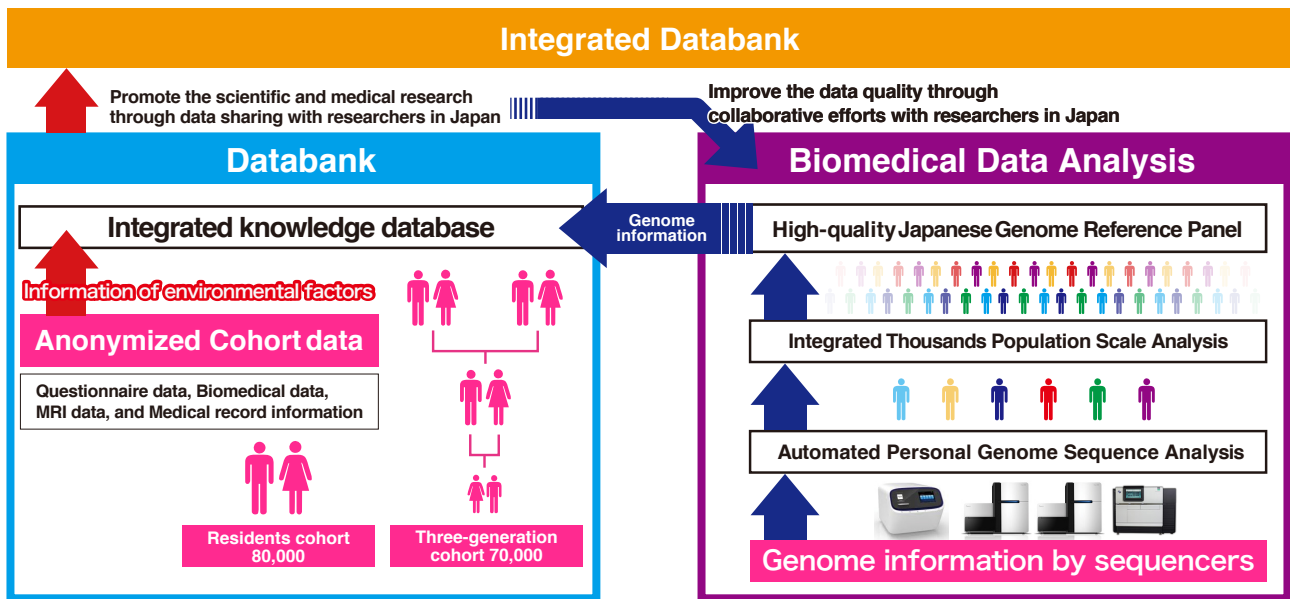


Missions of ToMMo Supercomputer System

1. Construct infrastructure of high quality genome information for Japanese.
2. Provide platforms and databases for large scale analyses of genotype-phenotype relationship and disease risk for Japanese.
3. Construct databank providing secure computational environment to accelerate “personalized healthcare” and “personalized medicine”.



As a prospective cohort project, Tohoku Medical Megabank Organization (ToMMo) plans to recruit 150,000 participants in the Tohoku area, primarily from Miyagi and Iwate prefectures, and gather blood samples, questionnaires, biochemical, and medical information. By the end of March/2014, around 70,000 participants are already recruited. This information is returned to participants health checks along with advice for improving their overall health condition. After anonymization, this information is stored into a secure unit of our supercomputer system as a part of our “Integrative/Knowledge” database.

Our supercomputer is equipped with 16,000 cores (400 TeraFLOPS) on 800 compute nodes, 12.3 Petabytes of high speed parallel file storage and 3 Petabytes of tape archive system. All of them are connected with 56 Gbps fiber channel for high speed data transfer. The system consists of four units, each of which is used for a different purpose and is separated in the network connection to keep the security for each unit.

ToMMo especially cares about the quality control of the information to be a useful data resource to achieve our ultimate goal of “personalized medicine” and “personalized healthcare” in Japan. To understand the risk of disease for each person, environmental factors and individual genetic background are very important. To this aim, last year, blood samples

from 1,000 healthy people were sequenced on the high performance DNA sequencers and the generated DNA sequences (in total 100 trillion bases) are processed on the units of the supercomputer to create the whole-genome Japanese reference panel (1KJPN). At the end of August 2014, the draft version of 1KJPN was constructed. The frequency of common single nucleotide polymorphism (SNP) in the reference panel was released to public domain. ([ijgvd ; http://ijgvd.megabank.tohoku.ac.jp](http://ijgvd.megabank.tohoku.ac.jp))

ToMMo is now shifting towards the sequencing of thousands whole-genome on this supercomputer to create the larger scale Japanese whole-genome reference panel. In parallel, to be a more qualified panel, the discovered SNP and other variants are verified and the reference panel itself is updated with the corroborative research with top researchers in Japan and in the world.

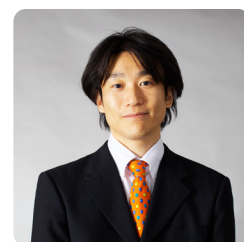
The qualified reference panel is joined to the advanced “Integration/Knowledge” database, which allows us to reveal the fine relationship between genetic factors and environmental factors (e.g. the risk of disease varying depending on individual) for the future of “personalized medicine” and “personalized healthcare”

To satisfy the role of an integrative biobank, our supercomputer has two key responsibilities; the biomedical large-scale “data analysis” (especially whole-genome sequence data) and the

“databank” stored in the Integrative/Knowledge database. The stored data in our supercomputer will be distributed to research institutes and companies after the rigorous examination, e.g. take into account the contributions possible for citizens in disaster area and/or the research community, by the authorized committee in ToMMo.

Through these activities, our supercomputer will serve a prominent role for the future of Japanese healthcare in especially the Tohoku area.

Please visit our supercomputer room, we would like to hear your opinion regarding our integrative databank, our supercomputer system, and our project.



Prof. Masao Nagasaki

Tohoku Medical Megabank Organization
The group leader of *in silico* Analysis