



BRIEF REPORT

Opening of the National Biobank of Korea as the Infrastructure of Future Biomedical Science in Korea

Sang Yun Cho^a, Eun Jung Hong^a, Jung Min Nam^a, Bogkee Han^a,
Chaeshin Chu^b, Ok Park^{a,*}

^aDivision of Biobank for Health Sciences, Korea National Institute of Health, Osong, Korea.

^bDivision of Epidemic Intelligence Service, Korea Centers for Disease Control and Prevention, Osong, Korea.

Received: July 2, 2012
Revised: July 18, 2012
Accepted: July 19, 2012

KEYWORDS:

biobank,
biospecimen,
cohort,
Korea Biobank Project,
Korea Biobank Network,
National Biobank of Korea

Abstract

On April 26, 2012, the Korea National Institute of Health officially held the opening ceremony of newly dedicated biobank building, 'National Biobank of Korea'. The stocked biospecimens and related information have been distributed for medical and public health researches. The Korea Biobank Project, which was initiated in 2008, constructed the Korea Biobank Network consisting of the National Biobank of Korea (NBK) with 17 regional biobanks in Korea. As of December 2011, a total of 525,416 biospecimens with related information have been secured: 325,952 biospecimens from the general population obtained through cohort studies and 199,464 biospecimens of patients from regional biobanks. A large scale genomic study, Korea Association Resource (KARE) and many researches utilized the biospecimens secured through Korea Genome Epidemiology Study (KoGES) and Korea Biobank Project (KBP). Construction of 'National Biobank of Korea', a dedicated biobank building at Osong means that NBK can manage and check quality of the biospecimens with promising distribution of 26 million vials of biospecimen, which provide the infrastructure for the development of health technology in Korea. The NBK and the National Library of Medicine (to be constructed in 2014) will play a central role in future biomedical research in Korea.

1. Introduction

The human genome project has completed mapping the human genome, and the next-generation sequencing

technique expands the availability of personal DNA sequencing, which increases the possibility of customized medical treatment. This means that the bottleneck of customized medical treatment is transitioned from the

*Corresponding author.

E-mail: okpark8932@gmail.com

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

technical field to the fields of biological samples and data/knowledge management. The role of biobanks as an infrastructure of human biospecimen supply is important for the development of customized medical treatment, development of new drugs, and prevention of diseases [1,2]. A biobank stores human biospecimens from volunteer participants with various clinical conditions, conducts quality checks, creates a database to handle related information (e.g., clinical/epidemiological data), and distributes biospecimens for researches [3]. Well-documented, up-to-date epidemiological data, along with extensive human biospecimen information connected with clinical and biological information, are necessary to draw statistically significant, effective results in retrospective researches [3]. Furthermore, it is also the role of the biobank to ensure the participants' rights and to provide the ethical and legal framework for the use of biospecimen in biobanks [4].

The UK biobank, one of the major population-based biobanks, has secured biospecimens of 500,000 volunteers aged 45–69 years with related information and distributed these biospecimen to researchers in 2012 [5]. The Kadoorie Study collects Chinese human biospecimens; the baseline survey on 0.5 million participants was completed in 2008, and follow-up resurvey was subsequently performed [6]. In addition, many countries including the United States, EU, Mexico, and India, have constructed and run large-scale biobanks with more than 200,000 participants [7]. There are various biobank networks to draw collaboration among biobanks and standard operating procedures (SOPs).

Biobanks such as P3G (Public Population Project in Genomics; www.p3g.org) and BBMRI (Biobanking and Biomolecular Resources Research Infrastructure; www.bbMRI.eu) discuss challenges and critical issues for development, success, and business continuity for biobanks, and look for harmonization among biobanks [8–10].

