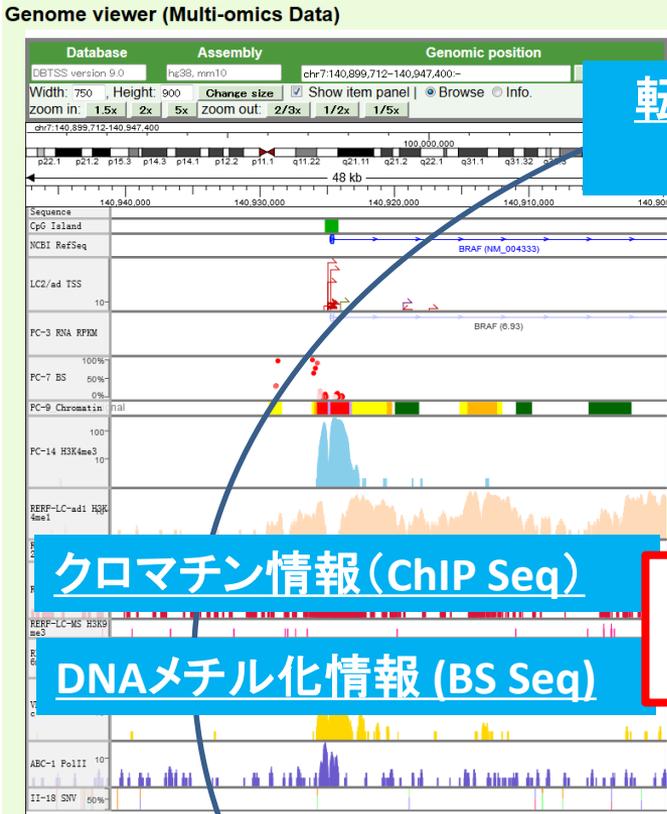


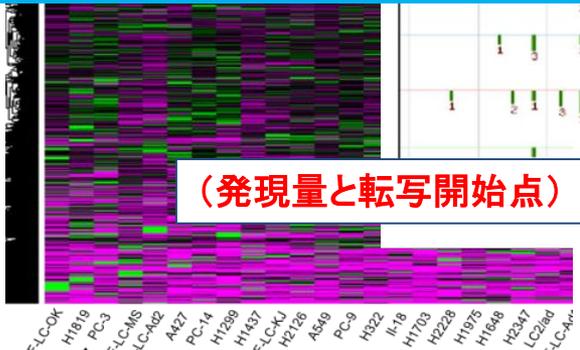
多階層オームクスデータ
DBKERO:
(Kashiwa Encyclopedia for Researches
of multi-Omic data)

鈴木穰

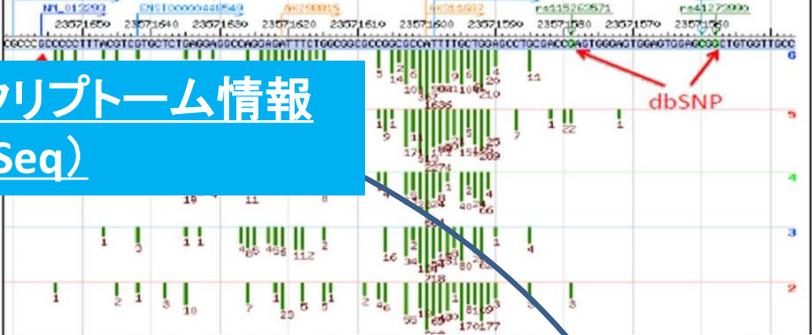
東京大学新領域創成科学研究科
メディカル情報生命専攻



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

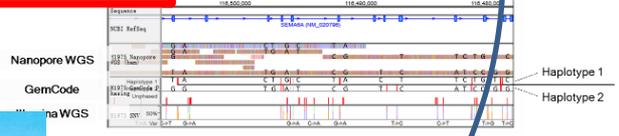
クロマチン情報 (ChIP Seq)

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

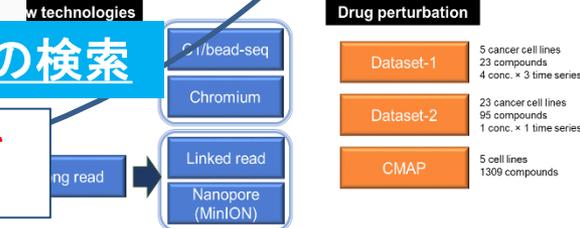
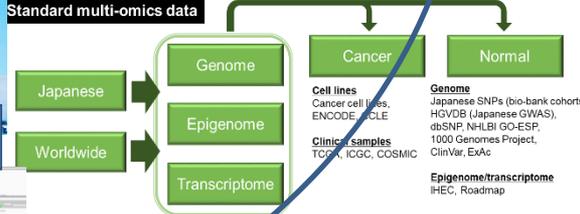
- Single cell 解析
- Long read 解析

統合DB (DBKERO)



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/



We recommend 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](#)

使い方の説明



Genome browser

Display tracks

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

Demo

chr1:99,849,997-100,150,000 Go x5 x2 x1.5 x2/3 x1/2 x1/5

100,000,000 200,000,000

300 kb

Sequence

CpG Island

NCBI RefSeq

PC=0 Chromatin positions ppm d

Movie

Overview movie in Japanese (53 min.)

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)

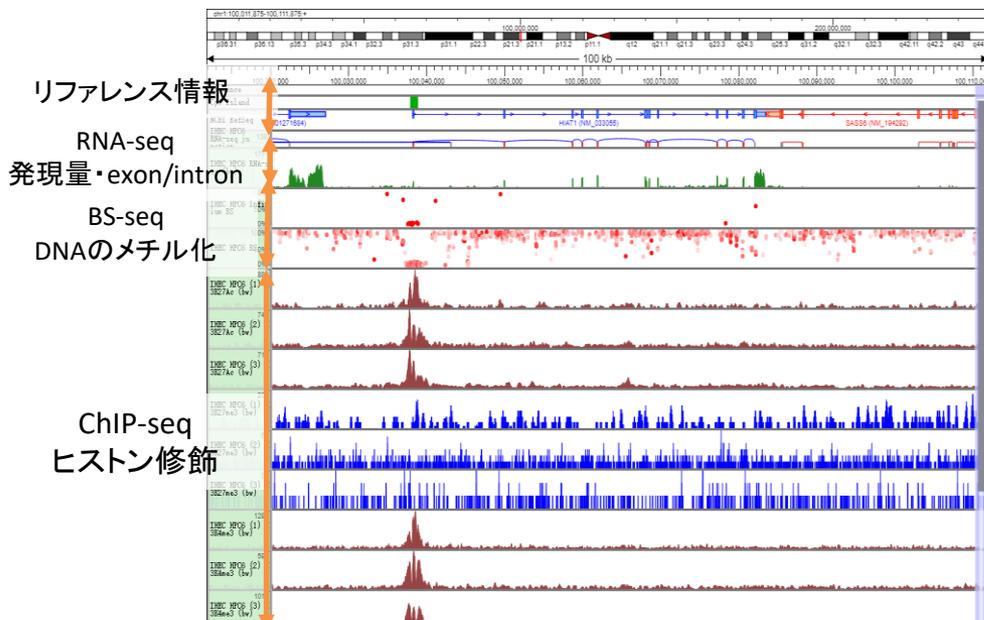
国際エピゲノムコンソシアム (IHEC)

