

Tohoku University Tohoku Medical Megabank Organization

Overview





Tohoku University Tohoku Medical Megabank Organization Past and Future

Masayuki Yamamoto
Executive Director, Tohoku University Tohoku Medical Megabank Organization

Tohoku University Tohoku Medical Megabank Organization (ToMMo) has been working on the following projects in response to the unprecedented damage caused by the Great East Japan Earthquake of March 11, 2011. ToMMo was established on February 1, 2012, with the aim of contributing to the recovery of medical care in the affected areas and to establish a system for the development of futuristic medicine to be delivered promptly to those affected by the disaster. Since then, we have developed many related businesses, and it has been more than 10 years since our establishment.

In this project, we conducted the Community-Based Cohort Study focusing on residents and the Birth and Three-Generation Cohort Study focusing on pregnant women and their families. The Community-Based Cohort Study was conducted jointly by Iwate Medical

University Iwate Tohoku Medical Megabank Organization, and we were ultimately able to obtain the cooperation of a total of 157,000 people for both cohorts.

We would like to express our sincere gratitude to the people of Miyagi and Iwate prefectures who participated in this project.

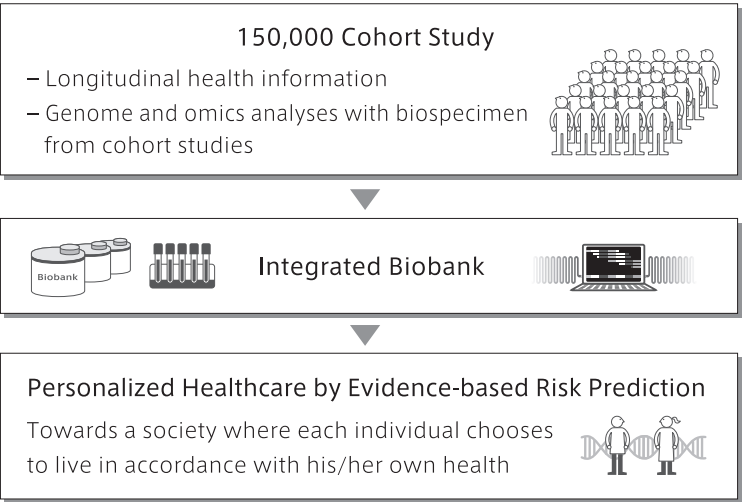
ToMMo has established a large-scale biobank to safely store the biological specimens deposited in these cohort studies, and has proceeded with whole genome analysis and omics analysis of the samples to make the database publicly available. We are also working to establish an efficient follow-up survey of cohort participants, and are reporting information on the health status of the participants and the medium to long-term effects of the earthquake on them to the participants and to local governments. In addition, we have established

a system that allows many researchers to make use of the biobank's specimens and information, and we are also participating in the Platform for Genomic Medicine by the Japan Agency for Medical Research and Development (AMED), making an active contribution as one of the three major biobanks in Japan. We are also promoting collaboration with many industries, including the Consortium for Integrated Analysis of Genome, Medical and Health Information that organized with pharmaceutical companies.

ToMMo is striving to build the future of medicine centered on personalized healthcare by synthesizing the results of cohort studies, genomic and omics analysis, and biobank operations. We are working as one to build a future in which each individual can choose the medical treatment and prevention that is best suited to him or her based on scientific evidence.

Concept

Tohoku University Tohoku Medical Megabank Organization (ToMMo) was founded to establish an advanced medical system to foster the reconstruction from the Great East Japan Earthquake. ToMMo has been developing a biobank that combines both medical and genome information during the process of rebuilding the community medical system and supporting health and welfare in the Tohoku region. The information from the brand-new biobank will allow a new medical system to be created. Based on the analytical findings, ToMMo aims to attract more medical professionals from all across the country to the area, promote industry-academic partnerships, create employment in related fields, and lastly reconstruct the medical system in the Tohoku region.