The Korea National Institute of Health (KNIH) organized a task force to plan the construction of a biobank in 2007 and started the Korea Biobank Project (KBP) in 2008 to set up biobanks and a biobank network. To systematically manage the biospecimens of population-based cohorts, which was already carried out and run in the KNIH, and collect biospecimens for chronic diseases and rare diseases such as cancers, diabetes, and hypertension, the Korea Biobank Network was established including the National Biobank of Korea (NBK) and 17 regional biobanks within 17 university hospitals (Figure 1).

This report discusses the details of the 2008–2012 biobank construction in Korea (the first term of the KBP) and major achievements (e.g., biospecimen collection and creation of database) along with a description of the facilities and instruments used in the NBK, and to share the experience in running a large-scale human biospecimen network.

2. Overview of the Korea Biobank Project

The KBP's (first term, 2008–2012) aim was to secure a total of 500,000 Korean human biospecimens



Figure 1. Location of National biobank and regional biobanks in Korea.

(consisting of 300,000 specimens from the general population and 200,000 from diseased participants). The NBK secured the general population's biospecimens from various cohorts organized by the KNIH since 2001. Regional biobanks secured disease-based biospecimens in 17 university hospitals, representing diseases such as cancers, chronic diseases, and rare diseases. As of December 2011, a total of 525,416 specimens (325,952 from the general population and 199,464 from diseased populations) have been secured, reaching the first term's goal 1 year early (Figure 2).

When KBP began in 2008, eight regional biobanks were established; four, one, and four additional regional biobanks were constructed in 2009, 2010, and 2011, respectively. Each regional bank can perform an independent biobank function to secure, store, and distribute by input of standardized human biospecimen information and clinical information under Biospecimen Information Management System (BIMS), run by NBK. The second term (2013–2015) plans to provide the user-friendly, one-stop distribution service of KBN in a coordination center.

NBK's functions include securing and distributing population-based cohort biospecimens, planning of the KBP project, providing education for regional biobank staffs, handling public relations, and construction and running of BIMS and the KBP portal site (<http://kbn.cdc.go.kr>).

3. Status of Population-based Human Biospecimens

The population-based human biospecimens come from population-based cohorts, separately run by the KNIH (Division of Epidemiology and Health Index, Division of Health and Nutrition Examination Survey),

including the Korean Genome Epidemiology Study (KoGES), Korea National Health and Nutrition Examination Survey, and research projects of the KNIH such as Polycystic Ovarian Syndrome Cohort and AIDS Cohort (Table 1).

KoGES is a large-scale Korean genome epidemiology cohort with 220,000 participants as of 2012. The cohort started from the Ansong–Ansan Community-based Cohort in 2001. The community-based cohort is composed of 10,038 participants aged 40–69 years (5,020 in Ansan, 5,018 in Ansong); it distributes questionnaires and performs biospecimen collection and physical examination every 2 years. About 2,000 items for each person are collected including lifestyle, medical history, physical activity, food consumption, disease-related blood test results and body measurement. The information and biospecimens from the rural Ansong community and urban Ansan community cohort are utilized in analytical studies on interaction of risk factors of chronic diseases and genomic characteristics and observational studies on environmental risk factors. The fifth follow-up examination was completed in 2011, and now the sixth follow-up examination is under progress (<http://biomi.cdc.go.kr>).

The production of human biospecimens is conducted by a company contracted with NBK. Blood samples from the nationwide collection sites using preconfirmed standard protocol are transported by vehicles equipped with refrigerators (set at 4 °C) and thermal monitoring systems, and are turned into biospecimens of blood plasma, serum, and DNA and temporarily stored at –80 °C. Next, the biospecimens in refrigerators are transported to NBK and undergo examination and quality check, then classified and stored according to SOPs into three classes: for distribution, backup, or permanent reference.