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。

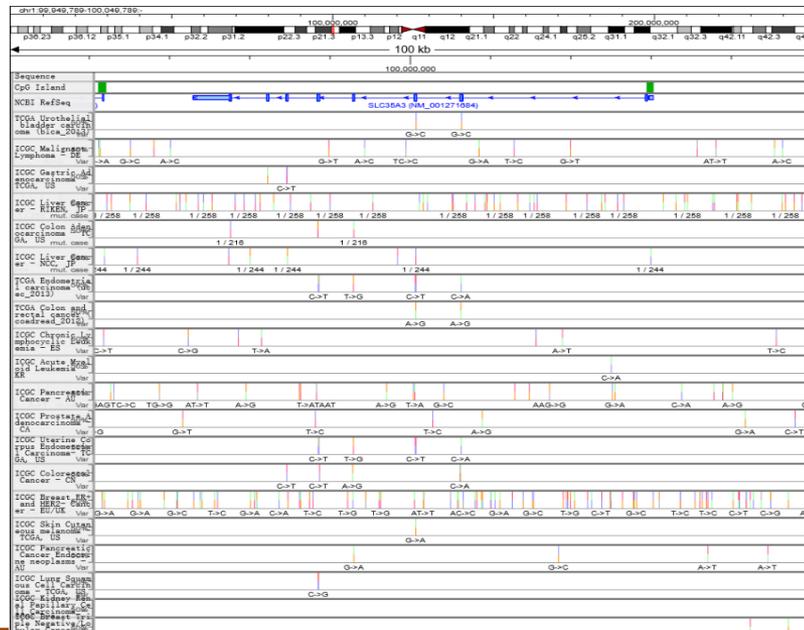
CREST-IHECデータの統合(日本人標準エピゲノム)

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ム情報の表示



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一覧

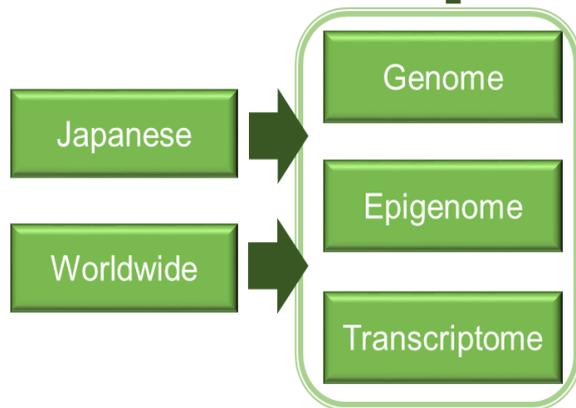
Sample	individuals	Expression		DNA methylation	Histone modifications					
		RNA-seq		BS-seq	H3K4me1	H3K4me3	H3K27ac	H3K27me3	H3K36me3	H3K9me3
		Read depth	Exon/intron junction							
Liver	8	8	8	8	13	13	13	13	13	13
Colon	11	11	11	11	11	11	11	11	11	11
Endometrial	15	30	0	13	14	15	15	15	15	15
Vascular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

Standard multi-omics data



Cell lines
Cancer cell lines, ENCODE, CCLE

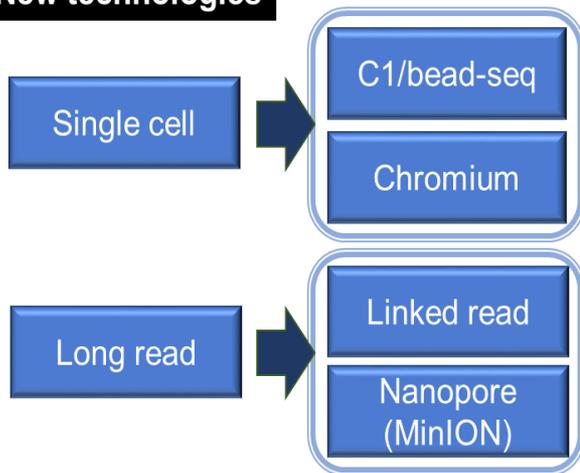
Clinical samples
TCGA, ICGC, COSMIC



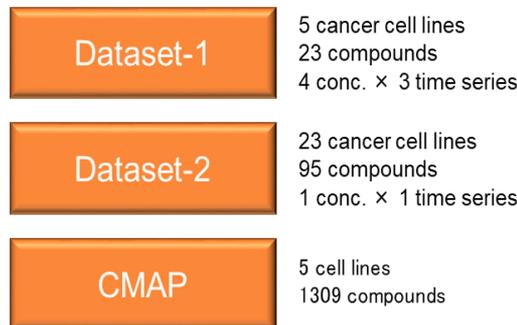
Genome
Japanese SNPs (bio-bank cohorts), HGVD (Japanese GWAS), dbSNP, NHLBI GO-ESP, 1000 Genomes Project, ClinVar, ExAc

Epigenome/transcriptome
IHEC, Roadmap

New technologies



Drug perturbation



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ミックス擾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

Release 9.0 Updated (July 9, 2016)
Based on UCSC hg19 build

About this database

Welcome to **DBTSS (Database of Transcriptional Start Sites)**

To support transcriptional regulation studies, we have constructed the DBTSS (Database of Transcriptional Start Sites), which represents exact positions of transcriptional start sites (TSSs) in the genome based on our unique experimentally validated TSS sequencing method, TSS-seq.

This database includes TSS data of a major part of human adult and embryonic tissues are covered. DBTSS now contains 451 million TSS tag sequences for collected from a total of 20 tissues and 7 cell cultures. We also integrated our newly generated RNA-seq data of subcellular-fractionated RNA and ChIP-seq data of histone modifications, RNA polymerase II and several transcription regulatory factors in cultured cell lines. We also included recently accumulating external epigenomic data, such as chromatin map of the ENCODE project.

In this update, we further associated those TSS information with public and original SNV data, in order to identify single nucleotide variations (SNVs) in the regulatory regions.

It is believed that single nucleotide variations (SNVs) in the transcriptional regulatory regions are responsible for many human diseases, including cancers. However, it remains difficult to identify functionally relevant SNVs from those having no explicit biological consequences. In this version of DBTSS, we attempt to associate SNVs with the genomic information of the surrounding regions. We used SNVs which we identified from genomic analyses of various types of cancers, including somatic mutations of 100 lung adenocarcinoma and lung small cell carcinoma. For germline variations, we used SNVs in dbSNP as well as our unique dataset of variations in 1000 Japanese individuals. We integrated those SNV information with our original datasets of TSS-seq, RNA-seq, ChIP-seq of representative histone modifications and Beutelle Sequencing of cytosine methylations of DNA. Particular, we present multi-omics data of 28 lung adenocarcinoma cells line for which TSS-seq, RNA-seq, ChIP-seq and BS-seq together with whole genome sequencing are collected from the same materials. We further connected the multi-omics data of model organisms by genome-genome alignment. We provide a unique data resource to investigate what genomic features are observed in a particular genomic coordinates in a wide variety of samples.

