Building the Human Cell Atlas
towards Medical Innovations

Jay W. Shin, PhD
@JayShin
Laboratory for Advanced Genomics Circuit
RIKEN IMS

2nd PSTC Japan Safety Biomarker Conference
April 17, 2019
The first comprehensive map of gene activity across the human body

- Comprehensive analysis of genomic regions involved in transcriptional regulation in various cells including normal cells and various tissues
- Mapped the sets of transcripts, transcription factors, promoters and enhancers active in the majority of mammalian primary cell types and a series of cancer cell lines, and tissue

https://www.nature.com/collections/fantom5
CAGE - Cap Analysis of Gene Expression
Capturing and Mapping the 5’-end of RNA

[Diagram showing RNA processing and expression levels in healthy and cancer cells]

Piero Carninci, RIKEN

https://cage-seq.com
Composite Promoter Architecture

B4GALT1

Combined profile of all human samples

CD4+ T-cells

CD14+ Monocytes

Aortic Smooth Muscle Cells

Carcinoma
### Annual Health Check 2012 July 18

<table>
<thead>
<tr>
<th>項目</th>
<th>單位</th>
<th>基準範圍</th>
<th>今 回</th>
<th>前 回</th>
<th>前 前 回</th>
</tr>
</thead>
<tbody>
<tr>
<td>身 長</td>
<td>cm</td>
<td></td>
<td>24.7.18</td>
<td>23.7.20</td>
<td>22.8.4</td>
</tr>
<tr>
<td>体 重</td>
<td>kg</td>
<td></td>
<td>178.2</td>
<td>178.2</td>
<td>178.5</td>
</tr>
<tr>
<td>B M I</td>
<td></td>
<td>18.5~24.9</td>
<td>22.5</td>
<td>23.9</td>
<td>23.9</td>
</tr>
<tr>
<td>腹 囲</td>
<td>cm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>脂肪度</td>
<td>%</td>
<td>-10~10</td>
<td>2</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>視 力</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>右</td>
<td></td>
<td>(1.5)</td>
<td></td>
<td></td>
<td>(1.0)</td>
</tr>
<tr>
<td>左</td>
<td></td>
<td>(1.2)</td>
<td></td>
<td></td>
<td>(1.0)</td>
</tr>
<tr>
<td>聴 力</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>右1000Hz</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4000Hz</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>左1000Hz</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4000Hz</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>血压</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 回目</td>
<td>mmHg</td>
<td>139以下 - 89以下</td>
<td>116 - 81</td>
<td>110 - 62</td>
<td>115 - 63</td>
</tr>
<tr>
<td>2 回目</td>
<td>mmHg</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>腎尿素窒素</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>細胞</td>
<td>mg/dl</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>脂肪</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>細胞</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>白 血球計数</td>
<td>/ µl</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>血小板数</td>
<td>/ µl</td>
<td>8~33</td>
<td>* 412</td>
<td>15</td>
<td>24</td>
</tr>
<tr>
<td>AST (GOT)</td>
<td>IU/l</td>
<td>6~32</td>
<td>* 100</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>ALT (GPT)</td>
<td>IU/l</td>
<td>49以下</td>
<td></td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>γ-G T P</td>
<td>IU/l</td>
<td></td>
<td></td>
<td>152</td>
<td>18</td>
</tr>
<tr>
<td>Z T T</td>
<td>KU</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDLコレステロール</td>
<td>mg/dl</td>
<td>140未満</td>
<td>96</td>
<td>125</td>
<td>122</td>
</tr>
<tr>
<td>HDLコレステロール</td>
<td>mg/dl</td>
<td>40以上</td>
<td>48</td>
<td>44</td>
<td>45</td>
</tr>
<tr>
<td>トリグリセライド</td>
<td>mg/dl</td>
<td>50~149</td>
<td>69</td>
<td>86</td>
<td>77</td>
</tr>
<tr>
<td>総コレステロール</td>
<td>mg/dl</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>尿 酸</td>
<td>mg/dl</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Annual Health Check

