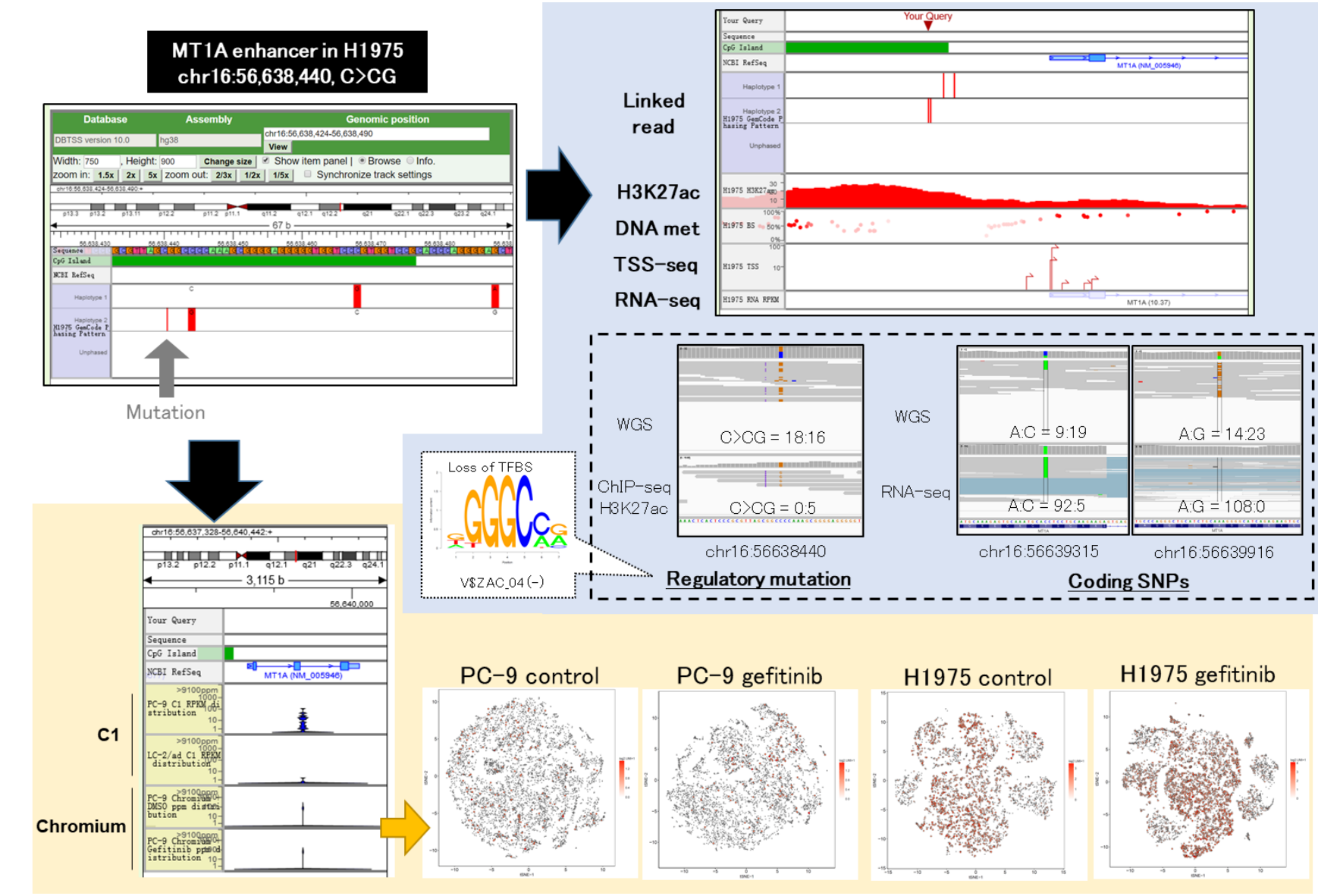
新しい手法でどんなデータが出てくるの？

シングルセル解析（C1, Chromium、その他）　ロングリード解析（GemCode, MinION）





シングルセルとロングリードのデータ閲覧



MT1A遺伝子での例．ゲノムブラウザを中心に目的の遺伝子に関連したプロモーター領域や関連のシングルセルのデータを確認することができる．

http://kero.hgc.jp/

DBKERO
Data Base of Encyclopedia of Regulatory Omics

We recommended to use Edge (V40 above), Google Chrome (V61 above) or Firefox (V56 above) for the DBKERO browsing. Internet Explorer has not been supported.

