Biobank network infrastructure in Japan for improving human health

Soichi Ogishima
Dept. of Informatics for Genomic Medicine, Tohoku Medical Megabank Organization, Tohoku University
Genomics to health
“Improving the health of all people”

Fig 2 The future of genomics rests on the foundation of the Human Genome Project.

Development of biobanks for improving human health

as a trusted infrastructure collecting genomics, phenotypic and life-style data (exposome data)
Tohoku Medical Megabank Project

Prospective genome cohort

Collection of life-course data

Healthy people

150,000

Follow-up study

Biobank

Focusing on the common diseases that involve an interplay b/w genetic and environmental events

Gene

Environment

Phenotype

Genomic data

Exposome data

Metabolome/proteome data

Phenotype data

whole genome data
SNP array data
life style data
questionnaire data
Laboratory data
clinical data

Epigenetic data

constitutions
lifestyle habits, exposures
physical condition
clinical data

health status
physiological data

Epigenetic data
Integrated Biobank and Databank

Tohoku Medical Megabank (TMM) is an integrated biobank retaining both biobank and genome / omics analytical facilities.
“dbTMM” integrates both health data and genomic/omics data concomitantly toward development of genome medicine.

**Search**

**Constitution** (Genomic data)
Chromosome 8 41519462 (rs515071) = TT

**Health status** (Lab test data) : HbA1c > 6.2

**Lifestyle** (QA) : Alcohol Drinking = Yes

**Disease history** (QA) : Type II Diabetes = Yes
Data Collection and Integration

Data Collection

- Cohort
  - Cohort ID
  - Clinical Data
  - Hospitals (Regional Hospital Information Network)

Data monitoring / cleaning

De-identification

Biobank

- DNA
- Sequencing

QC

Genome & Omics Data

Data Integration and Standardization

- Biobank ID
- Investigation ID
- Clinical ID

Tohoku Medical Megabank

Integrated Database “dbTMM”
Smooth Search of Genomic, Health, and Clinical “Big Data”

Integrated Database “dbTMM” stores genomic, health and clinical “big data” over ~3.7 million SNV sites for 34,000 participants

1.3 trillion data entries

Statistical Characterization of Stratified Population

Stratified Population
3000 participants

How can researchers understand narrowed-down population w/o browsing all the variable data?

Detection of statistically significant difference

Mother Population
150,000 participants

Statistical characterization of stratified population gives us a hint for research.
Detailed Search
Any variables, combinations, conditions with AND/OR

Category Facet
Cohort type, Disease classification, Omics data type

Data Table
Data table for stratified population

Graph for Stratified Population
Sex, Age, BP, Smoking, Alcohol Drinking, Disease History

Statistical Characteristic for Stratified Population
Detection of statistically significant difference
Integrated Database dbTMM Catalogue

Explanations and statistical graphs for all the variables stored in our integrated database “dbTMM”
Immanent Molecular-types and Deep Phenotypes

Questionnaire
Clinical Data
Observation
Laboratory Value
Clinical Score
Omics Data
Genome Data
Exposures
Disease
Subtypes
Deep Phenotype
Immanent Molecular-type
Molecular Network
Genotype

Development of Genomic Medicine and Drug Discovery for Immanent Molecular-type and Deep Phenotypes
Development of Biobank Network
Development of Biobank Network

12 Biobanks

- Kyoto Univ Biobank
- NCBN (6 National Center Biobanks)
- Biobank Japan
- Okayama Univ Biobank
- Tsukuba Univ Biobank
- Tokyo Medical & Dental Univ Biobank

- Hospital biobank
- Population cohort
For promotion of utilization of biobank toward realization of genomic medicine, this project aims at

- Research & development of biobank network by biobank cross-search system
- One stop service to access to biospecimen and data.
Participant

Sex
Disease history

Biospecimen

Biospecimen Type
Disease name
Age

Data

Data Type
Vendor
Platform

Biobank cross-search of biospecimen and data

Access to biospecimen and data

Academia

Industries
ISBER-IRL and BBMRI-ERIC Directory/Sample Locator

ISBER-IRL (International Repository Locator)

Search

Enter your search criteria below. Matching results will be shown here. You may also show all results.

Use text searching to find repositories:

- Text Search
  - Search for...

Found 41 matching repositories

You may also show all results. Click on a repository to view more information about it.

- Repository
  - International Melanoma Biorepository and Research Laboratory
  - Wesley Medical Research Tissue Bank
  - NIDDK Central Repository
  - Biobank of the Respiratory Health Network of the FRQS
  - Tumorbank@UZA
  - Spectrum Health Universal Biorepository (SHUB)
  - NHLBI BioLINCC

BBMRI-ERIC Directory 4.0

Biobank/Biorepository search

Sample Locator is under R&D for biospecimen search

German Biobank Node launched Sample Locator in Oct 2019
# Minimum Common Datasets

<table>
<thead>
<tr>
<th>Category</th>
<th>Property</th>
<th>Minimum information of biobank</th>
<th>Standards</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MIABIS</td>
<td>CCB</td>
</tr>
<tr>
<td>Last update</td>
<td></td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Biobank</td>
<td>Biobank ID</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Participant</td>
<td>Participant ID</td>
<td>○</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>○</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Previous disease code</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Previous disease name</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Biospecimen</td>
<td>Disease code</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Disease name</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Biospecimen type</td>
<td>○</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Age sampled</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Data</td>
<td>Data type</td>
<td>○</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Vendor</td>
<td>○</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Platform</td>
<td>○</td>
<td>●</td>
</tr>
</tbody>
</table>
Search by Disease Name/Code

