

# National population-based biobanks for genetic research

Helen Swede, PhD<sup>1,3</sup>, Carol L. Stone, PhD<sup>2</sup>, and Alyssa R. Norwood, BA<sup>3</sup>

Clinical practice guidelines derived from genetic research using population-based biobanks could dramatically change the nature of personal and public health medicine. Centralized population-based biobanks have been established or proposed in at least nine countries to date, and many lessons have been learned from these landmark developments. Scientific and governmental leaders in the United States are currently contemplating pending federal legislation regarding the establishment of centralized and networked biobanks. Public health practitioners and clinical care providers may be called on to serve pronounced planning roles at the state level. Possible responsibilities include: formulating legislation, gathering public comment, reviewing research proposals, and developing procedures for informed consent, participant withdrawal, and confidentiality protection. State health agencies may also need to create and/or administer banking facilities. Proper planning may ensure that individual rights are protected while research benefits are maximized. **Genet Med 2007;9(3):141-149.**

**Key Words:** MeSH tissue banks, biobanks, Human Genome Project, genetic predisposition to disease, genomics, public health genomics

## INTRODUCTION

With the first draft DNA sequence of the human genome published in *Nature*<sup>1</sup> and *Science*<sup>2</sup> in February 2001, and the full sequence published in April 2003, the completed Human Genome Project<sup>3</sup> ushered in a new era of promise for medicine. Sequencing of the entire human genome was an enormous scientific triumph that laid the groundwork for the daunting task of analyzing disorders that result from the dynamic interplay between genetic and environmental factors, yet translation into clinical benefit is not yet fully at hand.<sup>4</sup> Recognizing both the power of a population-based approach to search for complex disease susceptibility genes and the large number of samples required for these studies, a number of countries have established population-based biobanks.<sup>1</sup> Discussions about a biobank for the United States are expected to continue in the coming year given that one of the provisions of the Genomics and Personalized Medicine Act of 2006 (S.3822), currently under consideration in the United States Senate, calls for a feasibility study of a national biobanking research initiative.<sup>5</sup>

Introduction of the United States Senate bill marks a pivotal point in genetic research. Much discussion has taken place about establishing a central national biobank<sup>6,7</sup> or a network of linked repositories,<sup>8</sup> yet, agreement about a specific public health approach or even the necessity of the entire venture has been elusive.<sup>8,9</sup> As debate and planning now move forward, an informed public health community can contribute to the enlarged dialogue and develop a vision for enhancing workforce competencies in anticipation of a major national effort. We present the following overview of experiences encountered in the international community with an emphasis on the ethical, legal, and social implications (ELSI) of population-based biobanking. Should a nationally sponsored initiative come to fruition, state health departments may be required to serve a pronounced planning role among a variety of infrastructural matters, such as preparation of state legislation, formal consultation with the public, informed consent procedures, and confidentiality protections of data.<sup>8</sup> Guidance from practical experiences related to population-based biobanks remains warranted,<sup>10</sup> particularly for public health professionals who might not be familiar with the fundamentals or application of the ELSI framework in genetic research.

A discussion of these issues is also relevant for state governments that are considering establishing and administering population-based biobanks. On the near horizon, for example, is federal funding for biobanks to explore the genetic and environmental determinants of preterm birth, which may involve coordination within individual states between researchers and departments of public health, as well as among multiple states across the country.<sup>11</sup> This biobanking initiative is part of the PREEMIE Act of 2006, which was enacted into law by the U.S. Congress in December 2006 (Public Law 109-450).

From the <sup>1</sup>Connecticut Tumor Registry and <sup>2</sup>Family Health Section, Public Health Initiatives Branch, Department of Public Health, Hartford; <sup>3</sup>Department of Community Medicine & Health Care, University of Connecticut School of Medicine, Farmington, Connecticut.

Helen Swede, PhD, Division of Epidemiology & Biostatistics, Department of Community Medicine & Health Care, University of Connecticut School of Medicine, 263 Farmington Avenue, Farmington, CT 06030-6325. E-mail: hswede@uchc.edu

Submitted for publication October 30, 2006.

Accepted for publication December 29, 2006.

The authors declare no conflict of interest.