Tools

Genome Browser [GitHub]
- Search from Keyword or Genomic Position (human and mouse)
- Search from SNV-enriched Gene in Cancers

TF Binding Site Search
- Transcription Factor Binding Site Search [New]

Pathway Map
- Human Pathway Map

RDF for NGS Analysis Results (Trial) (Lung adenocarcinoma 26 cell lines: RNA-seq, ChIP-seq, SNV, BS-seq, TSS-seq)
- RDF Schema
- RDF Browser
- SPARQL Endpoint

Chromatin Features (for Lung adenocarcinoma 26 cell lines)
- Search from Genomic Position
- Search from SNP (dbSNP rsID)
- Search from SNV (COSMIC: somatic mutation)

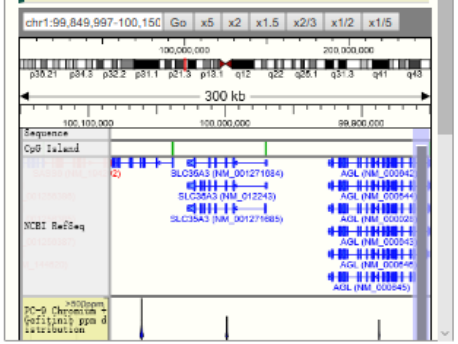
Overview of mutation frequency in patients
- SNV Summary in Cancers

Documents
- Experimental Procedures
- Data Contents
- Help
- Download
- Links
- References

Genome browser

Display tracks
- Single cell
- Long read
- Multi-omics
- Cancer genome
- Disease genome

Demo



Movie

Overview movie in Japanese (53 min.)

データポータル
（準備中）

http://kero.hgc.jp/

DBKERO Data Portal (Sample)

Project: ALL

Race: ALL

Omics Category: ALL

Show 25 entries Search:

track_id	ID	Sample	Race	Omics category	Project	Data type	Genome version	Data download
2295	wg_bw_VMRCLCD	VMRCLCD	japanese	genome	Cell_lines	Depth (Genome)	hg38	Download
2293	wg_bw_RERFLCOK	RERFLCOK	japanese	genome	Cell_lines	Depth (Genome)	hg38	Download
2291	wg_bw_RERFLCMS	RERFLCMS	japanese	genome	Cell_lines	Depth (Genome)	hg38	Download
2289	wg_bw_RERFLCKJ	RERFLCKJ	japanese	genome	Cell_lines	Depth (Genome)	hg38	Download
2287	wg_bw_RERFLCad2	RERFLCad2	japanese	genome	Cell_lines	Depth (Genome)	hg38	Download
2285	wg_bw_RERFLCad1	RERFLCad1	japanese	genome	Cell_lines	Depth (Genome)	hg38	Download
2281	wg_bw_PC9	PC9	japanese	genome	Cell_lines	Depth (Genome)	hg38	Download
2279	wg_bw_PC7	PC7	japanese	genome	Cell_lines	Depth (Genome)	hg38	Download
2277	wg_bw_PC3	PC3	japanese	genome	Cell_lines	Depth (Genome)	hg38	Download
2283	wg_bw_PC14	PC14	japanese	genome	Cell_lines	Depth (Genome)	hg38	Download
2275	wg_bw_LC2ad	LC2ad	japanese	genome	Cell_lines	Depth (Genome)	hg38	Download
2299	wg_bw_II18	II18	japanese	genome	Cell_lines	Depth (Genome)	hg38	Download
2305	wg_bw_H322	H322	worldwide	genome	Cell_lines	Depth (Genome)	hg38	Download
2325	wg_bw_H2347	H2347	worldwide	genome	Cell_lines	Depth (Genome)	hg38	Download