Activities of ToMMo: Establishment of Integrated Biobank for Next Generation Medicine

The Tohoku Medical Megabank Project is pursuing the following studies:

- Long-term health survey for community residents
- Establishment and operation of large-scale biobank that integrates both biological specimens and health-related information
- Analyses of biological specimens and information amassed by the biobank using advanced technology

Through these activities, the project is enhancing the progression of future-oriented health care and cutting-edge research, while contributing to reconstruction efforts.

Long-term Health Survey

The Tohoku Medical Megabank Project has accumulated the longitudinal health survey information and genetic information of over 150,000 people. This was only achievable due to the generous cooperation and consent of community residents, while facing local health issues, after the Great East Japan Earthquake.

This project conducts two cohort study programs: The Community-Based Cohort Study (TMM CommCohort Study) and the Birth and Three-Generation Cohort Study (TMM BirThree Cohort Study).

In a cohort study, information about the lifestyle habits of a large group of people are collected and studied to understand how their habits and their surrounding environment may correlate to disease. ToMMo's cohort studies also incorporate longitudinal MRI imaging and the collection of daily lifestyle information through smartphone applications to build a large data set.

Building and Operating an Integrated Biobank

Our biobank is collecting and storing de-identified biological specimens including DNA, plasma and serum from blood and corresponding health-related and clinical information obtained from the participants of cohort studies. Some biospecimens are analyzed and the information is incorporated into an integrated biobank. Researchers from various institutions are able to use these resources for health-related studies upon the approval of their research applications.

Detailed Analyses of Biological Specimens and Information

ToMMo is conducting whole genomic analyses of some of the participants

from the cohort studies. In November 2013, ToMMo completed whole genomic analysis of 1,000 individuals in Japan and is continuing to expand the analysis to tens of thousands of individuals. By the first half of 2024, we aim to have completed whole genome analyses for approximately 100,000 individuals. The largest whole genome analysis of a general population in Japan has contributed widely to many related fields, including industry and academia. That is, a whole genome reference panel cataloging the variants detected by this analysis has led to significant support for patient genome analysis. The reference panel will greatly contribute to the development of innovative drug discovery and is indispensable for the realization of personalized healthcare. ToMMo also performs a variety of omics analyses, including large-scale metabolomic analysis of plasma, which continues to contribute to the enrichment of the integrated biobank. Taken together, we are attempting to elucidate the mechanisms of disease caused by complex interactions between genetic and environmental factors; and to realize evidence-based, personalized disease risk prediction.



