Introduction of
Tohoku Medical
Megabank Project

Feb 12th, 2020
Masayuki Yamamoto

Tohoku University
Tohoku Medical Megabank Organization
Tohoku Medical Megabank Project
Towards Creative Reconstruction of Tohoku Region

Support of residence in tsunami suffered area

Support community medicine of the damaged area by establishing next generation medical systems

Increased recruitment of medical professionals to the suffered area

Circulating system of medical services

Large scale genome cohort and Integrated biobank

Personalized healthcare

Miyagi Medical & Welfare Information Network (MMWIN)

Tohoku Medical Megabank Project
Endeavor for “Society of Health and Longevity"

**Personalized Medicine**

- Onset of a disease
- Analysis of genome
- Precise diagnosis and selection of treatment

**Personalized Healthcare**

- Healthy people
- Analysis of DNA and lifestyle info
- Return of results
- Prevention by adequate healthcare
Toward Realization of Personalized Healthcare
ToM Mo Have Established Strategically Two Types of Cohort Studies
Residents Cohort and Birth & Three Generation Cohort

- **Community / Residents Cohort** (in Miyagi and Iwate)
  Recruit 80,000 residents from coastal areas in Miyagi and Iwate provinces

- **Birth & Three Generation Cohort** (in Miyagi)
  Recruit 70,000 people including offspring, parental and grandparental generations

- TMM established two types of cohort to strategically achieve our goal
- TMM genome cohorts are in a cutting edge design

We have finished recruit of
84,073 participants for Resident Cohort and
73,500 for Birth & Three Generation Cohort
**Total of 15,573**
ToMMo Community Support Center (CSC)

- Center for ToMMo Cohort Studies
- Center for GMRC activity
Birth and Three Generation Cohort Study

- **Target area:**
  Whole Miyagi Prefecture

- **Target subjects:**
  20,000 pairs of pregnant women and their fetus, siblings, fathers, grandparents and other family members (20,000 family, 70,000 participants)

- **Place for the recruitment:**
  - Obstetric clinics or hospitals and 7 Community Support Centers
  - We recruit pregnant women who are diagnosed pregnancy
  - After their participation, we ask their family to participate
  - Individually obtained informed consent is required
Scenes of Our Cohort Recruits

Community Cohort Recruits
- Join annual health surveys of local governments and recruit participants at the sites
- Recruit participants at the Community Support Centers

Birth and Three Generation Cohort
- Recruit pregnant mothers at OB hospitals by our GMRC*

*GMRC: Genome Medical Research Coordinator
Samples and Data from One Participants

Blood 34 ml (21ml for Storage) + Urine

<table>
<thead>
<tr>
<th>Volume</th>
<th>Sample Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 ml</td>
<td>Regular Blood Test + HbA1c, Blood Glucose, Cholesterol, AST, ALT, BUN, Cr, IgE, <em>Helicobacter pylori</em></td>
</tr>
<tr>
<td>7 ml (EDTA)</td>
<td>Plasma / Buffy Coat (DNA Extraction)</td>
</tr>
<tr>
<td>9 ml (Plain)</td>
<td>Serum</td>
</tr>
<tr>
<td>5 ml (Heparin)</td>
<td>Mononuclear Cells</td>
</tr>
<tr>
<td>Urine</td>
<td></td>
</tr>
</tbody>
</table>

Physiological Examinations
Cardio-Vascular, Respiratory, Ocular, Bone Density, Dental etc.

MRI (10,000 Participants, Test on Cognition and Depression)

**Questionnaire** (Japanese-Cohort Standard)
Age, Sex, Area, Job, Diet, Sleep, Exercise, Alcohol, Smoking History, Prescription, Gynecologic Problems, Mental Health, Psychosocial Factors

**Disaster-Related Questions**
Damage, Life-Style Change
MRI in Tohoku University

Aims
- Early diagnosis of Dementia and Alzheimer diseases
- Evaluate relationship of tsunami stress and brain structure changes
- Detection of biomarkers for disease prevention

Philips Ingenia 3.0T
(Two machines)

Volunteer basis
Relatively rigorous exclusion criteria

Around 12,000 imagings have been done
TMM Biobank

A system that collects, stores, and distributes biological specimens and related information for the advancements of medicine and science

Biobank is beneficial for the society

Large size biobank is good for
- Efficient use of resources
- Good quality control
- Reasonable use of resources

February, 2020
More than 3.8 million sample storage in total

From August 25th, 2015, ToMMo has started distribution of samples and information to research scientists
TMM Biobank Is an Integrated Biobank

Integrated Biobank

• TMM sets up an analytical center that executes standard analyses of samples
• To avoid rapid depletion of samples, TMM distributes analysis information first, and then bio-samples

Blood & Urine

Whole blood, serum, WBCs are stored
→ metabolome and proteome

Genomic DNA

DNA extracted from blood is also stored
→ genome and transcriptome

Questionnaire

Main part is for lifestyle (including food), psychological condition, experiences of the disaster

More than 10 physiological examinations, and cognitive and psychological assessment
Whole Genome Sequencing in TMM