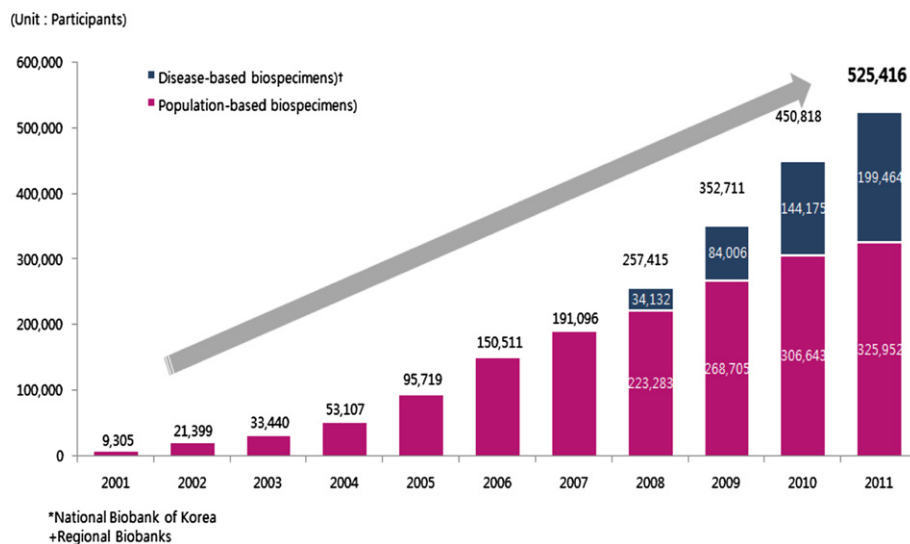


Figure 2. Number of participants in biobanks in Korea.

Table 1. Status of cohorts in Korea National Institute of Health (as of Jan. 2012)

Category	Cohort	Participant(s)
Korean Genome and Epidemiology Study	Population-Based Cohort	225,850
	Gene-Environmental Model Cohort	18,556
	Community-Based Cohort	12,046
	Other Cohort	3,070
<i>Subtotal</i>		259,522
Disease Prevention Research	The Korea National Health & Nutrition Examination Survey	46,127
	National Infectious Disease Serosurveillance Study	2,139
<i>Subtotal</i>		48,266
Immunology and Communicable Disease Research	National Measles Serosurveillance Study	16,251
	National Brucellosis Survey	7,439
	AIDS Cohort	1,072
<i>Subtotal</i>		24,762
Genome and Biomedical Science Research	Polycystic Ovarian Syndrome Cohort/Genome Research	2,704
	Major Birth Defect Cohort	2,533
	Korea Longitudinal Study of Aging	911
	Avellino Cohort	871
	Polycystic Ovarian Syndrome Cohort/Biomedical Science Research	633
	Primary Care Family Cohort	514
<i>Subtotal</i>		8,166
Total		352,952

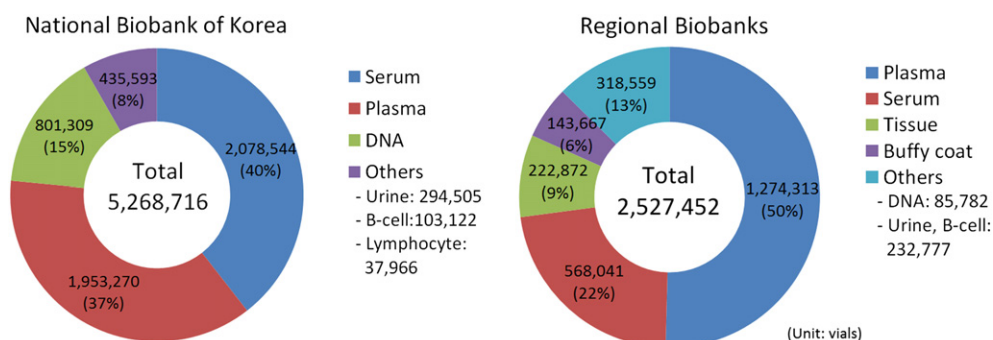
For population-based biospecimens of participants including community-based cohorts, 10 vials of 300 μ L of blood plasma and serum from a participant are collected and stored at -185°C , blood DNA of 80–100 μ g collected from whole blood is divided into two sets of 20- μ g vials, and the rest of the DNA is put into one vial and stored at -75°C . Midstream urine of 10 mL is collected and put into a 15-mL conical tube and stored at -75°C . For maximizing the efficacy of distribution, the automated biospecimen storage and distribution system operate with 300,000 vials of DNA.