**2012 July 18**

<table>
<thead>
<tr>
<th>項目</th>
<th>営業</th>
<th>基準範囲</th>
<th>今回</th>
<th>前回</th>
<th>前々回</th>
</tr>
</thead>
<tbody>
<tr>
<td>身長</td>
<td>cm</td>
<td>18.5〜24.9</td>
<td>24.7</td>
<td>23.7</td>
<td>22.8</td>
</tr>
<tr>
<td>体重</td>
<td>kg</td>
<td>71.4</td>
<td>178.2</td>
<td>178.2</td>
<td>178.5</td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td>22.5</td>
<td>23.9</td>
<td>23.9</td>
<td>23.9</td>
</tr>
<tr>
<td>肥満度</td>
<td>%</td>
<td>-10〜10</td>
<td>2</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>視力</td>
<td>右</td>
<td>(1.5)</td>
<td>(1.0)</td>
<td>(1.0)</td>
<td>(1.0)</td>
</tr>
<tr>
<td></td>
<td>左</td>
<td>(1.2)</td>
<td>(1.2)</td>
<td>(1.2)</td>
<td>(1.2)</td>
</tr>
<tr>
<td>聴力</td>
<td>右</td>
<td>1000Hz, 4000Hz</td>
<td>正常</td>
<td>正常</td>
<td>正常</td>
</tr>
<tr>
<td></td>
<td>左</td>
<td>1000Hz, 4000Hz</td>
<td>正常</td>
<td>正常</td>
<td>正常</td>
</tr>
<tr>
<td>血圧</td>
<td>1回目</td>
<td>mmHg</td>
<td>139以下〜89以下</td>
<td>116〜81</td>
<td>110〜62</td>
</tr>
<tr>
<td></td>
<td>2回目</td>
<td>mmHg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>尿素窒素</td>
<td>mg/dl</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>クレアチニン</td>
<td>mg/dl</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>血小板数</td>
<td>/μl</td>
<td>8〜33</td>
<td>*</td>
<td>412</td>
<td></td>
</tr>
<tr>
<td>AST(GOT)</td>
<td>IU/l</td>
<td>6〜32</td>
<td>*</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>ALT(GPT)</td>
<td>IU/l</td>
<td></td>
<td>49以下</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Z TT KU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDLコレステロール</td>
<td>mg/dl</td>
<td>140未満</td>
<td>96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDLコレステロール</td>
<td>mg/dl</td>
<td>40以上</td>
<td>48</td>
<td>44</td>
<td>45</td>
</tr>
<tr>
<td>リンクレアチニン</td>
<td>mg/dl</td>
<td>50〜149</td>
<td>69</td>
<td>86</td>
<td>77</td>
</tr>
<tr>
<td>食塩</td>
<td>mg/dl</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
GOT1 expression

FANTOM5 CAGE phase 1 and 2 human hg38 (q20 TPM, min 1TPM CTSS)

- occipital pole, adult, pool1 : CNhs10643 cts [q20_tpm]
- pons, adult, pool1 : CNhs10640 cts [q20_tpm]
- heart, fetal, pool1 : CNhs10653 cts [q20_tpm]
- heart, adult, diseased, donor1 : CNhs11758 cts [q20_tpm]
- heart, adult, pool1 : CNhs10621 cts [q20_tpm]
- paracentral gyrus, adult, pool1 : CNhs10642 cts [q20_tpm]
- postcentral gyrus, adult, pool1 : CNhs10638 cts [q20_tpm]
- parietal lobe, adult, pool1 : CNhs10641 cts [q20_tpm]
- Keratocytes, donor1 : CNhs11337 cts [q20_tpm]
- temporal lobe, adult, pool1 : CNhs10637 cts [q20_tpm]
- melanoma cell line: COLO 679 : CNhs11281 cts [q20_tpm]
- signet ring carcinoma cell line: Kato III : CNhs10753 cts [q20_tpm]
- oral squamous cell carcinoma cell line: HO-1-u-1 : CNhs11287 cts [q20_tpm]
- Insula, adult, pool1 : CNhs10646 cts [q20_tpm]
- frontal lobe, adult, pool1 : CNhs10647 cts [q20_tpm]
- Fibroblast - Conjunctival, donor1 : CNhs11339 cts [q20_tpm]
- Nucleus accumbens, adult, pool1 : CNhs10644 cts [q20_tpm]

- liver, adult, pool1 : CNhs10624 cts [q20_tpm]
- heart, adult, diseased post-infarction, donor1 : CNhs11757 cts [q20_tpm]