News

- 09 Jul 2015: New 7 helper cell data of mouse are now available. Raw data accession: [DRAB00581](#) ([GEMM1102](#)) and [SRR007385](#) ([Genome Shen](#))
- 30 Jun 2015: New BRIC-Seq data (UPF1 RNA) are now available. Raw data accession: [DRAB00551](#) ([Genome Shen](#))
- 15 Sep 2014: New DBTSS opened.

Database Search

Keyword Search

Species: H. sapiens

Keyword: NM_

Search

Limit over: hg19

Search

Search from Genomic Position: chr1:90,000,000-90,000,000

Search

Search from SNP (dbSNP rsID): rs7522969

Search

Search from SNV (COSMIC):

Human Genome Variation DB

Release 9.0 Updated (July 9, 2016)
Based on UCSC hg19 build

Database Search

Keyword Search

Species: H. sapiens

Keyword: NM_

Search

Limit over: hg19

Search

Search from Genomic Position: chr1:90,000,000-90,000,000

Search

Search from SNP (dbSNP rsID): rs7522969

Search

Search from SNV (COSMIC):

Term	Genomic position	TSS viewer	Genome viewer	Human Variation DB
NM_000014	chr12:9,007,708-9,115,902	TSS viewer	Genome viewer	Human Variation DB
NM_000015	chr8:19,397,240-19,401,213	TSS viewer	Genome viewer	Human Variation DB
NM_000016	chr1:15,724,147-15,763,078	TSS viewer	Genome viewer	Human Variation DB
NM_000017	chr12:120,725,766-120,740,008	TSS viewer	Genome viewer	Human Variation DB
NM_000018	chr17:2,219,817-2,257,307	TSS viewer	Genome viewer	Human Variation DB
NM_000019	chr11:108,121,531-108,148,104	TSS viewer	Genome viewer	Human Variation DB
NM_000020	chr12:1,607,418-1,623,361	TSS viewer	Genome viewer	Human Variation DB
NM_000021	chr14:73,138,435-73,223,691	TSS viewer	Genome viewer	Human Variation DB
NM_000022	chr20:44,619,522-44,651,735	TSS viewer	Genome viewer	Human Variation DB
NM_000023	chr17:50,196,005-50,175,932	TSS viewer	Genome viewer	Human Variation DB
NM_000024	chr5:149,826,550-149,828,634	TSS viewer	Genome viewer	Human Variation DB
NM_000025	chr8:37,942,996-37,969,696	TSS viewer	Genome viewer	Human Variation DB
NM_000026	chr22:40,348,500-40,396,571	TSS viewer	Genome viewer	Human Variation DB
NM_000027	chr1:17,418,775-17,442,803	TSS viewer	Genome viewer	Human Variation DB
NM_000028	chr19:95,869,869-95,924,023	TSS viewer	Genome viewer	Human Variation DB
NM_000029	chr1:230,702,523-230,714,560	TSS viewer	Genome viewer	Human Variation DB
NM_000030	chr2:240,869,745-240,879,119	TSS viewer	Genome viewer	Human Variation DB
NM_000031	chr9:113,386,312-113,401,338	TSS viewer	Genome viewer	Human Variation DB
NM_000032	chr5:55,009,055-55,014,064	TSS viewer	Genome viewer	Human Variation DB
NM_000033	chr8:153,724,868-153,744,702	TSS viewer	Genome viewer	Human Variation DB
NM_000034	chr16:30,053,000-30,070,420	TSS viewer	Genome viewer	Human Variation DB
NM_000035	chr8:101,420,560-101,435,780	TSS viewer	Genome viewer	Human Variation DB
NM_000036	chr1:114,873,096-114,865,618	TSS viewer	Genome viewer	Human Variation DB
NM_000037	chr1:41,653,225-41,707,622	TSS viewer	Genome viewer	Human Variation DB
NM_000038	chr5:112,737,858-112,846,230	TSS viewer	Genome viewer	Human Variation DB

