



REVIEW ARTICLE OPEN



National Center Biobank Network

Yosuke Omae^{1,2}✉, Yu-ichi Goto^{1,3} and Katsushi Tokunaga^{1,2}

© The Author(s) 2022

There are six national centers (6NCs) for advanced and specialized medicine in Japan that conduct basic and clinical research on major diseases that have a substantial impact on national health. Disease-specific bioresources and information collected by each NC are stored in a separate biobank. The National Center Biobank Network (NCBN) was established in 2011 and coordinates the biobanks and researchers of the 6NCs via an open-access database (Catalogue Database: http://www2.ncbiobank.org/Index_en) as an efficient means of providing registered biological resources and data for use in research communities. The NCBN resources are characterized by their high-quality and rich medical information and are available for life science research and for the development of novel testing methodologies (biomarkers), new treatments, and drugs for future health care in the scope of personalized medicine through a deeper understanding of disease pathogenesis. Here, we explain the activities of the NCBN and the characteristics of the NCBN Catalogue Database.

Human Genome Variation (2022) 9:1–6; <https://doi.org/10.1038/s41439-022-00217-6>

INTRODUCTION

The pathogenesis and pathophysiology of many common chronic diseases (e.g., cancer, cardiovascular disease, dementia) are extremely complex and require a multifaceted and integrated research approach to elucidate and overcome them. For rare and undiagnosed diseases, the development of effective treatments requires a similar research approach that extends from fact-finding and basic research to clinical practice. Human-derived samples (bioresources) have traditionally been collected and stored for research purposes, and their effective utilization has promoted cutting-edge disease research. Following the development of omics analysis technologies in recent years, including genomics, the importance of bioresource banking has been recognized, and many countries in Europe and the United States are competing to establish biobanks^{1–3}.

ESTABLISHMENT AND ACTIVITIES OF THE NATIONAL CENTER BIOBANK NETWORK (NCBN)

The six national centers (6NCs) for advanced and specialized medicine in Japan are equipped with hospitals and research institutes and conduct basic and clinical research in the following major fields: cancer (National Cancer Center, NCC), cardiovascular diseases (National Cerebral and Cardiovascular Center, NCVC), psychiatric, neurological, and muscular diseases (National Center of Neurology and Psychiatry, NCNP), chronic and infectious diseases (National Center for Global Health and Medicine, NCGM), pediatric diseases (National Center for Child Health and Development, NCCHD), and geriatric diseases (National Center for Geriatrics and Gerontology, NCGG). Each NC has established a disease-based biobank where disease-specific bioresources and information are collected (Table 1). Since the fiscal year 2011, and with the support of the Ministry

of Health, Labor and Welfare, the NCBN has been constructing a networked and federated organization of the 6NC biobanks to utilize patient-derived bioresources for medical and research purposes (Fig. 1).

The Secretariat of the NCBN was established within the NCGM for the purpose of coordinating its implementation. Its activities include (1) standardization of ethical review (comprehensive informed consent acquisition), clinical information (common medical interview questionnaires and disease name registration), and sample handling (standardization and publication of standard operating procedures, SOPs) among the 6NC biobanks; (2) release of the catalogue database; and (3) public relations activities. The first 5-year phase was completed in FY2016, and the second phase began in FY2017, with continuing activities related to clinical information, bioresources, and ethical, legal, and social issues to build a new common platform, as well as activities to further develop the features of the NCBN and promote collaboration with other biobank organizations.

In FY2017, the NCBN established joint meetings with the Research and Development Committee of the Japan Pharmaceutical Manufacturers Association (JPMA), and in FY2020, the GAFREE project was started with the support of the Japan Agency for Medical Research and Development (AMED) to develop a disease-specific information database integrating multiomics data and clinical information. Since FY2018, the NCBN has participated in the “biobank cross-search system” through the Biobank-Construction and Utilization biobank for genomic medicine Realization (B-Cure)⁴. This collaborative research and development project is led by AMED and promotes the utilization of bioresources among the three major biobanks in Japan: the NCBN, BioBank Japan (BBJ), and Tohoku Medical Megabank Project (ToMMo)^{5,6}.