The numbers and types of biospecimens secured in the first term of KBP are shown in Figure 3. A total of 5,268,716 vials and tube of human biospecimen have been secured: 2,078,544 sera, 1,953,270 plasmas, 801,309 DNAs, and 435,593 others [including urine and lymphoblastoid cell line (LCL), B-cell]. The human

biospecimens have been utilized in an intramural research of KNIH, Korean Association Resource (KARE) project, and also on extramural research.

4. Status of Disease-related Human Biospecimens

KBP's disease-related human biospecimens are secured through 17 regional biobanks located in 17 university hospitals. Figure 4 shows the number and type of biospecimens obtained from the first-term KBP in 17 regional biobanks. The proportion of neoplastic diseases is 45%, whereas that of non-neoplastic diseases is 55%. The characteristic of a regional biobank determines the kinds of biospecimens that are collected. For example, Kyung-sang University Hospital specializes in childhood

**Figure 3.** Total number of collected biospecimens in KBP (as of Dec. 2011).

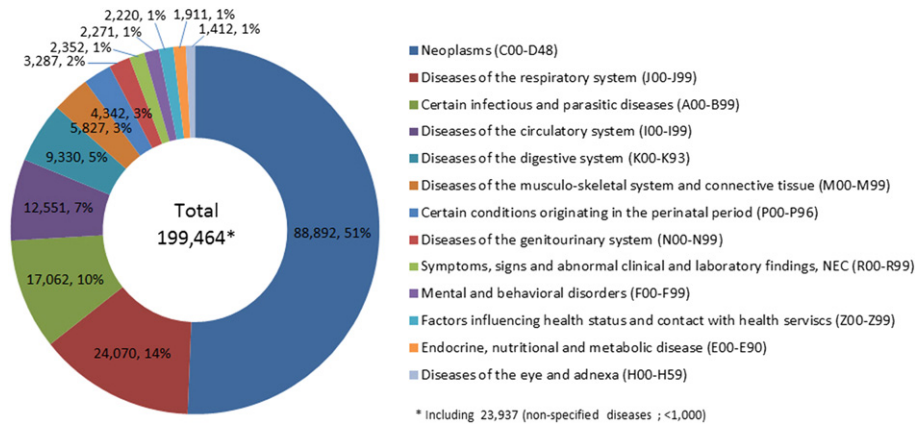


Figure 4. Number of disease participants in regional biobanks of Korea Biobank Network. Disease categories : KCD-6 (Korean Standard Classification of Disease), Unit : Participants.

diseases, of gastritis, and colitis, Kyungpook National University Hospital concentrates on intractable geriatric diseases, Wonkwang University Hospital places emphasis on immune diseases and gastric disease, and Soonchunhyang University specializes in asthma and lung diseases. The kind of biospecimens collected on cancer and chronic diseases depends on the hospitals in which the regional biobanks are located. The paired biospecimens of cancer and normal tissues are common biospecimens in all regional biobanks and stored in paraffin block, optimal cutting temperature block, and fresh frozen forms.

The information on a disease-related biospecimen found in the EMR (Electronic Medical Record) in a hospital synchronizes with that in BIMS, and automatically inputs the predefined items into BIMS in some biobanks. For reference in this procedure, Biobank Data Dictionary, which NBK has developed and distributed, is used for understanding and systematic and consistent integration of terminology. All information on participants is made anonymous to protect their identity and guidelines on research ethics are strictly followed.

As of December 2011, the regional biobanks have amassed a total of 2,572,452 vials: 1,274,313 plasmas, 568,041 sera, 222,872 tissues, 143,667 buffy coats, and 318,559 others (DNA, urine and LCL) (Figure 3).