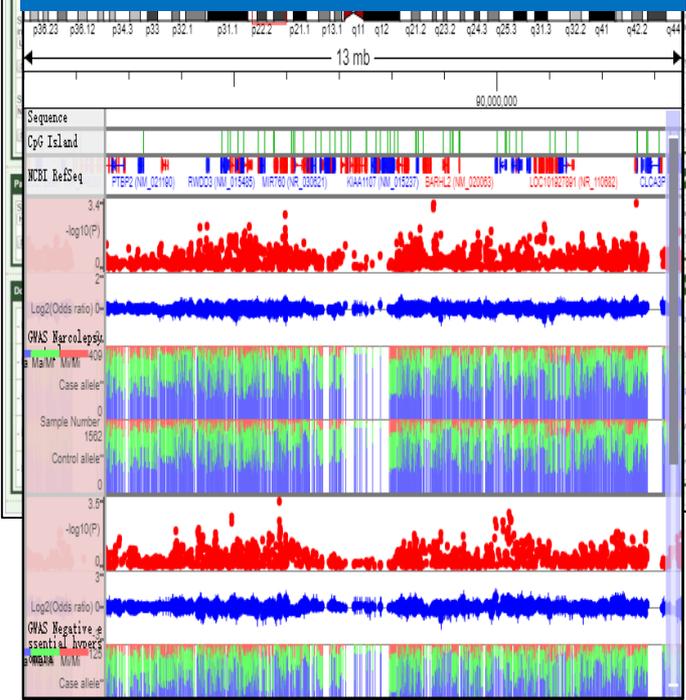
HGVDへのリンク機能

Human Variation DB

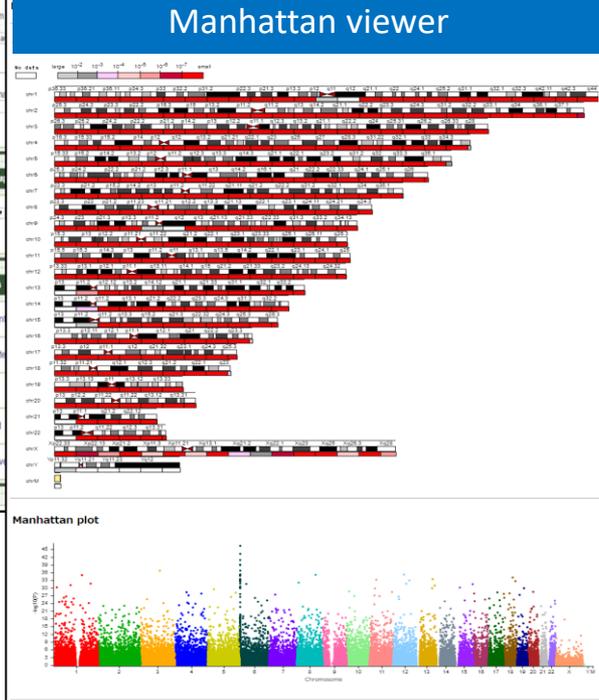
Search

Human Variation DB

Genome browserへGWAS情報の可視化



Whole genome viewer, Manhattan viewer



Genomic position	rs ID	NM change info	Amino acid change	Gene ID	Gene name	Hetero/Homo	Disease	V-ID	Case/Control with this mutation	P value	OR(95%CI)	Type of studies	
12:52301883	rs167099243	C	G	100	5	173	0.028	C	C	C	C	Japanese_TKPU	HKFU
12:52301818	rs3782479	C	A	100	26	152	0.148	A	A	A	A	Japanese_TKPU	HKFU
12:52302217	rs7989275	C	T	100	26	152	0.148	C	C	C	C	Japanese_TKPU	HKFU
12:52302304	rs55929712	A	T	100	2	176	0.011	A	A	A	A	Japanese_TKPU	HKFU
12:52302338	rs77898410	G	A	100	38	140	0.213	A	A	A	A	Japanese_TKPU	HKFU
12:52302443	rs7956340	T	G	100	26	152	0.148	A	A	A	A	Japanese_TKPU	HKFU
12:52302742	rs11189951	C	T	100	27	151	0.152	C	C	C	C	Japanese_TKPU	HKFU
12:52302869	rs78782411	T	C	100	26	152	0.148	A	A	A	A	Japanese_TKPU	HKFU
12:52304399	rs11189953	C	T	100	110	68	0.618	C	C	C	C	Japanese_TKPU	HKFU
12:52304791	rs14379294	T	G	100	4	174	0.022	C	C	C	C	Japanese_TKPU	HKFU
12:52305069	rs3782480	G	T	100	46	132	0.255	C	C	C	C	Japanese_TKPU	HKFU
12:52305700	rs12578436	C	A	100	47	131	0.264	A	A	A	A	Japanese_TKPU	HKFU

NBDC ヒトDBとの連携へ

何ができるか(2)

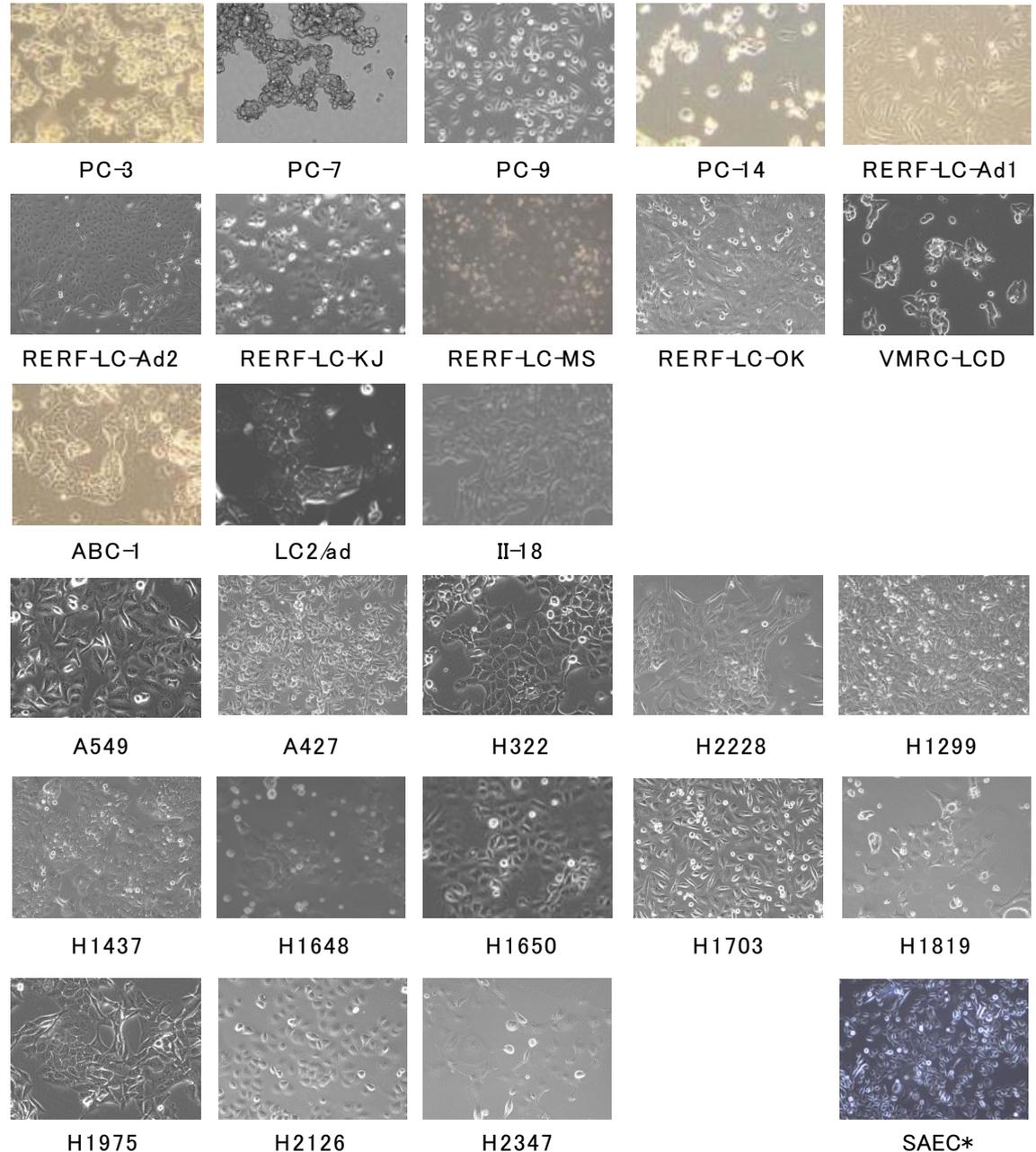
どの培養細胞をモデルにするのか？
を選ぶオーミクスデータ

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	
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

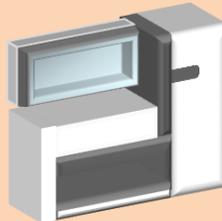
- Whole-genome sequencing
- Long-read sequencing (10X/MinION)