ファイルの選択（準備中）

http://kero.hgc.jp/

日本人正常オーミクスデータ（IHEC）

がん培養細胞オーミクスデータ

シングルセル・ナノポアデータ

DBKERO
Data Base of Encyclopedia of Regulatory Omics

初めての方は、“For the first users”へ

We recommended to use Edge (V40 above), Google Chrome (V61 above) or Firefox (V58 above) for the DBKERO browsing. Internet Explorer has not been supported.

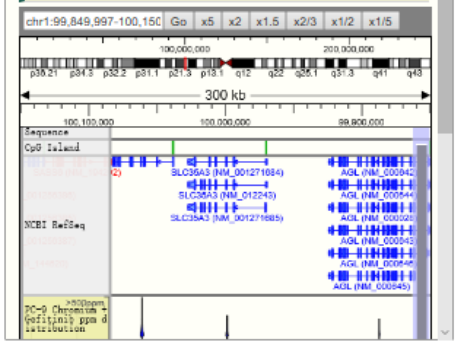
とりあえず
ブラウザへ
クリック！

Tools

- Genome Browser [GitHub]
  - Search from Keyword or Genomic Position (human and mouse)
  - Search from SNV-enriched Gene in Cancers
- TF Binding Site Search
  - Transcription Factor Binding Site Search [New]
- Pathway Map
  - Human Pathway Map
- RDF for NGS Analysis Results (Trial) (Lung adenocarcinoma 26 cell lines: RNA-seq, ChIP-seq, SNV, BS-seq, TSS-seq)
  - RDF Schema
  - RDF Browser
  - SPARQL Endpoint
- Chromatin Features (for Lung adenocarcinoma 26 cell lines)
  - Search from Genomic Position
  - Search from SNP (dbSNP rsID)
  - Search from SNV (COSMIC: somatic mutation)
- Overview of mutation frequency in patients
  - SNV Summary in Cancers
- Documents
  - Experimental Procedures
  - Data Contents
  - Help
  - Download
  - Links
  - References

Genome browser

- Display tracks
  - Single cell
  - Long read
  - Multi-omics
  - Cancer genome
  - Disease genome
- Demo



Movie

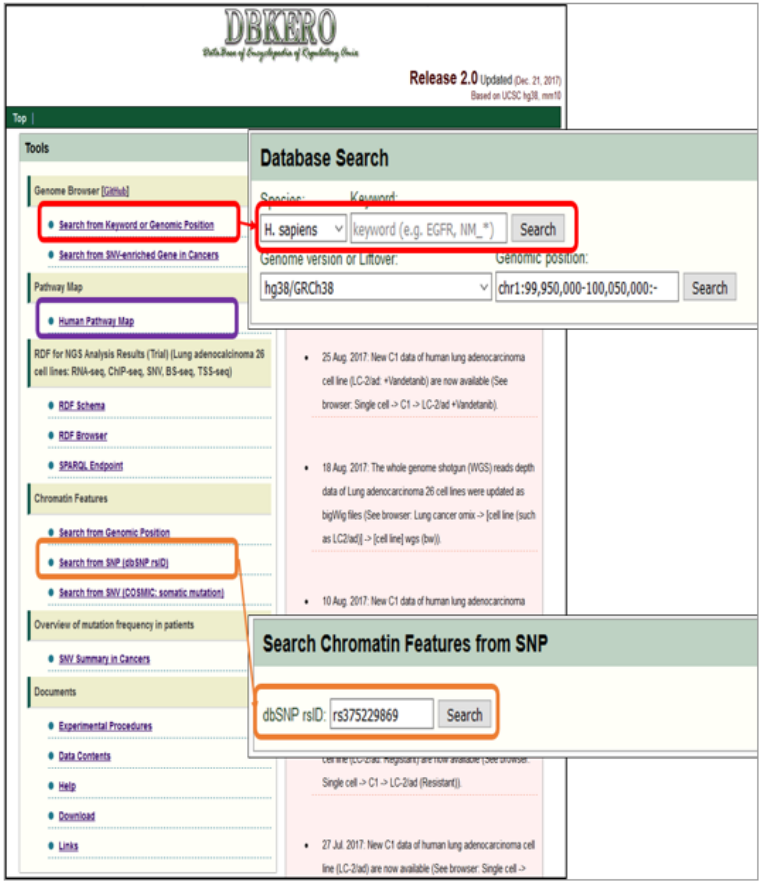
Overview movie in Japanese (53 min.)