¹Central Biobank, National Center Biobank Network (NCBN), Tokyo, Japan. ²Genome Medical Science Project (Toyama), National Center for Global Health and Medicine (NCGM), Tokyo, Japan. ³Medical Genome Center, National Center of Neurology and Psychiatry (NCNP), Tokyo, Japan. ✉email: yomae@ri.ncgm.go.jp

Received: 31 May 2022 Revised: 2 September 2022 Accepted: 3 October 2022
Published online: 4 November 2022

Table 1. Overview of 6NC biobanks in NCBN.

NC	Main target diseases	Main types of biological samples	Year of NC establishment	Number of enrolled patients in biobank ¹	Relevant NC URLs
NCC	Cancer	Plasma, DNA, RNA, tumor/non-tumor tissue	1962	54,431	https://www.ncc.go.jp/en/index.html
NCVC	Cardiovascular diseases	Plasma, serum, DNA, pathological tissue, body fluid	1977	22,541	https://www.ncvc.go.jp/english/
NCNP	Psychiatric, Neurological diseases	Plasma, serum, DNA, cerebrospinal fluid, brain tissue, muscle tissue	1986	19,764	https://www.ncnp.go.jp/en/
NCGM	Chronic and Infectious diseases	Plasma, serum, DNA, pathological tissue	1993	20,126	https://www.ncgm.go.jp/en/index.html
NCCHD	Pediatric diseases	Serum, DNA, umbilical cord, placenta, chorionic plate	2002	2,468	http://www.ncchd.go.jp/en/index.html
NCGG	Geriatric diseases	Plasma, serum, DNA, cerebrospinal fluid, urine, feces	2004	12,161	https://www.ncgg.go.jp/english/index.html

¹As of 31 March 2022.

CHARACTERISTICS OF NCBN BIORESOURCES

High-quality and rich medical information

The primary characteristic of the NCBN is that it includes a variety of patient-derived bioresources, including detailed medical information, which has been collected by the physicians of each specialist NC. In addition to the collection of basic medical information through linkage to electronic medical records, disease-specific clinical information is also available in each NC biobank. For example, the disease severity of each psychiatric disease patient is assessed by the clinical psychologist at the NCNP. Brain imaging test records, such as MRI and PET, are collected at the NCNP and NCGG, and cardiovascular imaging test records, such as electrocardiograms and echocardiograms, are collected at the NCVC. These data are available as accompanying medical information in the NCBN. These patient-derived bioresources and information provide high-quality research resources that can serve as the basis for basic research as well as advanced medical research in drug discovery, personalized medicine, and the development of regenerative medicine^{7–12}. This characteristic is based on the physical proximity of the medical and research sites at each NC, which enables the selection of optimal sample collection methods and the gathering of necessary clinical information from the perspectives of both physicians and researchers with sound knowledge of both the disease and the research. An important difference between NC researchers and researchers at universities and other institutions who are studying a particular disease is that NCs are continuously conducting research on the diseases in question, and bioresources are accumulated over a long span of 10 to 20 years. In fact, the NCNP has collected more than 16,000 frozen muscle tissue samples since 1978. This continuity is an indispensable advantage of the NCBN when collecting and using bioresources for rare diseases, and it is necessary for the NCBN to maintain a management policy that does not undermine this advantage.

Types of bioresources

An overview of the bioresources that can be provided by the NCBN is shown in Table 2. In general, the target of genomic analysis is peripheral blood DNA. However, it is often necessary to confirm the expression and protein levels to assign significance to the identified genomic mutations. In addition to omics analysis, including genomic analysis, biochemical and pathological analyses are also required in life science research. The NCBN intends

to create a system that can appropriately process, store, and utilize a wide variety of bioresources for such purposes. This wide range of uses is a characteristic feature of NCBN bioresources.

Genomic information

In addition to the bioresource and clinical information characteristics described above, whole genome sequencing analysis for 9,850 NCBN samples has been conducted since FY2020 with support from AMED. These can be utilized as control subjects in research on cancer and rare diseases in the Japanese population, and their variant frequency information will be registered to the NBDC Human Database and Medical Genomics Japan Variant Database (MGeND) after publication^{13,14}. It is now possible to search the catalogue database for bioresources with genomic information.