Epigenome

- Bisulfite sequencing
- ChIP-Seq
- ATAC-Seq

Transcriptome

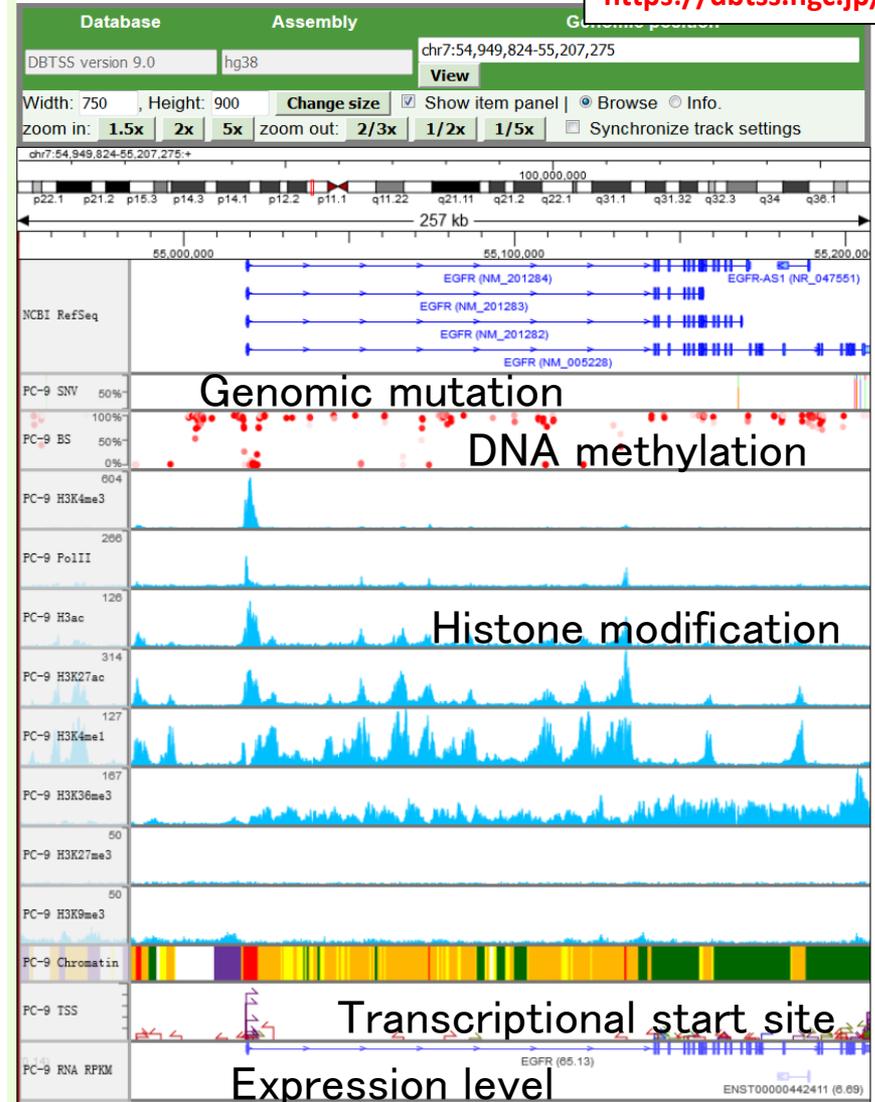
- TSS-Seq
- RNA-Seq
- miRNA-seq



Genome viewer (Multi-omics Data)

DBTSS

<https://dbtss.hgc.jp/>



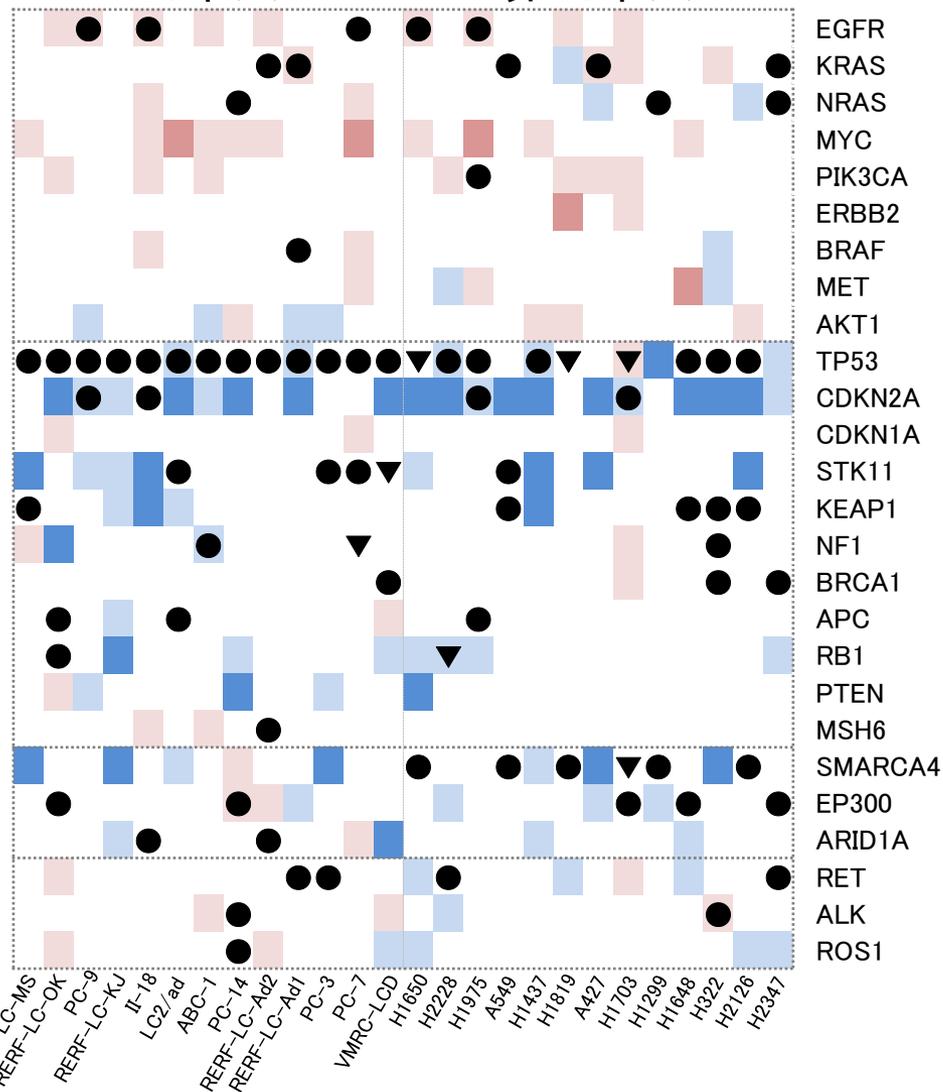
Suzuki *et al.* 2014 *Nucleic Acids Research*

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

日本人

非日本人



ドライバー
変異

- Non-synonymous SNVs/short indels on CDS
- ▼ SNVs/short indels on splice sites
- Highly copy number gains
- Copy number gains
- Homo losses /large deletions (>1 Kb)
- Copy number losses