使い方の説明

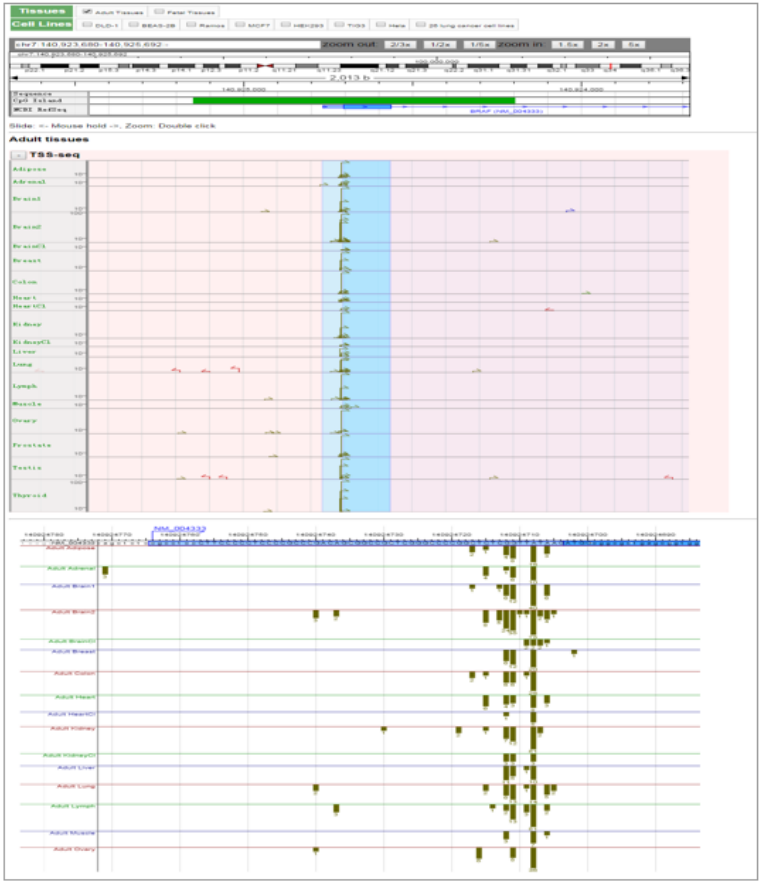
Tutorialの動画

他にも
・ChIP-Atlasとの連携
・RDF化のスキーマの作成
を進めています。

A.



B.



A. Top page of DBKERO. A simple search for "TSS Viewer" and "Genome Viewer" can be made by specifying a keyword, such as a gene name 'BRAF' in the Database Search at the left frame (red box). Search by "SNV Summary in Cancers" and "Pathway Maps" can be made from the positions indicated by orange and purple boxes, respectively.

B. A part of the TSS Viewer display for the BRAF gene. The overview and the detailed positions of the TSSs are shown in the upper and lower panels, respectively. Many of the fields are expandable.

http://kero.hgc.jp/

ysuzuki@hgc.jp

展示会場でブースを出していますので、ぜひ、お立ち寄りください！

DB使用例

2P-0005:シングルセル（Talkは終了）
2P-0047: 培養細胞モジュール化
2P-0048:培養細胞抗がん剤試験（Talkは終了）
3P-0006:ナノポアがんゲノム
3P-0026:病原体ロングリードアセンブル
3P-0514:培養細胞mRNA全長ロングリード
4P-0015:On site sequencing