OPEN-ACCESS BIORESOURCE SEARCH OF THE NCBN CATALOGUE DATABASE

Overview of information registered in the catalogue database

An overview of the catalogue data is shown in Fig. 2A. In addition to basic patient information such as age and sex, other common interview items include information on medical history, family history, allergy history, history of drinking and smoking, and the genome (Table 3). The disease information for primary diseases and comorbidities is registered as the ICD10 code or corresponding MEDIS number. Bioresource information includes the type of specimen (e.g., plasma/serum, DNA, solid tissue, spinal fluid), date of collection, the amount obtained, and method of storage. For pathological specimens, the database also includes SOPs of the storage methods for various types of specimens (<https://www.ncbiobank.org/sample/index.php>). On 31 March 2022, the total number of registered bioresources exceeded 400,000, and the total number of registered patients was 120,081, with a wide distribution of diseases (Table 2, Fig. 2B). These statistics are updated regularly, and the latest figures can be found in the NCBN Catalogue Database (http://www2.ncbiobank.org/Index_en).

How to use the open-access NCBN Catalogue Database (Fig. 3)

When one accesses the NCBN Catalogue Database via the URL (http://www2.ncbiobank.org/Search/Search_en), a search condition window appears. One can search for a specific disease by name and further narrow the search by including basic patient information (sex, age), type of bioresource, current and previous