がん抑制
遺伝子

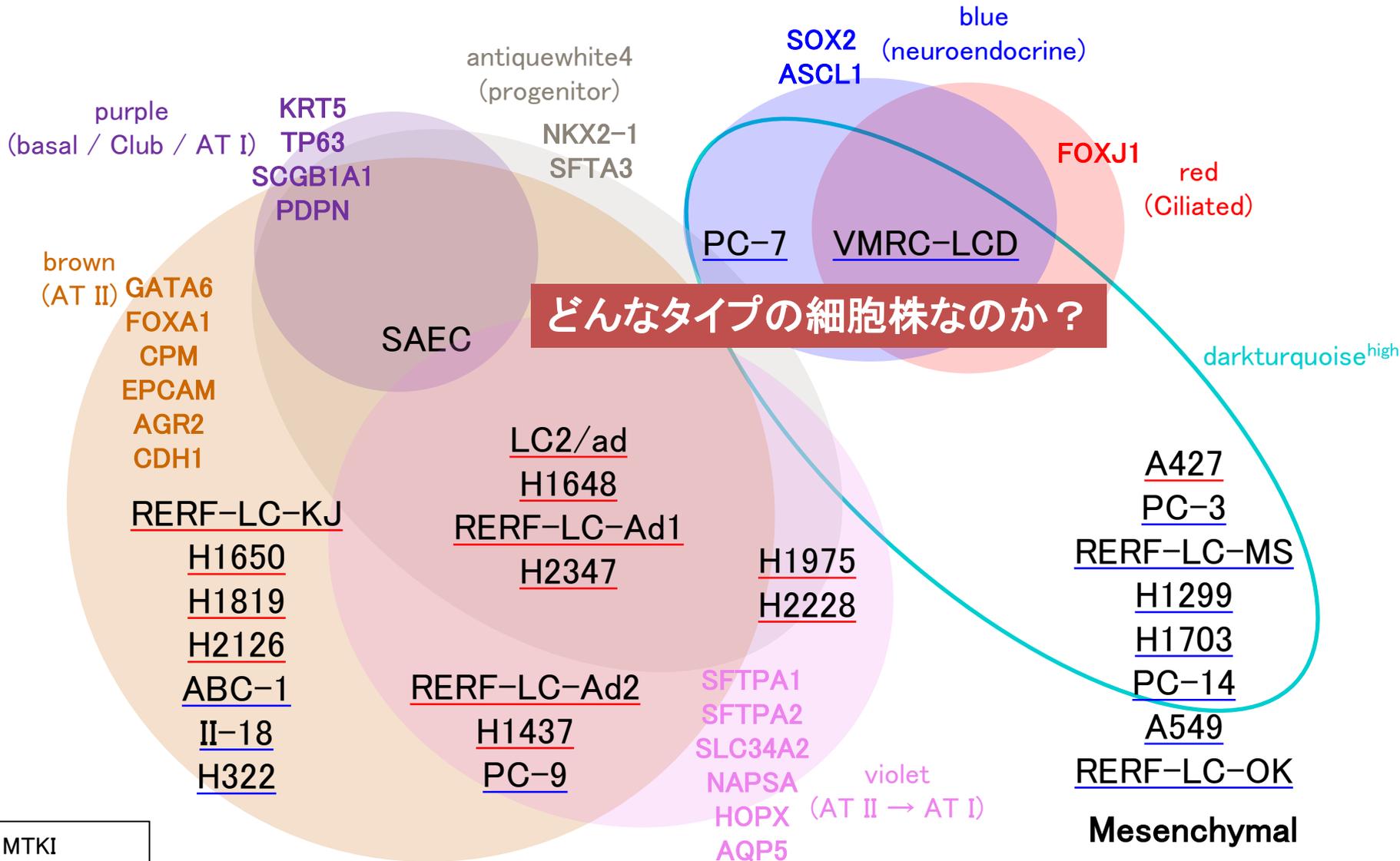
がん関連
遺伝子

融合遺伝子



肺腺癌細胞株のTranscriptome/Epigenome層別化

Pan-modules



どんなタイプの細胞株なのか？

代表的な薬剤をかけるとどう動くのか？

- MTKI
- Sensitive
 - Resistant

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

LC-SCRUM-Japan
(151 institutes in 46 prefectures participating as of Dec. 18th 2013)

Clinical sequencing also started in Japan

JSPS 文部科学省科学研究費助成事業「新学術領域研究「学術研究支援基盤形成」

先進ゲノム解析研究推進プラットフォーム

先進ゲノム支援

Genome-Aids

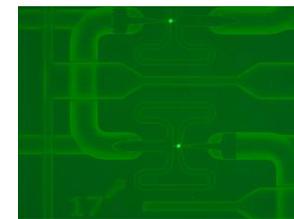
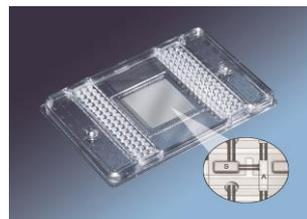
こちらにもブースにおこしく下さい。

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)

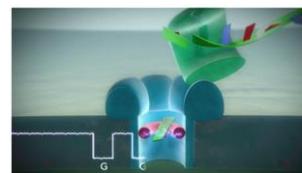
“Chip” of C1

Image of single-cell capture on C1



PI: Yutaka Suzuki.
Univ. Tokyo

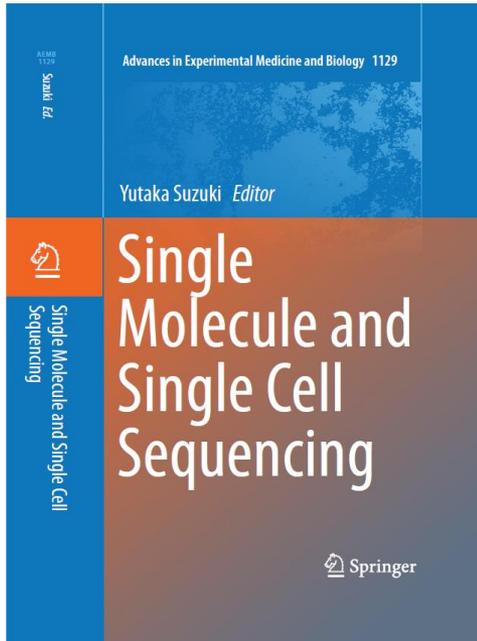
Operators:
Technicians 5
Programmers 4



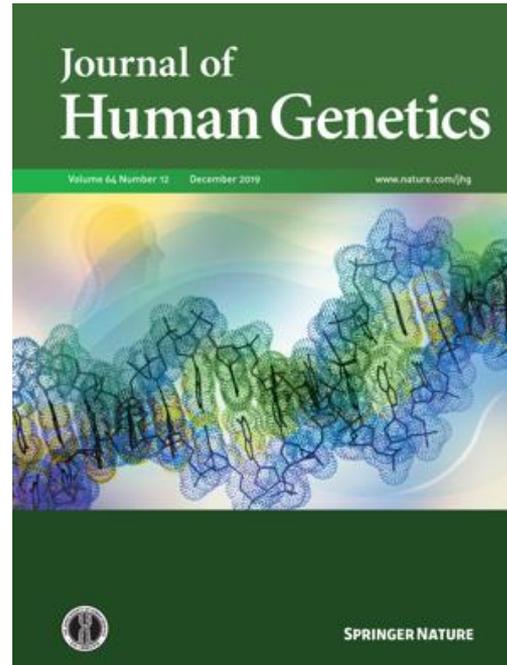
Beta-test of nanopore sequencer

ysuzuki@hgc.jp

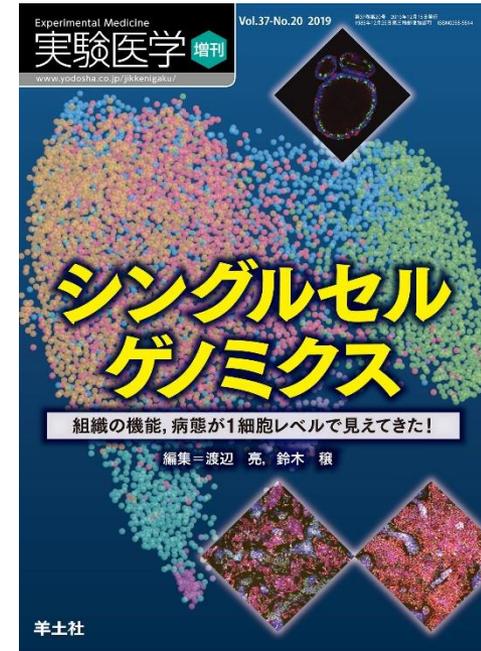
参考文献(シングルセル+ロングリード)