5. Status and Results of the Distribution in the First Term KBP

The first term KBP (2008–2012) focuses on securing Korean human biospecimens and constructing and running the network. Since 2001, when biospecimen distribution of KNIH started, until the end of 2011 (when the first term of KBP ended), collected biospecimen and information have been distributed for a total of 454 research projects: 117 projects from NBKs and 347 projects from regional biobanks.

DNA biospecimen distribution takes the lion's share (163,939, or 87%) of the total number of distributions

(188,380). It is followed by distribution of serum (14,971, 8%), plasma (6,515, 3%), and other specimens including cells (2,958, 1.6%) (Figure 5). Most studies have been conducted by using population-based human biospecimens of the Ansan and Ansong Cohort [11]. A large-scale genomic study, the Korean Association Resource (KARE) project, analyzed GWAS (genome-wide association study) utilizing 8,842 Korean genomic DNA samples secured through KoGES and KBP, and reported the identified genomes related to hypertension, the waist/hip ratio, and bone mineral density [12]. In 2011, Cho found and reported the common loci for metabolic diseases in East Asians after a large-scale GWAS analysis such as repetitive comparison analysis of 30,395 participants' DNA using Korean Association Resource and the Health Examinee cohorts with Biobank Japan genome-wide association study and Health2 and Shanghai Jiao Tong University Diabetes Cohort in China [13]. The information and biospecimens, which were distributed to cohort researchers and general researchers, made it possible to identify the diet patterns causing certain diseases [14], index of sleeping pattern

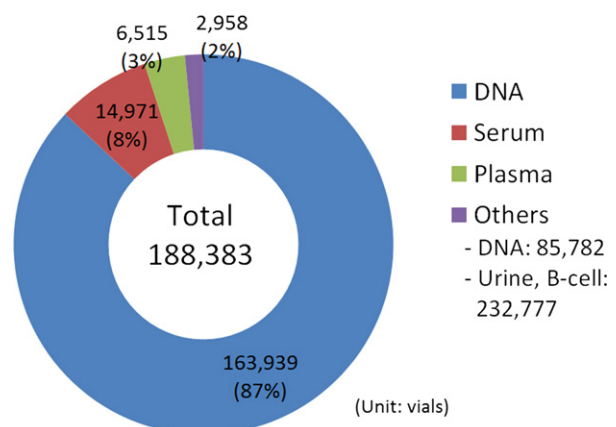


Figure 5. Number of distributed biospecimens (2001–2011).

(e.g., snoring) [15], and genome relating to the height of Koreans [16].

As the KBP continues to operate, requests for distribution increase. The automated biospecimen storage and distribution system was introduced for quick distribution of preserved biospecimens. The system has distributed 33,331 biospecimens since its launching in June 2010.

A total of 89 papers have been produced through the studies using distributed human biospecimens from NBK: 11 intramural research studies from the KNIH, 18 extramural research, and 60 collaboration researches (Figure 6).

6. Quality Check of Biospecimens in KBN

The standardization of management on human biospecimen, the manual, regulations, guidelines for biospecimen distribution, and working-level guidelines are documented in “Work about Collection, Storage, Distribution of Biospecimens,” which obtained ISO 9001:2008 quality management system accreditation from the Korea Standard Association in November 2005 and has maintained this accreditation status with periodic post-evaluations.

Quality assurance on human biospecimens consists of DNA stability test by electrophoresis, DNA purity test by spectrophotometer, microorganism contamination test by polymerase chain reaction, cross-contamination test by analysis of homogeneity. All biospecimens collected in KoGES undergo DNA degradation test and

purity test, and the quality assured data with biospecimens are saved. Ten percent of stocked biospecimens are randomly selected and tested for the quality assurance test to ensure good quality of DNA biospecimens. The same is true for LCL quality assurance; 10% of the stocked biospecimens are randomly selected to undergo survival rate investigation with cross-contamination test.