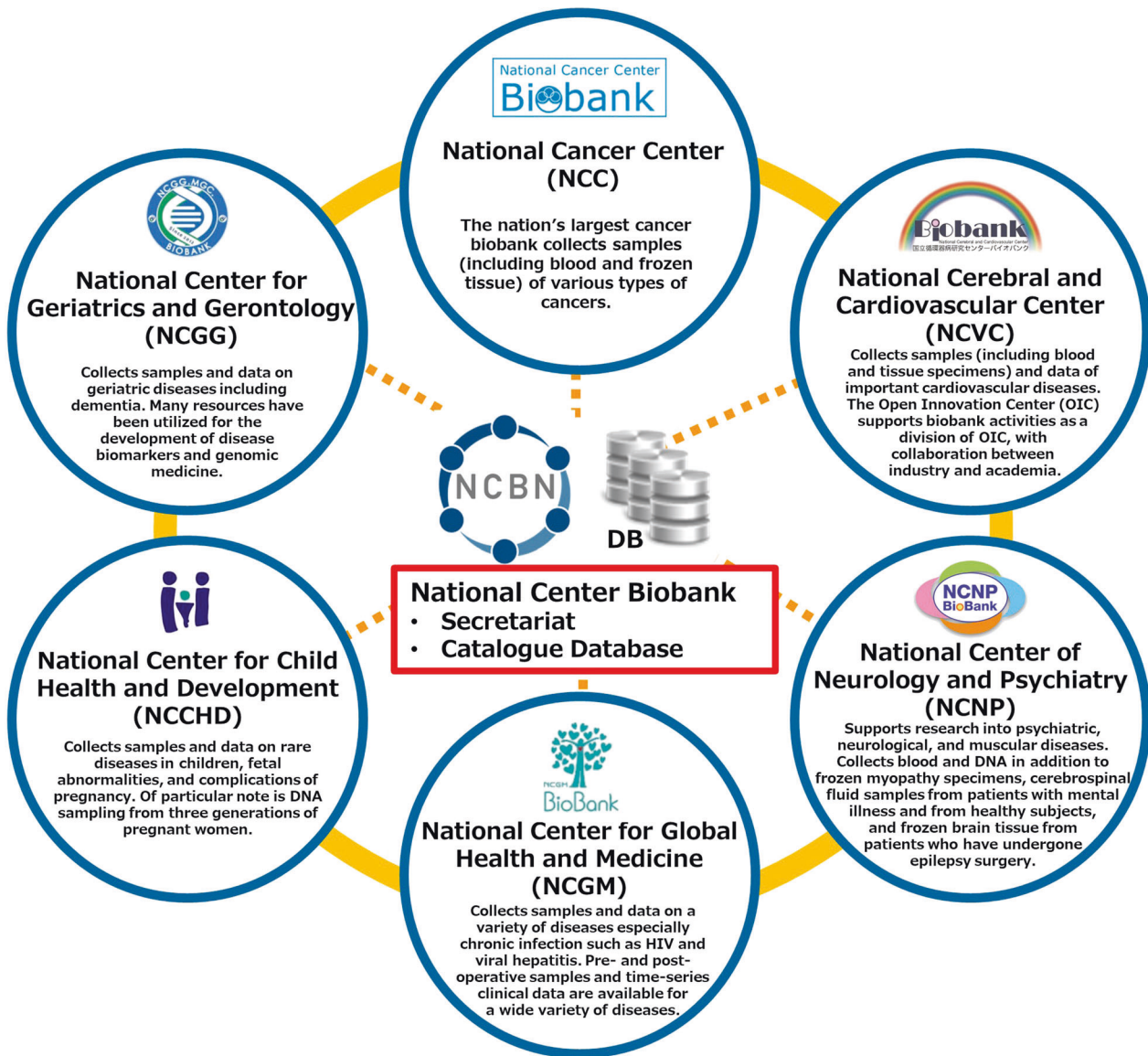


Fig. 1 Overview of the NCBN. The NCBN is a collaborative project by six national centers in Japan. National Center Biobank coordinates secretariat and catalogue database to promote medical research and development.

Table 2. Overview of bioresources in the NCBN catalogue database.

NC	Total number	Bioresources						
		Plasma	Serum	DNA ¹	RNA	Cerebrospinal fluid	Solid tissue	Others ²
6NC	414,046	99,832	77,699	128,491	44,169	4898	34,161	24,796
NCC	154,330	44,169	0	44,169	44,169	0	21,823	0
NCVC	119,386	25,347	26,295	43,216	0	0	531	23,997
NCNP	44,592	7474	5951	16,139	0	4511	10,517	0
NCGM	55,229	11,183	31,691	12,140	0	0	0	215
NCCHD	3239	0	1040	1450	0	0	749	0
NCGG	37,270	11,659	12,722	11,377	0	387	541	584

¹Including pre-extraction.

²Urine, feces, body fluid, etc.