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



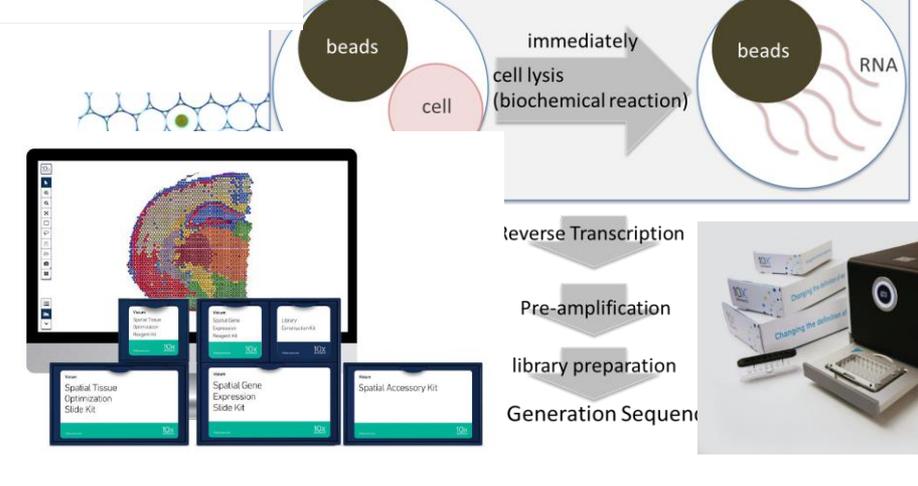
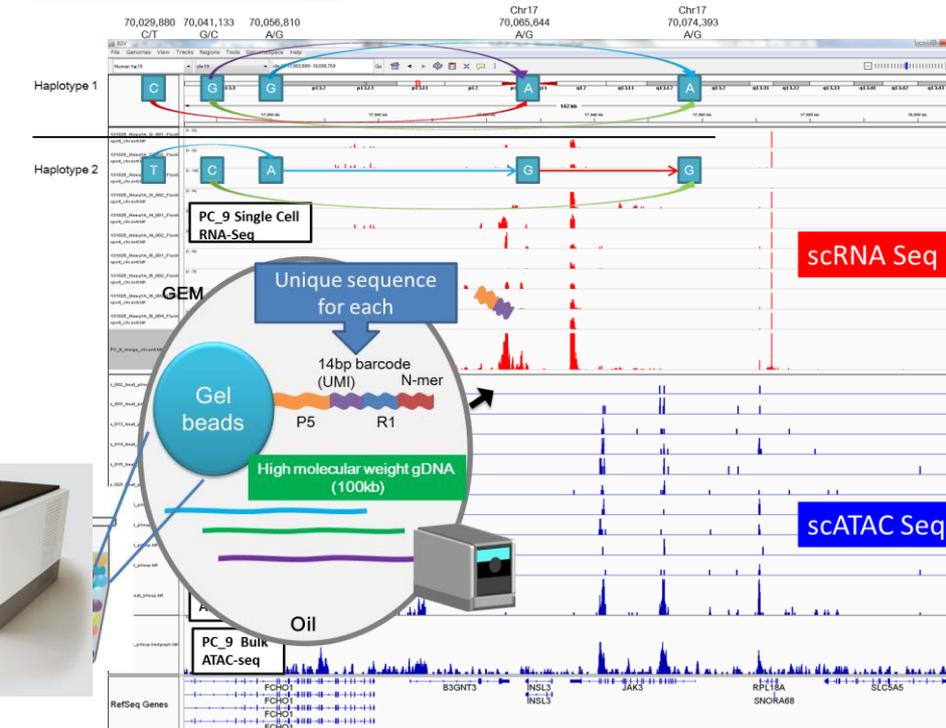
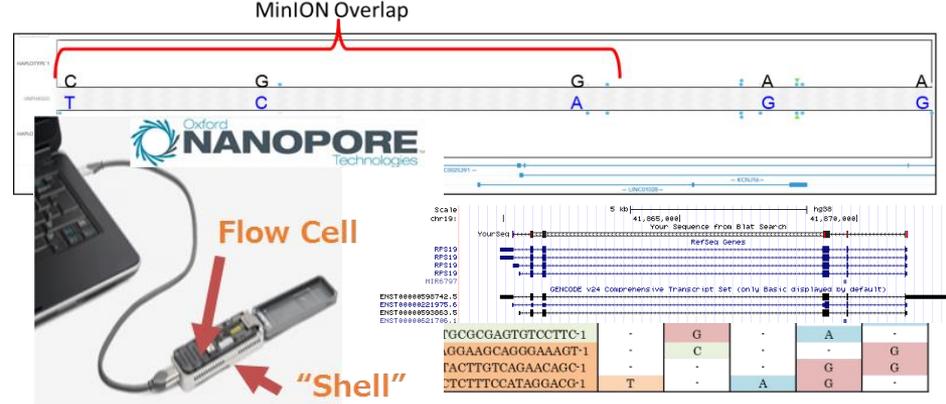
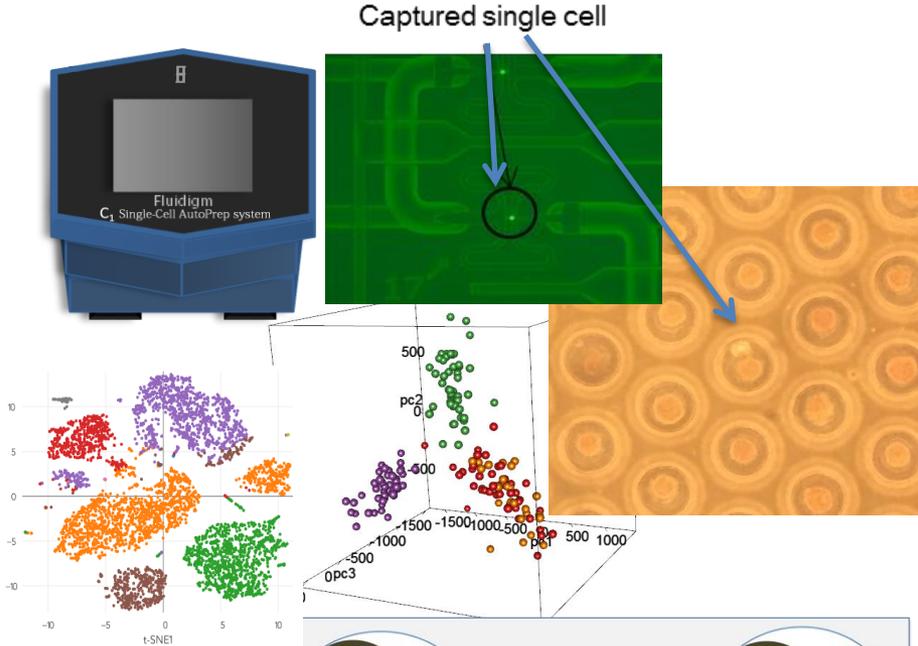
Journal of Human Genetics
Special issue for “New type Sequencer”
Jan (2020)
(Yutaka Suzuki ed.)