Disease name

Disease code
Sample/Data Search of Your Interest

Sample Type

Data Type
Search Results

Biospecimen/data of your interest

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2019.12.12</td>
<td>ToMMo</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2018.12.11</td>
<td>関大バイオ...</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2017.12.10</td>
<td>NCBN</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2016.12.9</td>
<td>BBJ</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2015.12.8</td>
<td>ToMMo</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2014.12.7</td>
<td>関大バイオ...</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2013.12.6</td>
<td>NCBN</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2012.12.5</td>
<td>BBJ</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2011.12.4</td>
<td>ToMMo</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2010.12.3</td>
<td>関大バイオ...</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2009.12.2</td>
<td>NCBN</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2008.12.1</td>
<td>BBJ</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2015.12.8</td>
<td>ToMMo</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2014.12.7</td>
<td>関大バイオ...</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2013.12.6</td>
<td>NCBN</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2012.12.5</td>
<td>BBJ</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2011.12.4</td>
<td>ToMMo</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2010.12.3</td>
<td>関大バイオ...</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2009.12.2</td>
<td>NCBN</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
<tr>
<td>2008.12.1</td>
<td>BBJ</td>
<td>123456780</td>
<td>A000</td>
<td>アジアコレラ</td>
</tr>
</tbody>
</table>
Participant Details

Participant

Biospecimen

Phenopacket

Data

- Metabolomics
- Illumina
- Japonica Array v2
- SNP array
- Affymetrix
- Human OmniExpressExome v1.0
Inquiry to Biobank w/ Specimen/Data of Your Interest

Inquiry to Biobank by email

ToMMo
otaia.wase@megabank.tohoku.ac.jp

ToMMo
TMM123456789
TMM123456789
TMM123456789
TMM123456789

BBJ
otaia.wase@megabank.tohoku.ac.jp

BBJ
Blood
Biobank Network launched in Oct 2019

Biobank cross-search of biospecimen and data

Access to biospecimen and data
バイオバンク横断検索システム
2019年秋 初版リリース

https://biobank-search.megabank.tohoku.ac.jp/

研究者のみなさま、
このようなお詫びはありませんか？

バイオバンク横断検索
システムをご利用ください！

① ウェブサイトへアクセスする
https://biobank-search.megabank.tohoku.ac.jp/
こちらから利用IDを取得してください

② 検索条件を選択する
（性別、病名、試料種類など）

横断検索システムの使い方

開発スケジュール

2019年 春
プロトタイプリリース

2019年 秋
初版リリース

2020年 秋
第2版リリース

Press Release on Oct 28, 2019
Current Statistics of Biobank Network in Japan

Participants

- **152,291**
- **153,357**

Age sampled

- 0
- 10
- 20
- 30
- 40
- 50
- 60
- 70
- 80
- 90
- 100

Disease history

- 3811 Diseases
- >10 1517 Disease

Biospecimen

- 654,553
- DNA: 291,119
- Blood: 238,462
- Blood plasma: 59,966
- RNA: 20,591
- Tissue (frozen): 10,168
- Brain homogenate: 2,619
- Buffy coat: 1,671
- PBMC: 1,320
- Tissue (PAXgene): 1,155
- Tissue (RNAlater): 950
- Serum: 377
- Urine: 318
- Bronchial lung fluid: 84
- Pleural fluid: 46
- Peritoneal fluid: 25
- Tissue (FFPE): 1
- Peritoneal fluid (3D): 0
- Blood cell: 0
- Saliva: 0
- Sputum: 0
- Stools: 0
- Others: 25,680

Data

- 198,968
- SNP array: 170,115
- SNP array (imputed): 22,689
- Genomics: 0
- Metabolomics: 0
- Proteomics: 0
- Transcriptomics: 0
- WES: 0
- WGS: 0
- Others: 6,164
Biobanks for improving human health

Collection of life-course data

Biobank

Gene

Environment

Phenotype

Genomic data
Epigenetic data

constitition

lifestyle habit, exposures

health status

physical condition

clinical data

Exposome data
(Life style data)

Metabolome/proteome data
Laboratory data
Physiological data

Phenotype data

as a trusted infrastructure collecting genomics, phenotypic and life-style data (exposome data)
Acknowledgement
People in ToMMo

ToMMo has more than 380 members including GMRC / TCF

Thank for your help and cooperation!
Acknowledgement

AMED Biobank Network project in Japan

Biobank Japan (BBJ)
  Yoshinori Murakami
  Koichi Matsuda
  Takayuki Morisaki
  Seiya Imoto

National Center Biobank Network (NCBN)
  Yuichi Goto
  Kotaro Hattori
  Yoshihiro Miytamoto
  Ryo Matsumura

Tohoku Medical Megabank (TMM) project
  Hiroshi Yokota
  Kichiya Suzuki
  Fuji Nagami
  Satoshi Nagaie

Toyama International University
  Toshihisa Takagi
  Shin Kawano

Okadai Biobank (OKB)
  Mizuki Morita

Keio University
  Hiroshi Nishihara

CIBER
  Hiroki Nakae

Thank for your attention