**Finished 4,773 whole genome sequencing**
WGS in single laboratory, single protocol, and single facility with high coverage is first in the world

**Catalogue more than 61 million SNVs and 26 million are new SNVs**
Of the new SNVs, more than 99% are rare variants

### August, 2014
Parts of the SNV information was open from ToMMo and NBDC.

### June, 2016
All SNV information was open from ToMMo.

### June, 2018
SNV information on X Chromosome and mitochondria were released (3,552 people).

### Sep, 2019
4,773 people SNV, INDEL, X Chromosome and mitochondria were released.

**4.7KJPN**

<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence Variation and Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>3458697</td>
<td>C: 70% T: 30%</td>
</tr>
<tr>
<td>8768942</td>
<td>A: 99.9% G: 0.1%</td>
</tr>
</tbody>
</table>

**Residential cohort 4,773 people**

4,773 people WGS analysis

- Y
- X chromosome
- 22 chromosome
- 2 chromosome

1 chromosome
TMM Whole Genome Variation Database

Position and frequency of sequence variation

- Characteristics of Japanese genome structure are getting clear
- TMM whole characteristics of genome database will facilitate clinical sequence studies

Segmented drug development

TMM integrated data will be of important for new drug development for specific group of people

Personalized healthcare and ethnic array

Generate a special array that enable efficient imputation of Japanese genome
Risk Assessment of Common Diseases Is Important for Personalized Healthcare

Target Disease of Genome Medicine

- Rare Diseases
- Cancer
- Undiagnostic Diseases
- PGx

Common Diseases
**DNA Array Highly Adopted for Japanese Population**

Japonica Array® is designed for various cohort studies in Japan, and will contribute to personalized healthcare and medicine

- Based on Japanese WGS data (3.5KJPNv2)
- Japonica Array is designed to minimize the number of probes but to maximize capacity of genotype imputation for Japanese
- Inexpensive: providing the low cost WGS information will realize mega-size survey of genes responsible for common diseases

Japonica Array® was started marketing ver.1 in 2014, ver2. in 2017, and NEO in 2019.

---

**WGS by NGS**

($1000-2000 / person)

- TMM WG Ref Panel
  - 1000 (2015)
  - 3500 (2018)
  - 4700 (2019)

**Japonica Array**

(>$100 / person)

- Does not exist in SNP array
- Exists in array
- Linkage Disequilibrium Genotype Imputation
- Imputation
Japanese Multi omics reference panel: jMorp

Metabolomics and proteomics reference database from 15,000 cohort participants

https://jmorp.megabank.tohoku.ac.jp/
Supercomputer is essential for biobank activities
Supercomputer is divided into Unit A through Unit C
- Unit A is for open database
- Unit B is for data sharing / visiting
- Unit C is for informatics and data processing

- Users can access AMED super commuter and dbTMM from remote security rooms set up in distantly located area in Japan
- This system contributes to the wide ranging data visiting of TMM
Integrated Biobank and Database

Tohoku Medical Megabank (TMM) is an integrated biobank retaining both biobank and genome/omics analytical facilities.

Tohoku Medical Megabank Integrated Database “dbTMM”

Scientists in Academia and Industry
Toward Overcoming Problems Surrounding Biobank

- Budgets for secured managements
- Return of data / intellectual properties
- Technical issues on long term sample storage
- Ethical issues (Incidental findings, etc.)

International collaborations are essential to overcome these issues

Merits of collaborations
- Enable elaborate and large scale meta-analyses
- Standardization of sample collection, preparation and storage
- International standardization of questionnaire
- Approach to ethnic difference / similarity of genomes
- Large scale collaboration
We have established Tohoku Medical Megabank to realize personalized healthcare (PHC) and personalized medicine.

In the Tohoku Medical Megabank Project, we have designed and are operating of two types of cohorts; community cohort and birth-and-three-generation cohort, which in collaboration will realize cutting edge accomplishments in Longitudinal Population Studies.

We have established an integrated biobank and are conducting genomics analyses for genome medicine.

http://www.megabank.tohoku.ac.jp/index.php
Directors / Professors
Masayuki Yamamoto
--
Shigeo Kure
Kengo Kinoshita
Nobuo Fuse
--
Inaho Danjo
Hiroaki Hasizume
Atsushi Hozawa
Hiroshi Kawame
Shinichi Kuriyama
Fumiki Katsuoka
Seizo Koshiba
Eiichi Kodama
Naoko Minegishi
Fuji Nagami
Tomohiro Nakamura
Soichi Ogishima
Kinuko Ohneda
Mika Sakurai
Ritsuko Shimizu
Junichi Sugawara
Kichiya Suzuki
Takako Takai
Yasuyuki Taki
Gen Tamiya
Hiroaki Tomita
Akito Tsuboi
Jun Yasuda
--
Nobuo Yaegashi
Sadayoshi Ito
Hiroshi Tanaka
Tadao Kobayashi
Yoshiyuki Sato

People in ToMMo

ToMMo has more than 380 members including GMRC / TCF
GMRC: genome medical research coordinators
TCF: ToMMo clinical fellows

Thank you for your help and cooperation!