DOI: 10.1097/GIM.0b013e3180330039

## WHAT IS A POPULATION-BASED BIOBANK?

Biobanks have been established to study DNA or other molecular markers derived from peripheral circulating blood, buccal cells (i.e., epithelial cells from the inner cheek or mouth), blood cells from umbilical cords, urine samples, or diseased tissue for use in biomedical research. There is no shortage of human specimens in the United States; an estimated 300 million human biological samples are stored in hospitals and academic research centers.<sup>12</sup> A population-based biobank is a repository consisting of a large collection of biological tissue donated by thousands of individuals from the general population who might or might not have a specific disease.<sup>13</sup> Genebanks or genetic databases are repositories that analyze DNA exclusively in the pursuit of the genetic determinants of diseases,<sup>14</sup> whether or not specimens were derived from the public at large.

The most comprehensive type of population-based biobank is designed to link biomarkers with medical history and lifestyle information. Diseases that result directly from single gene mutations are rare;<sup>4,15</sup> hence, this link provides a powerful tool that can contribute to our understanding of the genetic and environmental determinants leading to diseases such as diabetes, Alzheimers disease, asthma, schizophrenia, and cancer, and adverse outcomes such as preterm birth and congenital birth defects.

Another defining characteristic of a population-based biobank is that it can be used to estimate allele frequency (i.e., proportion of people with a particular DNA sequence at a particular location within a gene) in the population because they are comprised of large numbers of participants from a wide cross-section of ethnic groups.<sup>8</sup> Quantification of the public health burden related to most gene variants reported to date has been impossible to gauge because genotype prevalence is still unknown.<sup>8,15</sup> A paradigm to illustrate the need for prevalence information is the family of genes involved in drug metabolism. Although it has been established that genetic variation in a variety of these genes is responsible for major differences in patient response to certain medications,<sup>4</sup> the clinical and economic utility of which variants to screen for depends not only on the effect of the variation, but also on whether the genetic variation is relatively common. Further, emerging evidence indicates that allelic frequency of many of these gene variants differs substantially among ethnic groups,<sup>16</sup> implying that targeted screening may be more appropriate.

There are numerous biobanks of residual tissue at hospitals (i.e., remains after clinical purposes have been met) or of specimens acquired in the context of a specific research investigation in academia and industry. These collections cannot provide true estimates of allele frequency in the population, however. Nor can biobanks devoted to single-gene disorders comprising patients and their families estimate this valuable information. Population-based biobanks are a more powerful tool for translating genetic discoveries into clinical practice. Further, the contribution of genomic factors to most common diseases is not understood in sufficient detail to affect recommendations for public health initiatives or patient care.<sup>8</sup> Hence, research made possible by population-based biobanks

may contribute to disease prevention through knowledge of the interplay of genetic and modifiable risk factors.<sup>15</sup>

## NATIONAL INITIATIVES

### Overview

Centralized population-based biobanks have been established or proposed in at least nine countries (Tables 1 and 2), each with varying degrees of success. Table 1 describes the content, population size, funding source, oversight mechanism, and status of nine national biobank efforts. Countries in which specimens have been or are now being collected include Iceland, United Kingdom, Sweden, Japan, Latvia, Singapore, and Estonia. The biobank project in the Kingdom of Tonga was cancelled because of overwhelming public opposition. The Icelandic venture was stalled somewhat by a 2003 ruling of its Supreme Court that privacy protections were not adequate.<sup>17</sup> The smallest project is in Latvia, with an expected 60,000 sample size in its pilot phase, and the largest biobank will likely be in Estonia, with an anticipated one million participants (75% of citizenry). These projects are enormous logistical undertakings, with data collection alone taking many years. Not surprisingly, scientific publications on determinants of disease based on these databases are only just beginning to emerge.