本日先行発売!

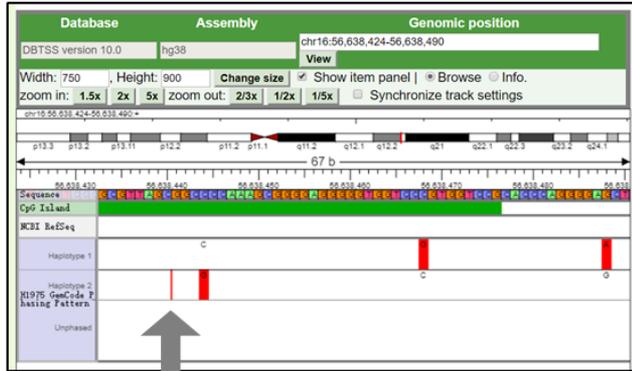
新しい手法でどんなデータが出てくるの？

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

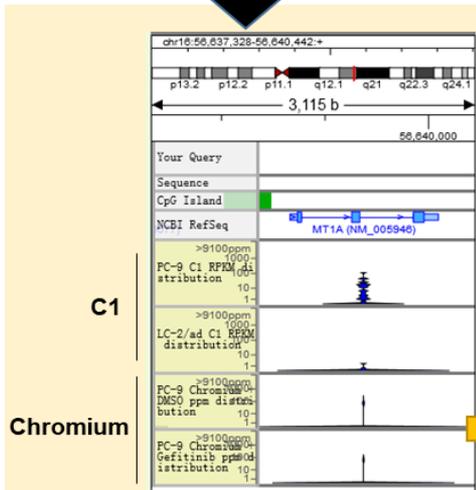


シングルセルとロングリードのデータ閲覧

MT1A enhancer in H1975
chr16:56,638,440, C>CG



Mutation

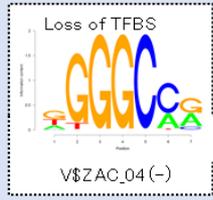
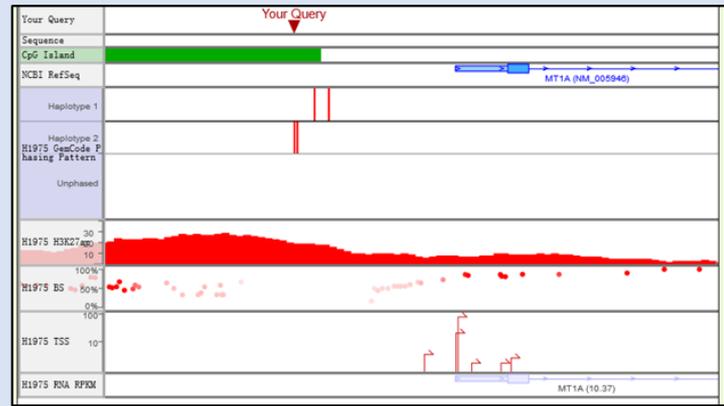


Chromium

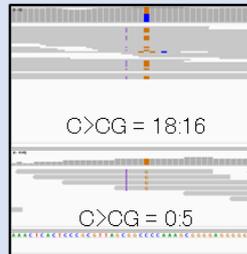


Linked read

H3K27ac
DNA met
TSS-seq
RNA-seq

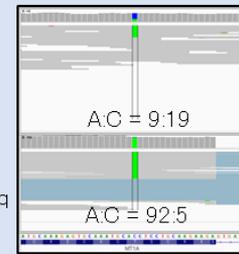


WGS

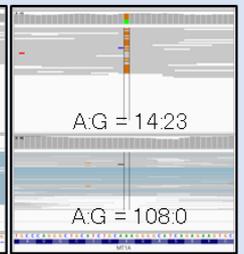


Regulatory mutation

WGS



RNA-seq



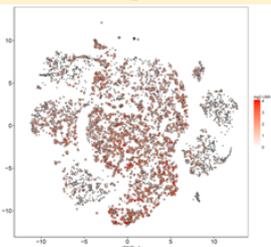
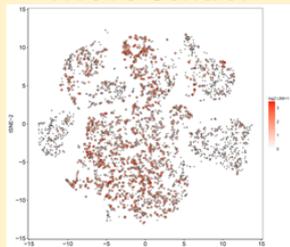
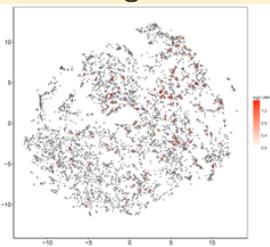
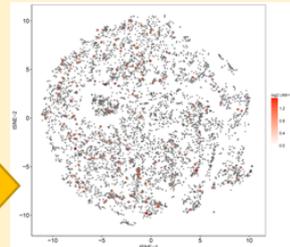
Coding SNPs

PC-9 control

PC-9 gefitinib

H1975 control

H1975 gefitinib



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/



We recommend 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

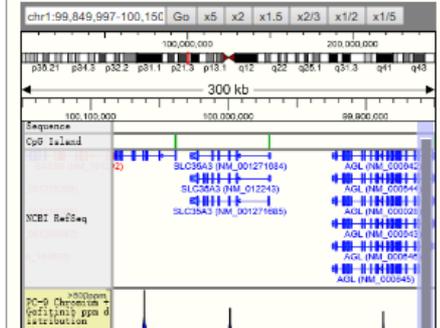
- [Experimental Procedures](#)
- [Data Contents](#)
- [Help](#)
- [Download](#)
- [Links](#)
- [References](#)

Genome browser

Display tracks

- [Single cell](#)
- [Long read](#)
- [Multi-omics](#)
- [Cancer genome](#)
- [Disease genome](#)

Demo



データポータル
(準備中)

Movie

Overview movie in Japanese (53 min.)



http://kero.hgc.jp/

DBKERO Data Portal (Sample)

Project:

Race:

Omics Category:

Show 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
2327	wg_bw_H2328	H2328	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

初めてのの方は、**"For the first users"へ**

We recommend 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

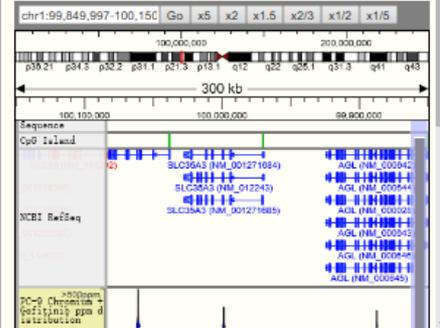
- [Experimental Procedures](#)
- [Data Contents](#)
- [Help](#)
- [Download](#)
- [Links](#)
- [References](#)

Genome browser

Display tracks

- [Single cell](#)
- [Long read](#)
- [Multi-omics](#)
- [Cancer genome](#)
- [Disease genome](#)

Demo



Movie

Overview movie in Japanese (53 min.)



とりあえず
ブラウザへ
クリック!

使い方の説明

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。

Tutorialの動画

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