- brain, adult, pool1 : CNhs10617 cts [q20_tpm]
- bile duct carcinoma cell line: TFK-1 : CNhs11285 cts [q20_tpm]
- gastric cancer cell line: A5221 : CNhs11286 cts [q20_tpm]
- left ventricle, adult, donor1 : CNhs11789 cts [q20_tpm]
- lung adenocarcinoma cell line: A549 : CNhs11275 cts [q20_tpm]
- extraskeletal myxoid chondrosarcoma cell line: H-EMC-SS : CNhs10728 cts [q20_tpm]
- Keratinocyte - epidermal, donor2 : CNhs11381 cts [q20_tpm]
- K562 erythroleukemia leukemia response to hemin, 01hr00min, biol rep2 : CNhs12689 cts [q20_tpm]
- ductal cell carcinoma cell line: KLM-1 : CNhs11100 cts [q20_tpm]

- skeletal muscle, adult, pool1 : CNhs10629 cts [q20_tpm]
- Retinal Pigment Epithelial Cells, donor3 : CNhs12733 cts [q20_tpm]

- medial frontal gyrus, adult, donor1258 : CNhs14221 cts [q20_tpm]
- squamous cell carcinoma cell line: JHU-nsk1 : CNhs11749 cts [q20_tpm]
Genes are controlled by complex regulatory network across different cells
Structure and function of long noncoding RNAs in epigenetic regulation. TR Mercer & JS Mattick 2013
CAGE-Associated Transcripts (CAT)

Number of Human IncRNA Genes

27,919

63,132 permissive
9,318 stringent

FANTOM CAT

Chung-Chau Hon et al., Nature 543, 199 (2017)
Essential Mediators of Life

Knockdown of lncRNA is context dependent

In healthy human dermal fibroblasts

Ramilowski J., Yip Chi-Wai, Michiel de Hoon, Jay W. Shin, Piero Carninci, FANTOM6 pilot study

http://fantom.gsc.riken.jp/cat
GOT1 ENSG00000120053.9

Associated with hepatocyte.; conserved TIR and conserved exon; Best associated with liver disease, among total 2 trait associations.; Best implicated in Decreased liver function vs hepatocyte, among total 1 trait vs sample ontology pairs.

Sample Ontology Association

Dynamic Expression

GOT1 viral infection
RPV infect 293SLAM
MEG3 ENSG00000214548.10

Most coexpressed with ENSG00000100811.6, among total 1 eQTL linked co-expressed mRNAs.; Best associated with pituitary gland, among total 21 sample ontology associations.; conserved TIR and conserved exon; Best associated with Type 1 diabetes, among total 2 trait associations.; No implications in trait vs sample ontology pairs.

Basic Information
- Official Name: maternally expressed 3 (non-protein coding)
- Official Symbol: MEG3
- Gene type: IncRNA intergenic

Sequence Conservation
- TIR: conservation
- exon: conserved

Sample Ontology Association

Dynamic Expression
- MEG3 iPSC differentiation
- iPSC to neuron wildtype
- log2FC: -5.0 -2.5 0.0 2.5 5.0
- Expression CPM
Summary Part I

- FANTOM5 provides rich database of gene expression across human tissues/cells
  - Single biomarker may not be sufficient
- Gene Signatures or Gene Regulatory Network are essential to understand disease and function
- IncRNAs are integral part of our genome
- FANTOM5 study was done in bulk tissues/cells
Profilng the human body and disease

There are 37 trillion single cells in the adult human body

50+ different organ systems

Elucidating

Genetic associations to disease

Normal bone Osteoporosis

Profile Tissues at the single cell resolution

Kidney

Cortex

Podocyte Capillary Mesangial cell

Kidney Profiling the human body and disease and disorders

Heterogeneity of Cancer

Regenerative Medicine

Genetic associations to disease

Drug development

Basic research Medical community Industry applications

Single Cell Database
Human Cell Atlas Mission

To create a comprehensive reference map of the types and properties of all human cells, the fundamental unit of life, as a basis for understanding, diagnosing, monitoring, and treating health and disease.
The Impact on Science and Society

- This international collaboration will help answer fundamental questions in all aspects of biology as well as serve as a guide to unravel the secrets of human disease.
- The data/technologies from this project will lead it to applications, such as diagnosis and drug discovery, that expand to industry.

**Regenerative medicine**

An atlas of cell types that are lost in disease will enable efforts to generate such cells faithfully. Similarly, an atlas of healthy human tissues and the matching organoids or in vitro differentiated cells will help determine if the engineered samples faithfully represent normal tissue composition and identify ways to complete any missing components.