In Canada, three population-based biobank initiatives are in various stages of development. The CARTaGENE project in Quebec proposes to enroll 1% of its citizens between 25 and 69 years of age who will be randomly chosen to form a representative sample.<sup>18</sup> Two other population-based genetic biobank projects planned in Canada are a national birth cohort and a longitudinal study on aging.<sup>19</sup>

Two projects followed most closely in the American media are the Iceland and United Kingdom projects. Iceland is unique in both the homogeneity of its citizens and its strong interest in genealogy.<sup>20,21</sup> These two characteristics made it a compelling prospect for the establishment of the world's first population-based biobank. After a vigorous public debate, deCODE Genetics, a Delaware-based biotechnology company, successfully partnered with the Icelandic Parliament (i.e., Althingi) in 1998 to create and operate a centralized database of nonidentifiable health data.<sup>14</sup> Known as the Icelandic Health Database (IHD), the aim of this computerized compilation of the country's medical records is to ultimately improve health and health services. Althingi entered into a 12-year contract with deCODE for exclusive commercial rights to the IHD database (e.g., for developing DNA-based genetic tests, drug targets) as well as a license to assemble genetic and genealogical databases.<sup>14</sup> The United Kingdom conceived its population-based biobank framework with careful attention to stakeholder concerns, and a number of feasibility studies were accordingly commissioned.<sup>22</sup> Launched in April 2002, the United Kingdom project began soliciting volunteers in March 2006 and plans to follow the health of its subjects for up to 30 years.<sup>23</sup> The United Kingdom Biobank is funded by the United Kingdom Department of Health, Medical Research Council, and the Wellcome Trust, and, unlike Iceland's biobank, has

**Table 1**  
Population-based biobanks in nine countries

	Data	Population	Funding	Status
<b>Estonia</b> Estonian Genome Project www.genomics.ee	Links genotype, medical records, and genealogical information	1 million; 75% of country's citizens	\$5 million for pilot project; initially funded by Egeen (Mountain View, CA); supplemented by government since 2004	Collection in progress
<b>Canada</b> CARTaGENE (Quebec) www.rmga.qc.ca/en/index.htm, national birth cohort, and longitudinal study on aging cbac-cccb.ca/epic/internet/incbac-cccb.nsf/en/Home	Genetic, environmental, medical, and genealogical information	CARTaGENE - 60,000 volunteers, ages 25–74 yrs; birth cohort up to age 20 yrs; aging study 50,000 volunteers age ≥40 yrs	Awaiting funding	Reevaluation of procedures based on concerns voiced by experts and public at workshops
<b>Iceland</b> Icelandic Biobank www.decode.com	Personal health, genetic, and genealogical	270,000; homogeneous population	\$212 million	80,000 samples genotyped; privacy protection ruled unconstitutional in 2003; future unclear
<b>Japan</b> Biobank Japan www.src.riken.go.jp/eng/src/project/person.html	Genetic, lifestyle, and medical history	300,000 citizens affected by 47 common diseases	\$180 million, funded by Japanese Education Ministry	100,000 DNA and serum samples collected
<b>Kingdom of Tonga</b>	Personal health, genetic, genealogical	108,000; entire population	Unspecified from national government	Plans cancelled because of public opposition
<b>Latvia</b> Latvian Genome Project www.bmc.biomed.lu.lv.gene	Genetic, medical, and genealogical	60,000 sample for pilot study	Potential from government, private resources, or Latvian Genome Research Foundation	Collection of pilot samples in progress
<b>Singapore</b> Singapore tissue network www.stn.org.sg	Collection of DNA samples of both normal and disease origin	Unspecified	Agency for Science, Technology and Research	Sample collection in progress
<b>Sweden</b> Medical Biobank www.biobanks.se/medicalbiobank.htm	Genetic with self-reported health and lifestyle	70,000 volunteers aged 40–60 yrs in Vasterbotten Co; samples preserved from previous study	Swedish National Healthcare System	Data analysis and repeat sampling at regular intervals from original cohort
<b>United Kingdom</b> UK Biobank www.ukbiobank.ac.uk	Personal health, genetic, physical exam information	Random sample of 500,000 volunteers aged 45–69 yrs	\$110 million from Wellcome Trust, Medical Research Council, Dept. of Health	Recruitment began March 2006

Adapted [and expanded] by permission from Macmillan Publishers Ltd: Maschke K. Navigating an ethical patchwork—human gene banks. *Nat Biotechnol* 2005;23(5):539–545<sup>38</sup> and Austin MA, Harding S, McElroy C. Genebanks: a comparison of eight proposed international genetic databases. *Commun Genet* 2003;6:37–45. Reprinted with permission of S. Karger AG, Medical and Scientific Publishers, Basel, Switzerland.<sup>14</sup>

provided no commercialization rights at present but recognizes that, if new therapies are to arise as a result of research efforts, industry's involvement is inevitable and necessary.<sup>24</sup>

The Icelandic and United Kingdom experiences, and others like them, have fueled much of the debate surrounding ELSI frameworks and other societal concerns. Table 2 summarizes how issues such as consultation with the public, informed consent, confidentiality, and legislative concerns have been addressed to varying degrees.