7. Opening of Newly Constructed NBK Building and Facilities

Construction of the NBK was finally completed (area = 12,301 m²), and the opening ceremony was held on April 26, 2012 at the Osong Health Technology and Administration Complex, Chungcheongbuk-do, Korea (Figure 7). The three-story building including a basement houses storage rooms that can contain 26 million vials of biospecimens; this volume is expected to increase as soon as the automated large archive and related facilities will be installed. The structure also houses laboratories (for quality assurance of biospecimen and pathogens), an auditorium with 150 seats, a main video conference room, nine conference rooms, and staff offices. The budget for the project was secured from the Korean central government in 2008, and it took 24 months to complete the construction (starting from April 2010), shortly followed by a preliminary operation period.

As of 2012, NBK has used two different types of freezer to store the biospecimens steadily for a long

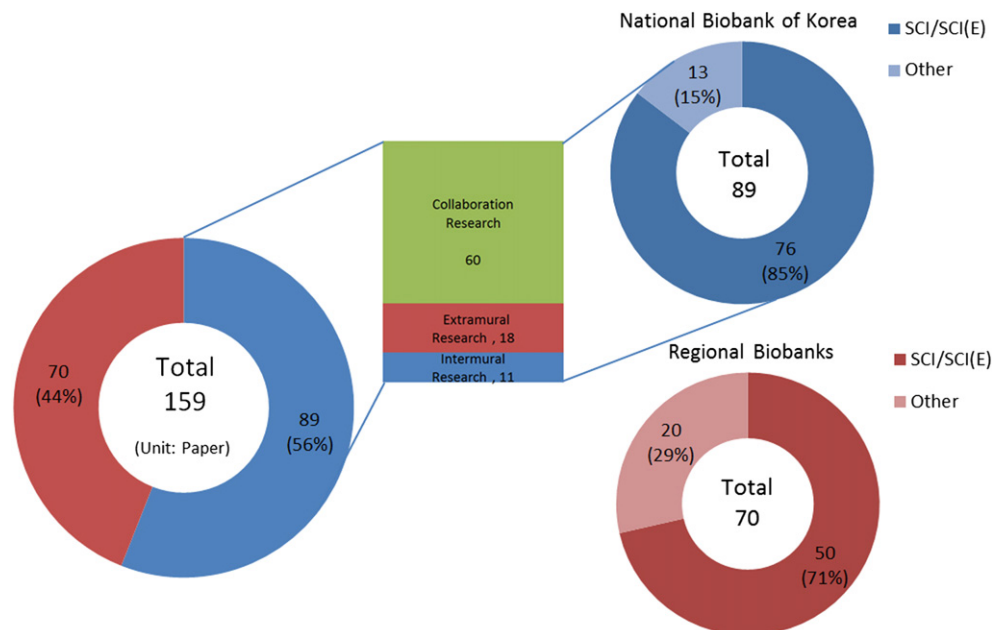


Figure 6. Total number of publications (As of Dec. 2011).