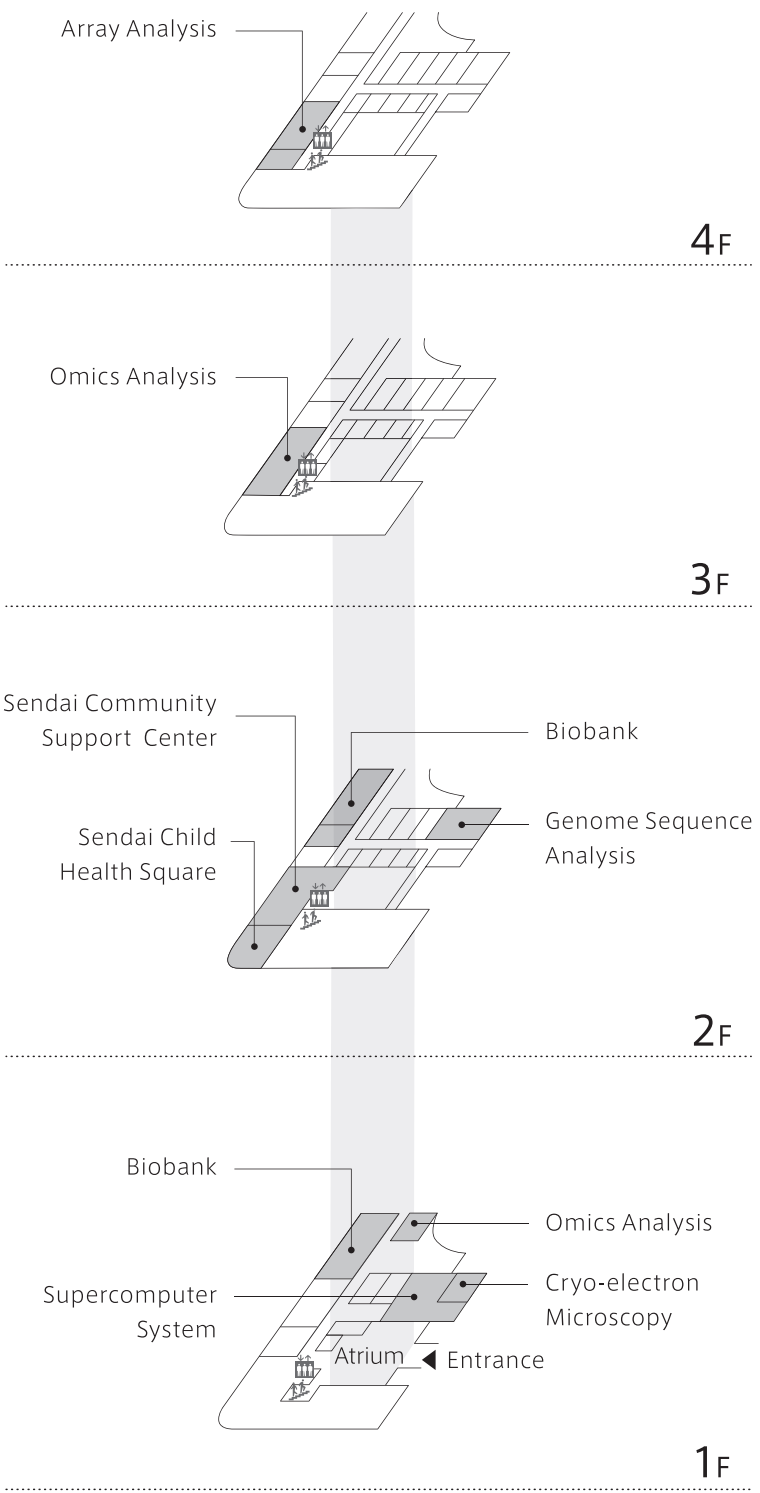
Main Research Facilities & Equipment

The Tohoku Medical Megabank Building was completed in 2013 at the Seiryō Campus of Tohoku University. The building is a seven-story structure with a building area of 3,425.62 m² and total floor area of 18,017.67 m². It has a prestressed concrete structure and base-isolated structure, and is connected to the adjacent five-story School of Medicine Building No. 6 on each floor and partially connected to the existing School of Medicine Building No. 5 on the second floor.

The building is equipped with biobank facilities (biological sample storage), MRI equipment, NMR equipment, sequencing analysis equipment, supercomputer, etc. A café is located in the atrium on the first floor, and the Sendai Community Support Center and Sendai Child Health Square are located on the second floor, where ToMMo's community residents The third floor and above are laboratory spaces including omics analysis facilities and array analysis facilities.



The floor tiles in the atrium are designed to resemble the "double helical structure" of DNA.



Supercomputer System

The supercomputer system is the data infrastructure for storing various data generated from a 150,000-people genome cohort, as well as the analysis infrastructure supporting large-scale genomic and omics analysis. As a data infrastructure, the system is equipped with a massively parallel hard disk drive that enables high-speed data analysis and storage consisting of tape libraries/object storage/optical disk libraries that can store data for a long period of time, making it the largest system in the life science field in Japan. As an analysis platform, the system is equipped with a large number of CPUs to perform large-scale genomic and omics analysis, which is essential for the construction of next-generation medicine, and enables high-speed data analysis. The system is designed to efficiently perform the analysis necessary to realize personalized healthcare while ensuring a high level of security for the diverse data from the complex biobank.