### Public consultation

Population-based biobanks depend on participation from the public. Acceptance of the research vision and trust in the

safeguards for participants are, therefore, unequivocal necessities. To those ends, Godard et al.<sup>25</sup> propose that public consultation should: 1) begin as early as possible in the process; 2) respond to criticism; 3) be a transparent process to establish legitimacy and gain public trust; and 4) educate the public on the issues involved.

The Icelandic and United Kingdom approaches to public consultation typify one model of early discussion. By consulting and engaging the public early on in the political process, 75% of its population supported the bill on the eve of its passage, and similarly, 75% of the members of the Althingi voted for its passage.<sup>25</sup> The public debate in Iceland has been generally characterized as vigorous and constructive<sup>26</sup> because it included 700 newspaper

**Table 2**  
Ethical, legal, and social issues

	Consent	Confidentiality	Oversight	Legislation
<b>Estonia</b>	Informed consent for initial sample collection; samples destroyed if subject withdraws	Samples identifiable only through coded systems; subjects have access to their information	Unspecified	2000: Act on Human Genome Research
<b>Canada</b>	Informed consent; option for varying degrees of re-consent as part of initial consent	Data anonymized in phase I of project; encoded in phase II	External advisory committee	Unspecified
<b>Iceland</b>	Informed consent for genetic sample collection; presumed consent for previously collected samples; restricted “opt-out” option	Data protection by third-party encryption	Data protection committee, national bioethics committee	1998; Act for health sector database 2000: Act for Biobank
<b>Japan</b>	Informed consent for genetic sample collection; comprehensive consent required to use existing samples for new research	Data protected by separate databases, encoding, and firewalls	Japanese Medical Association	Unspecified
<b>Kingdom of Tonga</b>	Unspecified	Unspecified	Unspecified	Unspecified
<b>Latvia</b>	Informed consent for initial sample collection; sample destroyed if subject changes mind	Samples identifiable only through encoded systems; subjects have access to their information	Central medical ethics committee; state data inspection	2003: Human Genome Research Law
<b>Singapore</b>	Informed consent obtained from previous donors for each new project	Sample coding and anonymized data	Bioethics advisory committee; central steering committee	Unspecified
<b>Sweden</b>	Informed consent obtained from previous donors for each new project	Academic researchers and UmanGenomics AB (Umea, Sweden) have access only to encoded samples	Swedish Medical Research Council	2003: Act on Biobanks in Health Care
<b>United Kingdom</b>	Informed consent to enable medical follow-up and to use genetic samples for analysis; no re-consent once data are anonymized	Anonymous storage; no genotype information shared with donor	Ethics and governance council; scientific committee	2004: Human Tissue Act

Adapted [and expanded] by permission from Macmillan Publishers Ltd: Maschke K. Navigating an ethical patchwork—human gene banks. *Nat Biotechnol* 2005;23(5):539–545<sup>38</sup> and Austin MA, Harding S, McElroy C. Genebanks: a comparison of eight proposed international genetic databases. *Commun Genet* 2003;6:37–45. Reprinted with permission of S. Karger AG, Medical and Scientific Publishers, Basel, Switzerland.<sup>14</sup>

articles, more than 100 radio and television programs, and several town meetings all across Iceland. Some contend, however, that the time allocated for community consultation was insufficient in that the biobank bill was introduced with only 2 months left in the parliamentary session.<sup>27</sup> Faced with severe public criticism in the early stages of its biobank, and with awareness of the Icelandic experience, United Kingdom Biobank administrators responded quickly to public concerns, forming the Ethics and Governance Council (EGC) to oversee the United Kingdom Biobank.<sup>28</sup> An independent body composed of ethicists, lawyers, scientists, and experts on consumer issues, the EGC ensures accountability to the public by acting as an independent guardian of the ethical framework.<sup>29</sup> As a result, the stated interest of the United Kingdom project to promote health throughout society has become more credible.<sup>24</sup>