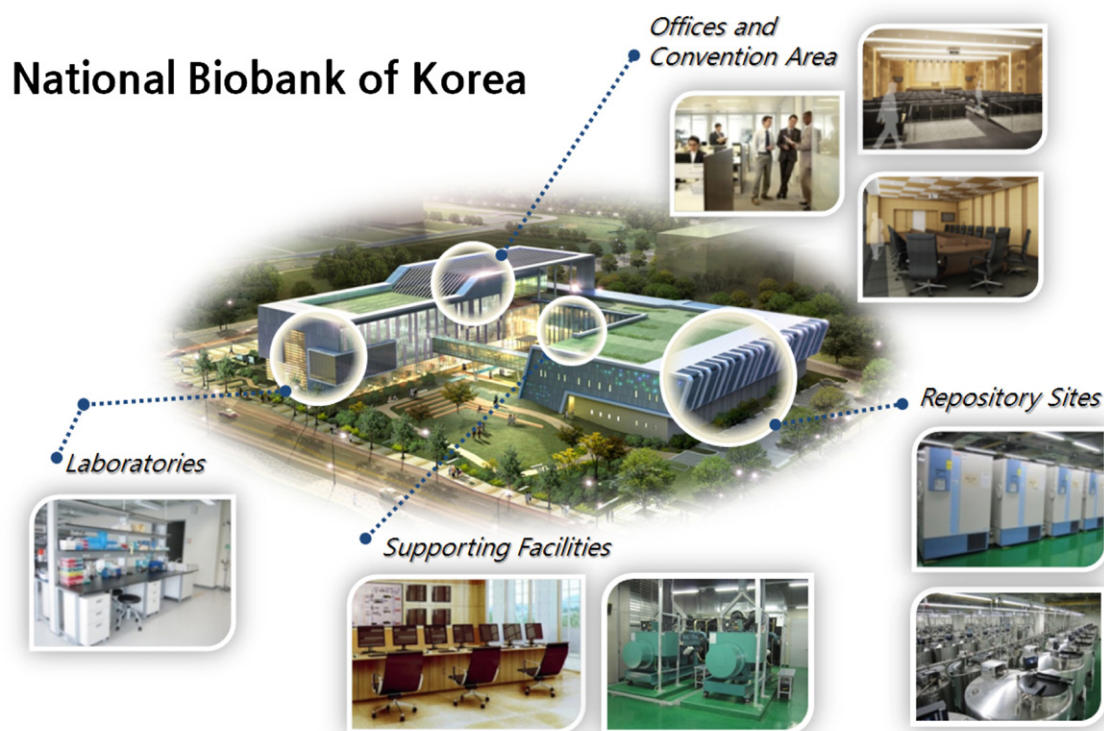


Figure 7. Sectors in the National Biobank of Korea.

time: 102 deep freezers (working temperature: -75°C) and 172 liquid nitrogen containers (working temperature: -185°C). Those freezers are racked to maximize the storage space and to distribute the biospecimens efficiently. One deep freezer stores 48,600 vials of DNA biospecimens or 8,640 urine biospecimens, while one liquid nitrogen container holds 34,600 or 70,000 vials of sera and plasma. The automated biospecimen storage and distribution system administers 300,000 vials of specimens to maximize the distribution of biospecimens.

The main storage rooms (an area of $3,202\text{ m}^2$) consist of two floors; the first floor is for the deep freezers while the second floor is for the liquid nitrogen containers. The backup storage space is $1,256\text{ m}^2$. The current storage space holds a maximum of 219 deep freezers and 544 liquid nitrogen containers with a maximum storage capacity of 25,357,540 vials of human biospecimen. If the planned automated large archive is installed, the capacity will expand drastically.

The main storage equipment, such as deep freezers, are electrically operated with mechanical compressors, so that the backup is available in case of power failure or malfunction. All deep freezers are equipped with an automated voltage regulator; if it experiences malfunction, the operator will be notified by the temperature monitoring system. In case of a power failure, a liquid nitrogen backup system is available. Two 600-KW emergency generators are designed to maintain the

maximum number of deep freezers and liquid nitrogen containers. An additional option is provided by means of a separate oil tank, which can operate for 48 consecutive hours.

To prevent the possible occurrence of suffocation among the staff (as a result of the use of large-scale liquid nitrogen containers), each storage room is equipped with a measuring device to measure oxygen concentration; the concentration of oxygen is displayed outside the room. Inside the room, double ventilation and air-conditioning system are always turned on. In case of fire, the automated fire extinguisher system is designed to disperse inert gas (HFC-23) to avoid harming the equipment or biospecimens. Closed-caption TVs are installed inside and outside the building with a 24-hour, 7-day continuous video-recording system to protect the facilities and information.