Biobank

The TMM biobank is one of the largest population-based biobanks in Japan that stores millions of biological specimens in the state-of-the-art storage systems. Automated sample storage system can hold 4.5 million tubes which contain blood components, urine, breast milk and DNA at either -80°C or 4°C. The system handles the specimens autonomously, minimizing the frequency of human errors and sample damage. The liquid nitrogen storage system can hold 1.3 million tubes in the vapor phase of liquid nitrogen below -180°C. The tubes contain peripheral and cord blood mononuclear cells, along with their derivatives. The specimen-associated information is managed by the Laboratory Information Management System (LIMS).



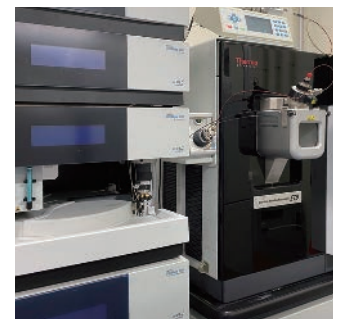
Genome Sequence Analysis

ToMMo sequencing facility has different types of next-generation sequencers. It has an annual output capacity of more than tens of thousands of high-grade whole human genome sequences. To perform such high throughput sequencing without any sample mix-ups, highly automated experimental pipelines for sample preparations. Using not only short-read sequencers but also long-read sequencers, various types of diversities with the human genome, such as single nucleotide variants and structural variants, are analyzed. In order to sequence a whole human genome with the best quality possible, 20 to 100 times the amount of data (up to 300 billion bases) must be collected for analysis. These massive amounts of data are protected by a highly secure information network and analyzed with our supercomputer system.



Omics Analysis

The cutting-edge facility for omics analysis allows the analysis and quantification, to a high degree of precision, of the various metabolites and proteins that are contained within the blood and urine samples stored in the biobank facility. Three high-field and high-sensitive nuclear magnetic resonance (NMR) spectroscopy system (800MHz and 600MHz) and various types of mass spectrometry system (LC-MS, GC-MS, etc.) have been installed to enable a wide variety of analysis to take place. Each system has an auto-sampler, allowing analyze several thousands of samples to be analyzed per year. The obtained data are then stored in the supercomputer system. Our study provides the standard omics information (types and quantity of metabolites and proteins)



Array Analysis

SNP genotyping using microarray techniques are performed in the facility designated for array analyses. Several types of microarray systems are used according to the purpose of each project. Each system is equipped with a Laboratory Information Management System (LIMS) to record samples and reagents information, conditions of dispensing robots, and information regarding to process control. The microarray data are then sent to our supercomputer system where quality checks and genotype imputation are performed.



Common Terminal Room

The common terminal room is equipped with several terminals that can be connected to the supercomputer system in a secure room with biometric authentication facilities. These are also dozens of remote security areas with similar function and access to the ToMMo supercomputer in Sendai from remote locations throughout Japan and the integrated database dbTMM can be used and analysis and calculations can be performed using the supercomputer.



Nihonbashi Branch

The Nihonbashi Branch was established in July 2020 as the Tokyo office of ToMMo. Three thin client terminals are installed for common terminal room and are connected to the ToMMo supercomputer via a VPN line. Researchers from distant locations can effectively use the various analysis information stored in the ToMMo supercomputer and its computing power without having to travel to Sendai.



Community Support Center

For cohort studies, we have established seven Community Support Centers in Miyagi Prefecture (Kesennuma, Osaki, Ishinomaki, Tagajo, Sendai, Iwanuma, and Shiroishi). The centers are equipped to handle various types of surveys after confirming consent from each individual by a Genome Medical Research Coordinator (GMRC). In Sendai, a dedicated facility for conducting health surveys with particularly for children, called Sendai Child Health Square, was opened in June 2017. Also MRI equipment has also been installed in Sendai, and MRI imaging is provided to cohort participants who have been taking health surveys in the centers and wish to do so.