A different model of consultation adopted in Tonga led to a dramatically different result. The government made plans to create a biobank of samples collected during routine health checkups.<sup>14</sup> A private biotechnology company based in Australia (Autogen Ltd.) was awarded an exclusive licensing agreement with the Tongan government without a public forum for discussion. The Tonga Human Rights and Democracy Movement, the Tongan National Council of Churches, and other advocacy groups expressed vehement criticism, particularly citing the lack of public consultation.<sup>30</sup> Licensing of data generated from national population-based biobanks continues to draw concern,<sup>13</sup> and there is a wide variation in approaches to commercialization among national projects.<sup>14</sup>

A cornerstone of the public consultation process is public education, and Estonia represents a success story. In 1999, a

group of scientists who organized themselves into the Estonian Genome Project presented to government officials and citizens an idea to create a centralized gene and health database.<sup>31</sup> Publicity for these efforts generated enormous public interest in genomics. Tartu University in Estonia experienced a sharp increase in the number of students enrolling in genetic courses and pursuing life science and biotechnology majors. Further, the Estonia Genome Project garnered a high degree of public participation (i.e., more than 75%), and participant awareness was reported to be strong.<sup>31</sup>

## ELSI practices

### *Informed consent*

Obtaining informed consent is one of the grounding ethical requirements for performing research on human subjects in the United States.<sup>6</sup> Emanating from the Nuremberg Code,<sup>32</sup> and the Declaration of Helsinki,<sup>33</sup> informed consent stipulates that a research participant must be provided the opportunity for voluntary informed consent before the study begins. The underlying principle to be satisfied by proper informed consent is respect for autonomy of the individual.<sup>34</sup> Oversight for informed consent, confidentiality, and other human protections in the United States falls under the purview of institutional review boards at individual research institutions and are governed by principles established by the Office of Human Research Protections in the Department of Health and Human Services.<sup>35</sup> Elements of a traditional model of informed consent include an explanation of the proposed research, its purpose, a description of potential risks and benefits to the individuals participating, and a statement that participation is voluntary and can be discontinued at any time (called an opt-out clause).<sup>36</sup> Informed consent also includes an explanation of whom to contact for questions about the research and subjects' rights.<sup>36</sup> In the United States, informed consent is not required for the use of tissue in research if no personally identifiable information is connected to the sample.<sup>6,35</sup>

Administrators of the biobank in Sweden required that consent be obtained for every use of an individual's information, with a waiver possible from an ethics committee only if subsequent proposals involved similar uses to which the donor originally consented.<sup>37</sup> In the United Kingdom (and the United States), if individual data have been made permanently unidentifiable, and if an ethics committee deems that the research is unlikely to harm the individual, re-consent for future projects is not required.<sup>38</sup> In contrast to Sweden and the United Kingdom, the Icelandic legislation challenged the traditional model of informed consent by incorporating what is called presumed consent.<sup>26,39</sup> The Icelandic model restricted the opt-out clause by stating that anyone who decides to withdraw from the study after 6 months cannot have personal health information removed from the database.<sup>39</sup> Objections to this legislation continue, despite the contention that presumed consent has greatly facilitated the gathering of information and maximized the quality of Iceland's database.<sup>40</sup>

The proposed Canadian model for informed consent also differs from the traditional informed consent model by its use of an authorization model.<sup>19</sup> Although informed consent is a requirement at the time of genetic sample collection in this model, participants can also authorize subsequent research, specifying permitted and excluded uses of their genetic material and associated data. Participants can also specify the subsequent degree of re-consenting they wish to maintain. Sheremeta<sup>19</sup> argues that this model strikes a reasonable balance between the potentially restrictive traditional and more general, open-ended presumed consent models. Hansson et al.<sup>34</sup> argue in favor of allowing broad consent (i.e., for multiple purposes without providing as much detail as required by traditional informed consent policies) as well as future consent (i.e., for as-yet unspecified research) regarding participating in a population-based biobank project, based on the assumption of minimal harm and the right to withdraw in the future. For the Icelandic Health Database, Gulcher and Steffansson<sup>26</sup> support the use of broad consent, stipulating that this flexibility is necessary so that the database can function as a "revolutionary method" for studying the interplay between genetics and environment in human disease and health.