8. Conclusion

A biobank is an indispensable infrastructure of the post-genome era and continues to evolve in a sophisticated manner. To facilitate the next generation of genome studies for customized medicine and development of biomarkers, the support of an advanced biobank is necessary. To enhance the quality and availability of specimens, a biobank network system is run, and the networking will be even more enhanced. In the future,

the SOPs and guidelines suggested by this network will dominate the biobank management and research on biobank specimens. To keep up with this trend, the NBK has obtained accreditation of biobank management system (ISO 9001: 2008) and the establishment of human biospecimen management standard protocol, so that it can enhance the stability and quality of specimens by constructing NBK with 17 regional biobanks and ensuring good quality specimen and information. With the introduction of an automated biospecimen storage and distribution system, the NBK enhances the precision and efficiency of the distribution of human biospecimens, and continues to develop quality assurance techniques and research on the utilization of biospecimens. In the future, the NBK will not only participate in domestic and international biobank networks, but will also seek to encourage the standardization of biospecimen quality and harmonization of biobanks. By doing so, it seeks to promote biomedical researches for the genome protein study to treat and prevent diseases, as well as the development of new medicines, that will eventually lead to realizing customized medical treatment.

References

1. Zika E, Gurwitz D, Ibarreta D. Pharmacogenetics and pharmacogenomics: state-of-the-art and potential socio-economic impact in EU. European Commission DG JRC/IPTS. EUR 2006;22214.
2. Austin MA, Harding S, McElroy C. Gene-banks: a comparison of eight proposed international genetic databases. *Community Genet* 2003;6(1):37–45.
3. Yuille M, van Ommen G, Brechot C, et al. Biobanking for Europe. *Brief Bioinform* 2008 Jan;9(1):14–24.
4. Cambon-Thomson A, Rial-Sebbag E, Knoppers BM. Trends in ethical and legal framework for the use of human biobanks. *Eur Repir J* 2007 Aug;30(2):373–82.
5. Watts G. UK Biobank opens its data vaults to researchers. *BMJ* 2012 Mar;344:e2459.
6. Chen Z, Chen J, Collins R, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow up. *Int J Epidemiol* 2011 Dec;40(6):1652–66.
7. GBI research, Biobanks – 2011 Yearbook (2011) Survey report.
8. Knoppers BM, Fortier I, Legault D, Burton P. The public population project in genomics (P3G): a proof of concept? *Eur J Hum Genet* 2008 Jun;16(6):664–5.
9. Yuille M, van Ommen GJ, Brechot C, et al. Biobanking for Europe. *Brief Bioinform* 2008 Jan;9(1):14–24.
10. European Commission. Final report on the networking meeting for EU-funded biobanking projects. Brussels: DG Research. Available from: http://cordis.europa.eu/fp7/health/past-events_en.html; 2008.
11. Lim S, Jang HC, Lee HK, et al. A rural–urban comparison of the characteristics of the metabolic syndrome by sex in Korea: the Korean Health and Genome Study (KHGS). *J Endocrinol Invest* 2006 Apr;29(4):313–9.
12. Cho YS, Go MJ, Kim YJ, et al. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat Genet* 2009 May;41(5):527–34.
13. Kim YJ, Go MJ, Hu C, et al. Large-scale genome-wide association studies in East Asians identify new genetic loci influencing metabolic traits. *Nat Genet* 2011 Sep;43(10):990–5.
14. Ahn Y, Park YJ, Park SJ, et al. Dietary patterns and prevalence odds ratio in middle-aged adults of rural and mid-size city in Korean Genome Epidemiology Study. *Korean J Nutr* 2007 Apr;40(3):259–69.
15. Joo S, Lee S, Choi HA, et al. Habitual snoring is associated with elevated hemoglobin A1c levels in non-obese middle-aged adults. *J Sleep Res* 2006 Dec;15(4):437–44.
16. Kim JJ, Lee HI, Pakr T, et al. Identification of 15 loci influencing height in a Korean Population. *J Hum Genet* 2010 Jan;55(1):27–31.