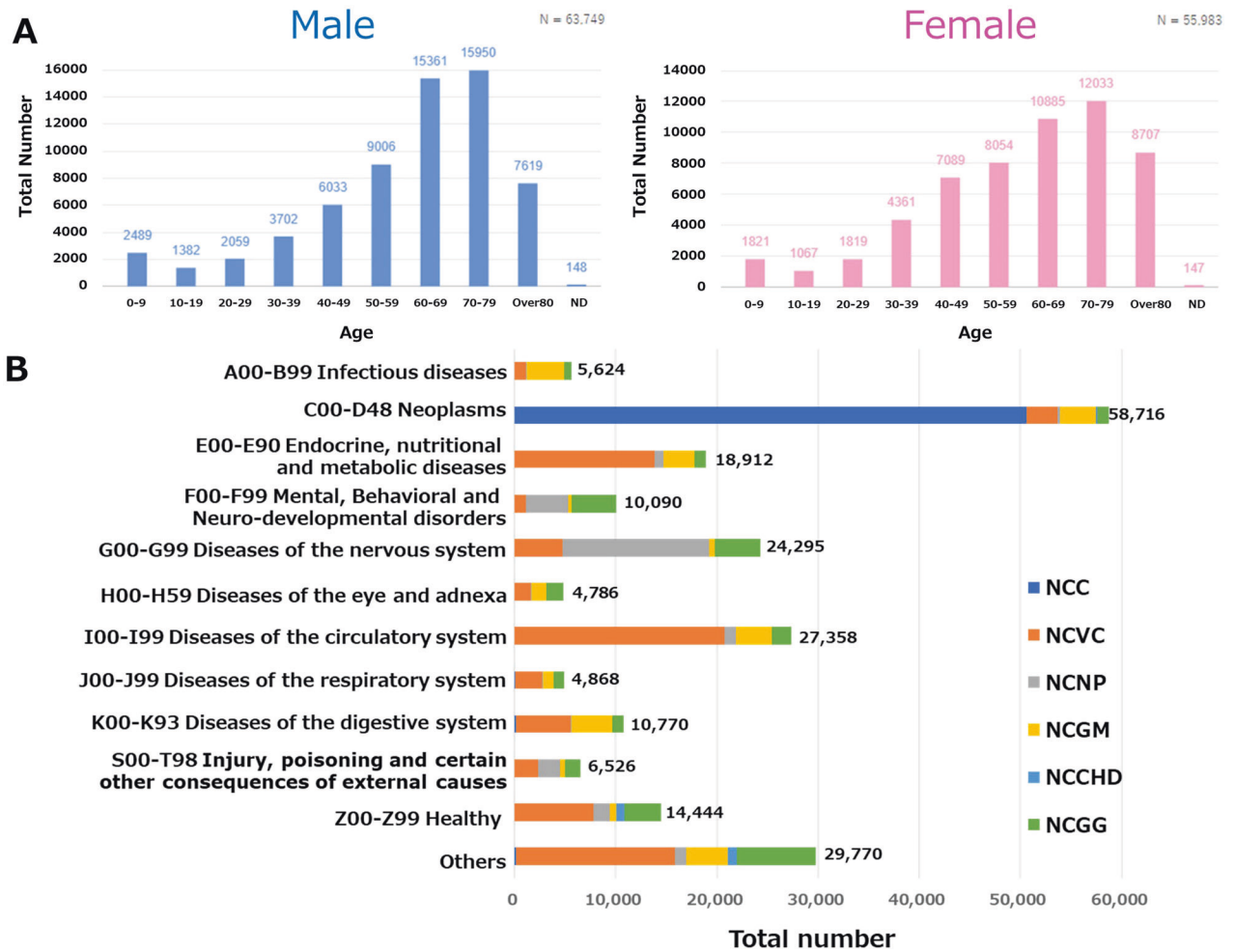


Fig. 2 Overview of registered people in the NCBN Catalogue Database [as of March 2022]. A Distribution of each sex and age classification. **B** Distribution of each disease classification.

Table 3. Overview of information in the NCBN catalogue database.

Basic patient information	Date of birth, sex, height, weight, blood pressure
Medical interview information (medical history)	Cancer, hypertension, diabetes, stroke, heart disease, kidney disease, liver disease, mental disease, current condition
Medical interview information (life history)	Family history, drinking history, smoking history, allergy history, surgical history, blood transfusion history
Disease information	ICD10 code of primary disease and comorbidities, or corresponding MEDIS number
Biological sample information	Type, date of collection, method of storage

medical history, family history, smoking or drinking history, and availability of genomic information. In the search by bioresource type, it is also possible to search for the number of samples that can be utilized in companies and academia through sample distribution.