The increasing emphasis on intellectual property rights related to discoveries made from donated bodily specimens has somewhat destabilized traditional understandings between human research participants and academic investigators<sup>41</sup> and has implications regarding informed consent. This predicament has prompted scholars to propose the Biotrust Model, in which a nonprofit trusteeship would govern a biobank for the benefit of the public good.<sup>41</sup> Prospective participants would be asked for explicit consent to transfer certain property rights to trustees. Proponents of this model suggest that such a mode of altruistic governance could engender public confidence and greater participation.<sup>41</sup>

Another critical ethical issue regarding informed consent is research involving minors. An example is the proposed Canadian national birth cohort, still under development, which would investigate the complex interaction among genetic, environmental, and psychological determinants of health from birth to age 20 years.<sup>42</sup> Should the academic debate confirm the scientific merit of this type of research, informed consent for children entering the cohort database would be obtained from a parent or guardian acting as the authorized representative of the minor.<sup>19</sup> An added provision is that the research not expose the minor to more than minimal risks because the potential for direct benefit is limited. Further consideration, however, has been suggested because the legality of enrolling minors in research is still unclear in Canada.<sup>19</sup>

### *Confidentiality*

The operational definition of confidentiality is the secure treatment of identifiable information that an individual discloses in a relationship of trust, with the expectation that it will not be divulged to third parties in ways that are inconsistent with the understanding of the original disclosure.<sup>35</sup> Privacy, a similar construct, is defined as control over the extent, timing,

and circumstances of sharing oneself (physically, behaviorally, or intellectually) with others.<sup>35</sup> The terms confidentiality and privacy are often used interchangeably, but we use them in accordance with these specified definitions.

Broadly, there are three methods used for protecting the identity of DNA donors.<sup>38</sup> The first is double-coding, in which anonymous project identification numbers are assigned to participants and associated specimens. The code linking the anonymous number to personally identifying information is placed in a secure location and is accessible by only a few authorized individuals. In the second method, samples may be anonymized: identifiers initially linked to the samples are later removed. Finally, samples may be fully anonymous from the start: all identifying links to the donor are destroyed at the time of collection.

There exists a delicate balance between donor and societal interests in research potential. Anonymity, of course, offers the utmost protection of confidentiality, yet disease markers are more easily discovered when genetic information can be linked to personal medical history. Maintaining identifiable information in a biobank also allows participants to receive information about the project's progress and to learn of new findings that might affect their health. The value of communicating with biobank participants remains a highly controversial subject and must be weighed against possible negative consequences,<sup>13</sup> such as anxiety from learning about their potential risk for cancer and other serious diseases or from a confidentiality breach from an inadvertent communication to someone in the same household. Interestingly, as a result of consultations with the public and experts, CARTaGENE opted to re-contact patients instead of monitoring medical files to gather follow-up information and changed its approach from anonymization of samples to double-coding.<sup>18</sup>

The degree of protection against breaches of confidentiality varies considerably among the nine national biobanks (Table 2). Recent events in Iceland underscore the importance of ensuring protections. In November 2003, the Icelandic Supreme Court ruled that the IHD was unconstitutional on the grounds that it did not adequately protect privacy. Specifically, the plaintiff argued that her deceased father's health records should not be entered into the IHD because it might be possible to infer information about her health from her father's records, even though the data were anonymized.<sup>17</sup> The court ruled that possibility of such identification is increased because the IHD allowed information to be linked with data from other genetic and genealogical databases. Research with currently available genetic samples continues, but it is unclear to what extent IHD is now fully operational.<sup>38</sup> Harsh concerns expressed by the Japan Medical Society regarding inadequacy of privacy and confidentiality safeguards resulted in the temporary suspension of recruitment activities for the population-based biobank in Japan.<sup>43</sup> The Japan Medical Society also criticized the use of municipal employees, untrained in the collection and protection of sensitive information, to solicit volunteers and distribute medical questionnaires. Biobank officials countered that the claims were unfounded because con-

fidentiality will be assured through use of separate databases, encoding, and firewalls.<sup>43</sup> Recruitment resumed and, as of January 2006, more than 130,000 participants have provided written informed consent on an annual basis.<sup>44</sup>