**Genes to drugs**

The cell atlas will enable researchers to identify the cell types in which a given genetic variant acts, thus helping to pursue therapeutic targets identified by genetic studies of disease.

**Diagnostics**

Knowledge of all the cell types in the body and their role in disease will enable updated and much more powerful versions of common diagnostic tools, such as the Complete Blood Count (CBC) and next-generation biopsy.

**Disease mechanisms**

Because the cell atlas will provide detailed maps of cells and their roles in tissues, researchers will be able to understand the mechanisms underlying any disorder at both the cell and the cellular-ecosystem level.

**Drug discovery**

The cell atlas will provide guidance as to which gene signatures to pursue in drug screens to represent desired cell phenotypes. For example, it can give us a molecular map of which genes and signatures drive cell development and how it goes awry in, say, cancer and provide targets for drug discovery.

**Toxicity**

It will be possible to determine where else in the body a particular gene is expressed, helping to identify potential off-target effects prior to drug trials.

**Drug efficacy and resistance**

The atlas will provide the tools necessary to understand why drugs work — or don’t — at the level of cells and tissues, both prior to and after treatment.
Single Cell Profiling Technologies are growing rapidly

A. Technology forum

B. Graph showing the growth of single cell profiling technologies

- Manual
- Multiplexing
- Integrated fluidic circuits
- Liquid handling robotics
- Nanodroplets
- Picowells
- In situ barcoding

Graph:

- scRNA-Seq studies
- Single-cell chromatin accessibility studies

- Tang et al.
- STRT-Seq
- SMART-Seq
- CEL-Seq
- Fluidigm C1
- MARS-Seq
- CytoSeq
- Drop-Seq
- inDrop
- 10x Genomics
- SPLiT-Seq
- sci-RNA-Seq
- DroNc-Seq
- Seq-Well

Study publication date:

2009: Tang et al.
2010: STRT-Seq, SMART-Seq, CEL-Seq
2011: Fluidigm C1
2012: MARS-Seq, CytoSeq, Drop-Seq
2013: inDrop
2014: 10x Genomics, SPLiT-Seq, sci-RNA-Seq, DroNc-Seq, Seq-Well

HCA white paper 2017
**Single-Cell CAGE**

**Fluidigm C1**

**10x Genomics 5'-seq**

Overview

C1 CAGE is a method for single-cell transcriptome analysis for molecular counting of RNA 5'-ends. Paired-end sequencing, random priming, and unique molecular identifiers are used for single-molecule fragment assembly of miRNAs and long non-coding RNAs, including non-polyadenylated transcripts.

**Protocol: C1 CAGE**

Duration (H:M): 12:00

- **Cell Load**
  - Capture
  - Stain
  - Wash
- **Sample Prep**
  - Lysis
  - RT
  - PCR
  - Harvest

Tsukasa Kouno, et al. *Nature Communications* 2019
Precise mapping of TSS based on single cell CAGE

* Promoter activity
* lncRNA, e.g. enhancer RNAs
* TCR/BCR profiling
Building the Gene Regulatory Network by inferring transcription factor activity

MYC’s DNA recognition sequence:
GGGCGCGAAAA
Many disease associated variants are found in enhancer elements
Detection of IncRNA such as enhancer RNA at the single cell resolution

eRNA: ADDG12020975224

Bidirectionality of eRNAs in bulk and Single cell CAGE

Establishment of human brain organoids from human iPS cells

Day18

Joachim Luginbuehl
Yan Jun Lan
Combined regulatory network reveals both known and unknown factors involved in neuronal development

Sequential generation of olfactory bulb glutamatergic neurons by Neurog2-expressing precursor cells.


There are 37 trillion single cells in the adult human body.

50+ different organ systems

Kidney

Elucidating

Brain complexity and disorders

Genetic associations to disease

Heterogeneity of Cancer

Profile Tissues at the single cell resolution

Regenerative Medicine

Drug development

Single Cell Database

Basic research
Medical community
Industry applications

Cancer in lymph nodes

Heart
Membrane around heart

Heart

Diaphragm

Left main bronchus

Lung lining

Chest wall lining

Trachea

Organoids

Cardiac cells
Muscle cells
Intestinal cells
Liver cells
Intestines
Muscle
Heart
Brain

iPSC
International Collaborations for Genome Science

- RIKEN (Japan), Broad Institute (US), Wellcome Sanger Institute (UK), and Karolinska Institute (Sweden) were designated as the core institutes of the project and expected to play a role as the leader of its area.
- RIKEN is invited to be a key organization because we have been contributed to genome science through international collaboration including Human Genome Project, ENCODE and FANTOM projects.
- The contribution of Japan’s Single Cell Genomics technologies to HCA is highly expected.