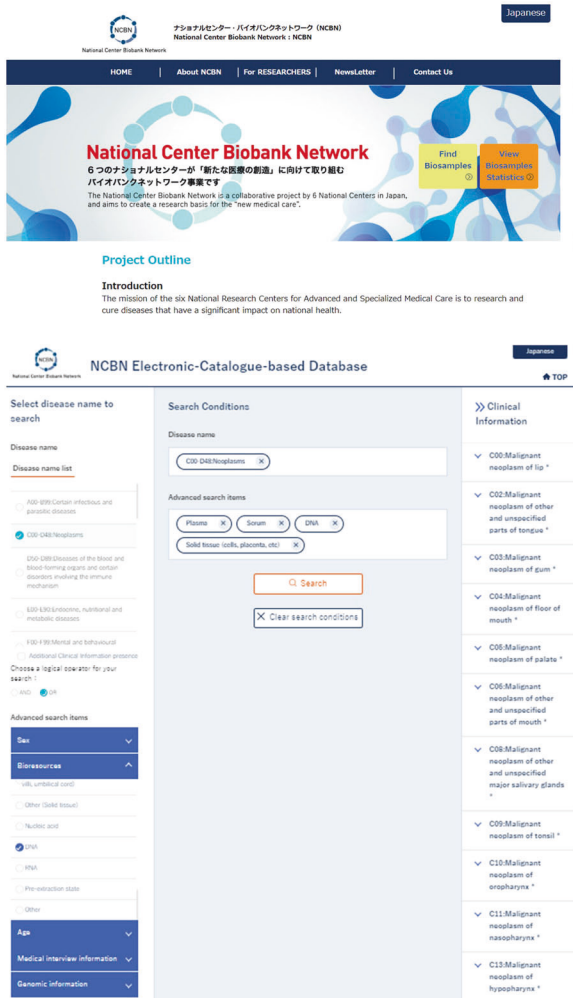
PROVISION OF NCBN BIORESOURCES

Contacting the NCBN secretariat

Inquiries from outside parties can be made by e-mail to the central biobank or to individual NC biobanks to obtain information regarding the availability of samples. The 6NCs cover a wide range

of diseases, and the types of samples and storage methods are diverse. Some NCs may have bioresources for which detailed data are not available in our catalogue database. If the catalogue database shows that samples matching the inquiry exist at more than one NC, or if more detailed information is required, the NCBN secretariat can be contacted using the “Contact” button. The referring NCBN Catalogue DB search result ID should be provided as the reference (<https://www.ncbiobank.org/en/contactus.php>).

The NCBN establishes a common platform among the 6NCs for comprehensive informed consent acquisition, enabling future genomic analysis and utilization for research development by companies as the result of our standardization of



How to use NCBN Catalogue database

1. Access the Catalogue via the database search page.
2. Select the target disease name and the required sample type. You can search by ICD10 code and MEDIS number, or select from the ICD10 classification list. You can also narrow your search by including medical interview information.
3. Find the number of samples matching your search condition for each of the 6 NCs. If you require more detailed information or wish to use the samples for your research, please contact the NCBN secretariat using the "Contact" button.

Result

	NCC	NCVC	NCNP	NCGM	NCCHD	NCGG
Total number of registered patients (people) in database	51099	22303	19138	14186	1744	12162
Appropriate number of patients (people)	50585	3080	200	3535	109	1184
Bioreources Stored Samples: Number of Available Samples						
Plasma	43757	3610	194	3561	0	1368
Serum	0	3701	176	7044	3	1405
Solid tissue	21661	0	61	0	107	0
DNA	43757	3006	127	3796	3	1423

Unit : people

[Contact](#)

Fig. 3 How to use the NCBN Catalogue Database. The bioresource information users need can be searched via the "Search Conditions" page in the open-access NCBN catalogue database. The NCBN secretariat can be contacted using the "Contact" button in the "Result" page.

Table 4. Summary status of utilization of 6NC biobanks.

NC	Total number of collaborative studies	Total number of collaborative institutions			Total number of provision ¹		Total number of relevant publications
		Company	Academia	Other	Collaborative research	Distribution	
NCC	294	237	308	117	39,208 (827)	0 (0)	933
NCVC	61	13	95	28	13,493 (119)	56 (8) ²	51
NCNP	104	26	107	84	11,036 (110)	2,108 (59)	310
NCGM	126	21	156	74	20,603 (31)	914 (27)	150
NCCHD	82	0	90	6	2185 (10)	352 (7)	172
NCGG	121	44	237	147	58,921 (145)	1,917 (15)	183

¹Total number of provision bioreources (times).

²Exceptional distribution for educational purposes.

ethical review. There are two types of provision of bioreources for research use: collaborative research and distribution. Collaborative research involves clinicians and researchers within each NC as coresearchers and is appropriate in the situation when a wealth of clinical information is needed, when the recipient needs to collect the desired medical information anew, or when clinicians and researchers familiar with the characteristics of the disease itself are needed. In contrast,

distribution is suitable for research that can be conducted with limited ancillary information and can be made into a Material Transfer Agreement (MTA) in which the provider biobank does not claim intellectual property rights through consultation. At present, only four NCs (NCNP, NCGM, NCCHD, and NCGG) can offer distribution services, but we are aiming to make a distribution available to all NCs and simplify and unify the procedures as much as possible.

Experience in sample distribution and collaborative research

The summary status of the utilization and publications of each NC biobank is summarized in Table 4. This status is also reported in the NCBN annual report, which is accessible through the NCBN website, and data from the provider institutions are also posted on the website for each NC biobank^{15–20}. The “Research activities and achievements” section of the website (<https://ncbiobank.org/research/index.php>) lists the number of collaborative research projects conducted using NCBN samples and information, the number of collaborative research institutions such as companies and academia, and information about results published since 2010. These are all listed by NC and by year.