Health Survey Detail

Informed consent, blood test, body composition, grip strength test, bone density measurement, hearing test, oral health check-up, carotid ultrasound imaging, respiratory function test, ophthalmic examination, questionnaire by touch screen, electro-cardiogram, MRI, cognitive assessment.



Location



Research Facilities of INGEM

INGEM (The Institute for the Advanced Research Center for Innovations in Next-Generation Medicine) is a university-wide organization in which ToMMo participates and was established in 2017. In the Tohoku Medical Megabank building INGEM's distinctive research facilities are also installed.



Clinical Biobank

As one of the facilities of INGEM, the clinical biobank operates along with the Tohoku University Hospital Personalized Medicine Center to collect precious biological specimens through medical care from the Tohoku University Hospital. Not only does it store liquid samples such as blood components, but also various types of other samples, including tissue specimens from surgery. All the biological specimens are anonymized and preserved at appropriate temperatures for future research.



Cryo-electron Microscopy

Cryo-electron microscopy (CRYO ARMTM 300 II, JEOL Ltd.) installed in the Tohoku Medical Megabank building is a device that enables structural analysis of various biological macromolecules with the world's highest resolution. In addition, a super-computer enables processing of large-scale measurement data. We will actively accept the use of this instrument not only within the university, but also from external companies and research institutions. It is expected to elucidate the mechanism of disease onset and realize drug discovery based on structural information.



About INGEM

In 2017, Tohoku University was selected as the first three Designated National Universities in Japan. The conferment of the Designated National University title is a recognition of the university's abilities to lead and shape global education and research. As a world-class research center and a leader of creativity and innovation, we have established a hub for next-generation medicine and founded INGEM as the organization that will play a central role within this hub. In cooperation with ten departments within Tohoku University, we have created a structure that brings together our collective capabilities.

Key Publication

- Fuse N *et al.* Establishment of Integrated Biobank for Precision Medicine and Personalized Healthcare: The Tohoku Medical Megabank Project. *JMA J.* 2, 2, 113-122, 2019.
- Kawame H *et al.* The return of individual genomic results to research participants: design and pilot study of Tohoku Medical Megabank Project. *J Hum Genet.* 67, 9-17, 2022.
- Kuriyama S *et al.* Cohort Profile: Tohoku Medical Megabank Project Birth and Three-Generation Cohort Study (TMM BirThree Cohort Study): Rationale, Progress and Perspective. *Int. J. Epidemiol.* 49, 1, 18-19m, 2019.
- Hozawa A *et al.* Study profile of The Tohoku Medical Megabank Community-Based Cohort Study. *J. Epidemiol.* 31, 1, 65-76, 2021.
- Minegishi N *et al.* Biobank Establishment and Sample Management in the Tohoku Medical Megabank Project. *Tohoku J Experiment Med.*, 248, 1, 45-55, 2019.
- Ogishima S *et al.* dbTMM: an integrated database of large-scale cohort, genome and clinical data for the Tohoku Medical Megabank Project. *Hum Genome Var.*, 8, 44, 2021
- Yasuda J *et al.* Genome analyses for the Tohoku Medical Megabank Project towards establishment of personalized healthcare. *J. Biochem.* 165, 2, 139-158, 2018.
- Takayama J *et al.* Construction and integration of three de novo Japanese human genome assemblies toward a population-specific reference. *Nat Comm.* 12, 226, 2021.
- Koshiba S *et al.* Omics research project on prospective cohort studies from the Tohoku Medical Megabank Project. *Genes Cells.* 23, 6, 406-417, 2018.
- Saito S *et al.* Oral microbiome analysis in prospective genome cohort studies of the Tohoku Medical Megabank Project. *Front. Cell. Infect. Microbiol.* 10, 604596, 2020.