HCA ASIA
- PIERO CARNINCI
- JAY SHIN

ASIA Members
- Korea
- Singapore
- India

JAPAN Members
- University of Tokyo
- CiRA (Kyoto Uni)
- Kyushu Uni
- Kanazawa Uni
- Ehime Uni
- (CREST/Sakigake project)
And more....
Integration of Data to increase the power of N

Human Cell Atlas: Growing rapidly

54 Countries, 638 Institutes, 1027 members
Data platform: draft design

- Data generators:
  - Lab 1
  - Lab 2
  - Lab 3
  - Lab 4
  - Lab n

- Ingest API:
  - Single ingestion with primary checks
  - Multiple ingestion brokers

- Ingestion and quality assurance:
  - Raw and derived data
  - Derived results

- Multiple external environments for running secondary analysis pipelines:
  - Pipeline 1
  - Pipeline 2
  - ... (run on demand)

- Community agreement

- Multiple portals and tertiary analysis tools
  - Portal 1
  - Portal 2
  - Portal 3
  - Portal 4

- Multiple replicated object stores (GCE/AWS/etc)
  - Obj store 1
  - Obj store 2
  - Obj store 3
  - ... (synced <= 1 day)

- Producer API

- Consumer API

- Syncing protocol

- Prepared jointly by HCA, Broad, UCSC, EBI/Sanger, CZI

- BSD licensed

- Governance:
  - Metadata
  - Data management
  - Analysis pipelines
Data platform: draft design

- Data generators: lab 1, lab 2, lab 3, lab 4, ..., lab n
- Ingest API
  - Multiple ingestion brokers
  - Single ingestion with primary checks
- Store API
  - *producer API*
  - *consumer API*
  - Replicated object stores (GCE/AWS/etc) (synced <= 1 day)
- Pipelines
  - Multiple external environments for running secondary analysis pipelines
  - Continuously running and endorsed by HCA
  - Community agreement
- Ingest
  - Raw and derived data
  - Ingestion and quality assurance
- Store
  - Multiple object stores
- Visualize/Analyze
  - Multiple portals and tertiary analysis tools
  - Multiple portals and secondary analysis pipelines

prepared jointly by HCA Broad UCSC EBI/Sanger CZI
The Human Cell Atlas Initiative: values

- **International** scientific collaboration
- Sample and science **equity**: geographic, gender, and ethnicity
- **Open Access**, with human subjects compliance
- Standards for **quality** and **reproducibility**
- **Flexibility**: intellectually, technologically, analytically
Unmet Seeds & Needs: prime opportunity with single cell technology

1. Unmet Seeds & Needs

- **Research community**
  - University A
  - Research Institute
  - Research Center

- **Clinical community**
  - Clinic
  - Medical university
  - Hospitals

**Lack of coordination to meet the seeds and needs of basic and clinical research (opportunity loss)**

**Clinical institutions**
- Access to human samples and clinical information
- Deep understanding of pathology and disease
- Finds value in genomics but limited time
- Lacks infrastructure and analytical know-hows

**Research Institutes**
- Possess technologies and infrastructure
- Bioinformatic skillset and handling of big data
- Experimental models and techniques
- Lacks human samples

2. Independent and separate database system

- **Data redundancies**
  - Multiple sites producing the same data

- **Inefficiency in cost**
  - And infrastructure

- **Lack of integration of omics data** (e.g. RNA-seq, epigenome, genome, etc)