CONCLUSION

In this review, we provide an overview of the NCBN and NCBN Catalogue Database. The six NCs have continuously collected a wide variety of patient-derived bioresources, including detailed medical information for the diseases in question over a long span of 10 to 20 years, and the NCBN constructs a networked and federated organization of the 6NC biobanks to utilize patient-derived bioresources for medical and research purposes. The NCBN Catalogue Database is an open-access database for the efficient provision of registered biological resources and data for use in research communities. Recently, the total number of obtained genome-wide data for NCBN bioresources has increased for many diseases through whole genome sequencing analysis or SNP array analysis, and the NCBN is creating a system that can provide both bioresources and genomic data to users. Quality control is also important to scientifically support the research use of various sample types. In this sense, it is essential to standardize the hardware and software aspects of registering information in terms of sample collection methods, processing processes, storage conditions, and storage periods. The global standardization of biobanking is steadily progressing, and the NCBN is preparing for ISO20387:2018 biobanking certification²¹.

REFERENCES

- Gaskell, G. & Gottweis, H. Biobanks need publicity. *Nature* **471**, 159–160 (2011).
- Sudlow, C. et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
- Denny, J. C. et al. The “All of Us” Research Program. *N. Engl. J. Med.* **381**, 668–676 (2019).
- AMED Biobank cross-search: <https://biobank-search.megabank.tohoku.ac.jp/v2/en/>.
- Nagai, A. et al. Overview of the BioBank Japan Project: Study design and profile. *J. Epidemiol.* **27**, S2–S8 (2017).
- Kuriyama, S. et al. ToMMo the tohoku medical megabank project: design and mission. *J. Epidemiol.* **26**, 493–511 (2016).
- Arakawa, A. et al. Vaginal transmission of cancer from mothers with cervical cancer to infants. *N. Engl. J. Med.* **384**, 42–50 (2021).
- Kuyama, N. et al. Circulating mature PCSK9 level predicts diminished response to statin therapy. *J. Am. Heart Assoc.* **10**, e019525 (2021).

- Saito, K. et al. Profiling of cerebrospinal fluid lipids and their relationship with plasma lipids in healthy humans. *Metabolites* **11**, 268 (2021).
- Sugiyama, M. et al. Serum CCL17 level becomes a predictive marker to distinguish between mild/moderate and severe/critical disease in patients with COVID-19. *Gene* **766**, 145145 (2021).
- Kasuga, Y. et al. Epigenetic changes in neonates born to mothers with gestational diabetes mellitus may be associated with neonatal hypoglycaemia. *Front. Endocrinol.* **12**, 690648 (2021).
- Akiyama, S. et al. JAMIR-eQTL: Japanese genome-wide identification of microRNA expression quantitative trait loci across dementia types. *Database.* **2021**, baab072 (2021).
- NBDC Human Database: <https://humandbs.biosciencedbc.jp/en/>.
- Kamada, M. et al. MGEND: an integrated database for Japanese clinical and genomic information. *Hum. Genome Var.* **6**, 53 (2019): <https://mgend.med.kyoto-u.ac.jp/>.
- Annual report of NCBN in FY 2021: https://ncbiobank.org/loadmap/pdf/annual_report_r03.pdf.
- Achievement in NCC biobank: <https://www.ncc.go.jp/jp/biobank/achievement/index.html>.
- Achievement in NCVC biobank: <https://www.ncvc.go.jp/oic/divisions/biobank/achievement/>.
- Activities in NCNP biobank: https://www.ncnp.go.jp/mgc/bio_31.html.
- Achievements in NCGM biobank: <http://biobank.ncgm.go.jp/achievements2.htm>.
- Activities in NCGG biobank: <https://www.ncgg.go.jp/ri/biobank/kenkyu/index.html>.
- International Organization for Standardization ISO 20387:2018 Biotechnology – Biobanking – General Requirements for Biobanking. (2018).

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Yosuke Omae.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022