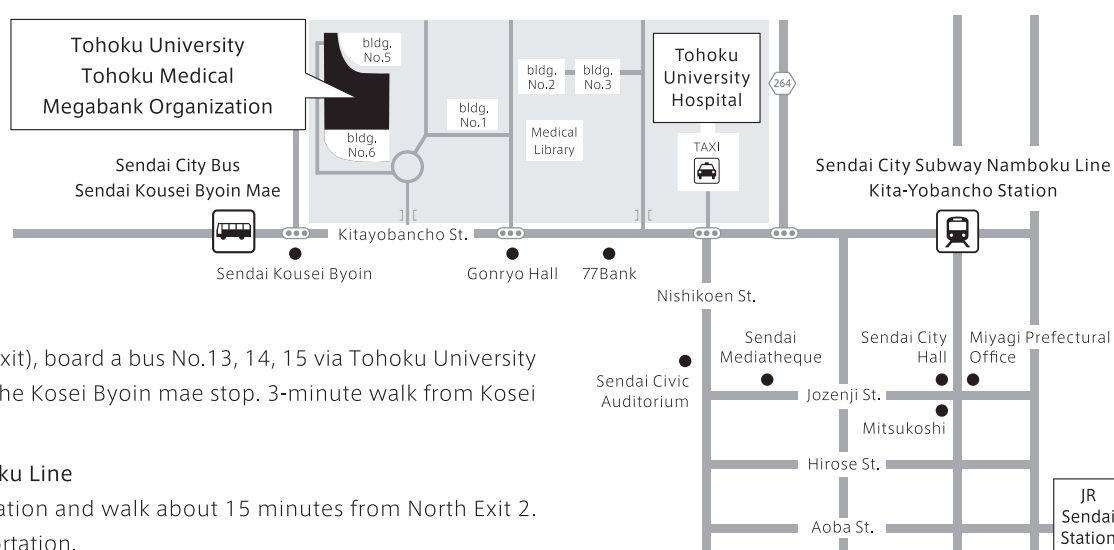
ToMMo analyzed the genome and omics data of the cohort study participants in the Tohoku Medical Megabank Project and published the statistical data to the jMorp(Japanese Multi Omics Reference Panel) web database.

The jMorp database was created in July 2015 to provide the analytical findings of multi omics data. Since then, it has been updated annually to include datasets for the genome, metabolome, methylome, transcriptome, PGx, and metagenome.

The following data are present in jMorp. (As of December 2023)

Genome Sequence	· Japanese reference sequence constructed from the de-novo assembly of three Japanese males.
Genome Variation	· SNV/INDEL allele and genotype frequency data, and HLA allele frequency data derived from the short-read whole genome sequencing of over 54,000 Japanese individuals. · Allele frequency data of copy number variations (CNVs) derived from short-read whole genome sequencing of more than 48,000 individuals · Allele and genotype frequency data of structural variations(SVs) derived from the short-read whole genome sequencing of more than 8,300 Japanese individuals, and the long-read whole genome sequencing of 222 Japanese individuals.
Other Genome-related Data	· Average depth information from short-read whole genome sequencing. · Linkage disequilibrium map derived from 300 haploids. · Lists of markers tiled on Japonica Arrays (SNP Arrays) developed by ToMMo.
Methylome	· Data on DNA methylation, gene expression, and allele frequency for three different types of blood cells in approximately 100 Japanese individuals.
Transcriptome	· RNA-seq analysis of whole blood samples from about 550 individuals in their 30s and 60s · Transcriptome analysis of three Japanese male individuals for lymphoblastoid cell lines (LCLs) using long-read sequencing technology (Iso-Seq).
Metabolome	· Metabolome analysis results obtained from around 61,000 Japanese plasma samples.
Phenome	· Drug sensitivity-related enzymes' genome variants and enzyme activity. · Microbiome analysis of plaque, saliva and fecal samples.
Other	· A repository for the TMM project's GWAS analysis results

Access



Sendai City Bus

From Sendai Station (West Exit), board a bus No.13, 14, 15 via Tohoku University Hospital and disembark at the Kosei Byoin mae stop. 3-minute walk from Kosei Byoin mae bus stop.

Sendai City Subway Namboku Line

Get off at Kita-Yobancho Station and walk about 15 minutes from North Exit 2.

* Please use public transportation.