- **Data protection and closed**

**New Medical + Research Alliance**

**Addressing Clinically Relevant Questions**

**Surgery**

**Biopsies**

**Pathology**

**Single cell technology**

**Bioinformatics**

**Experimental techniques**
HCA Japan Network will work with global HCA

Japan Single Cell Network

- CLINICAL SAMPLES & DATA
  - RESEARCH
  - HOSPITALS

- SINGLE CELL TECHNOLOGY
  - Spatial
  - Genotype

- SINGLE CELL DATABASE
  - scTranscriptome + Genome + images

Full Access to single cell data in Japan for Network members

Working with the International Community

Data Coordination Platform (DCP)
The Human Cell Reference Database

HCA Japan Network Coordination Team will support

Data infrastructure and integration to global DB

Sample SOP

Data analysis and reporting
国際共同研究を通じたオープン・サイエンス

HCA Ethical working group (EWG) の発足

- 2018年11月、「Ethical Working Group」が発足し、HCAプロジェクトにおけるデータ公開ポリシー・インフォームドコンセント・MTAなどに関する基本的なガイドラインについて議論される予定。
- HCAの生命倫理指針は、ゲノム科学とヒューマンヘルスに関するデータに特化し、これらのデータを公開するためのフレームワーク、ならびにデータ標準化を支援するため、2013年に発足した「GA4GH: Global Alliance for Genomics & Health」の枠組みに従う方針が発表された。

データ解析・管理基盤の構築

- HCAプロジェクトに特に関注すべき点として、2017年6月に設立が発表された「Data Coordination Platform」が挙げられる。
- プロジェクトが本格的な稼働する前のタイミングで、産出する大規模なデータを効率的に収集・整理・解析し、用途に応じた使いやすい形でオープンソース化するためのデータセンターモデルとしての取組み。
- Chan Zuckerberg Initiativeによる試みであり、実際の運営を協働担当する、英・EMBL-EBI、米・ブロード研究所、米・UC Santa Cruz Genomics Instituteに対して資金投下がなされている。

研究データの即時公開

- 2017年10月：英・サンガー研究所がコンソーシアム下で最初に収集した100万個のヒト免疫細胞の1細胞遺伝子発現プロファイルのデータセットの構築を発表。翌月には公開された。
- 2018年4月：「HCA Data Portal」というウェブサイトが立ち上がり、ヒト免疫細胞・ヒト脾臓・リンパ節腫瘍モデルマウスを対象とした1細胞RNA-seqデータセットが公開された。米・ブロード研究所および英・サンガー研究所の成果。

これらのデータは、いずれも論文として発表される前に世界中の研究者が自由に利用できる形で公開されている。

実験プロトコールの共有

- ライフサイエンス系の実験プロトコール共有サイト「protocols.io」にて「Human Cell Atlas Method Development Community」というコミュニティが作られ、新規開発された1細胞シーケンス法のみならず、各腎器についての分離法などが共有されている。

レジストリーに登録された「HCAプロジェクト」の産出データが全て登録され、即時公開される。

http://protocols.io/groups/hca

- 2018年8月現在、46の詳細なプロトコールが公開され、144人の研究者が参加している。
Human Cell Atlas: Meetings future

1. Launch
   London, October 2016

2. Technologies
   Stanford, February 2017

3. Computation
   Stockholm, June 2017

4. General
   Rehovot, October 2017

5. General
   Cambridge, 18-19 March 2018

6. General
   Boston, 1-2 November 2018

7. General
   Tokyo, 23-24 May 2019
EMBO Workshop supported by The Company of Biologists

Single cell biology

20 – 22 May 2019 | Plaza Heisei, Tokyo, Japan

ORGANIZERS: Piero Carninci (RIKEN, JP) Jay Shin (RIKEN, JP)

SPEAKERS

Stephen Quake
Stanford and HHMI, US

Ido Amit
Weizmann Institute of Science, IL

Kazuyo Moro
RIKEN/Osaka University, JP

Sarah Teichmann
Wellcome Sanger Institute, UK

Shyam Prabhakar
Genome Institute of Singapore, SG

Yutaka Suzuki
The University of Tokyo, JP

John Marioni
EMBL-EBI, UK

Wataru Fujibuchi
CiRA, Kyoto University, JP

Alicia Oshlack
Murdoch Children’s Research Institute, AU

Caroline Uhler
MIT, US

Peter Kharchenko
Harvard Medical School, US

Aki Minoda
RIKEN, JP

Takanori Takebe
Tokyo Medical and Dental University, JP

Muzlifah Haniffa
Newcastle University, UK

Nikolaus Rajewsky
Max-Delbrück-Center for Molecular Medicine, DE

Orit Rozenblatt-Rosen
Broad Institute, US

Shinichii Hashimoto
Kanazawa University, JP

Joakim Lundeberg
SciLifeLab Stockholm, SE

Timm Schroeder
ETH Zürich, CH

Lucas Pelkmans
University of Zürich, CH

Itoshi Nikaido
RIKEN, JP

Dana Pe’er
Sloan Kettering Institute, US