## CURRENT BIOBANKS IN THE UNITED STATES

One of the largest population-based biobanks in the country is the genetic repository of the Marshfield Medical Clinic in Wisconsin, which contains information from 40,000 participants living in northern and central Wisconsin. Some consider this effort to be a small step away from being characterized as a statewide population-based biobank.<sup>45</sup> New Mexico recently proposed legislation to study the feasibility of collecting DNA samples from all residents in the state. This initiative is not designed for genetic research, but rather for forensic purposes such as crime-solving, clearing the wrongly accused and convicted, and identifying unknown persons.<sup>46</sup>

Most population-based biobanks were developed at research institutions, drawing volunteers from neighboring areas, and many are linked to a wealth of well-characterized demographic information. Northwestern University, through its NUGene Project, is currently collecting and storing DNA linked with medical records of volunteer patients from neighboring hospitals and clinics.<sup>47</sup> Hospitals and clinics associated with Duke University participate in a DNA biobank,<sup>48</sup> as do the University of Alabama and Mayo Clinics.<sup>38</sup> The proposed African American Population Biobank, under the auspices of Howard University, plans to enlist 25,000 volunteers over 5 years.<sup>49</sup> The main objective is to develop a population-based genetic epidemiology resource for the study of common, complex diseases among members of the African Diaspora. The United States Department of Veterans Affairs has proposed a comprehensive biobank of veterans and their family members.<sup>50</sup>

A number of other large-scale tissue-banking projects, primarily building on existing epidemiologic studies, have been proposed or are in various stages of development using a variety of bio-samples.<sup>51–53</sup> These include: Physicians Health Study, Nurses Health Study, Multi-Ethnic Cohort, Women's Health Initiative, Cancer Prevention Study II of the American Cancer Society, and the National Health and Nutrition Examination Survey III of the Centers for Disease Control and Prevention.<sup>51</sup> The Framingham Heart Study,<sup>54</sup> a longitudinal study started in 1948, recently included a genetics component involving 9000 volunteers from within the study.<sup>55</sup>

Recent efforts at the National Institutes of Health have broadened the infrastructure to include multiple participating research institutions. The Environmental Genome Project at the National Institute of Environmental Health Sciences is a population-based effort to characterize a number of important genes and to relate variations in genotype with susceptibility to effects of chemical and physical agents.<sup>56</sup> Started in 1998, this project is assessing approximately 200 genes based on analyses of 1000 samples of human DNA. Also located at the National Institutes of Health is the recently launched Genetic Association Information Network, which is a public-private partner-

ship in conjunction with Pfizer Global Research, formed to study the genetic determinants of common diseases.<sup>44,57</sup> Genetic Association Information Network is structured to capitalize on the existence of enormous numbers of well-characterized biobanks housed in hospitals and academic centers across the United States. Via a competitive application process, the project is intended to augment existing, well-characterized case-control studies by providing free genotyping on existing DNA samples from these studies. All results will be made available in the public domain.

## FUTURE CONSIDERATIONS

### An informed public health workforce

Some leaders in the United States advise that a single, centralized national biobank may be an unrealistic goal given the very heterogeneous population of our nation and suggest that a network of linked repositories may be a more pragmatic option.<sup>8</sup> Others speak strongly in favor of a national population-based biobank, arguing that the endeavor could provide a powerful complement to other research approaches and that only through a large number of participants could an adequate sample size exist to appropriately study gene-environment interactions of common disorders.<sup>9</sup> Should biobanking research initiatives through the proposed Genomics and Personalized Medicine Act of 2006<sup>5</sup> and the PREEMIE Act of 2006<sup>11</sup> come to fruition in one form or another, competencies in a number of core areas may be needed for the public health workforce (Table 3). The National Office of Public Health Genomics at the Centers for Disease Control and Prevention, formerly the Office of Genomics and Disease Prevention, has developed professional training programs in public health genomics.<sup>58</sup> The Centers for Disease Control and Prevention has also funded two Centers for Genomics, at the University of Washington and the University of Michigan, to formulate how genomics can be best integrated into public health practice.<sup>59</sup> A primary aim of these centers is to provide tech-

nical assistance and training for the current and future public health workforce in the country. The following sections summarize current strengths and gaps in the public health community with regard to key knowledge areas.

### Protection of Identifiable Health Information

State health departments have a long and noble history of collecting and protecting identifiable health information, such as that required for vital records and reportable illnesses. More recently, the federal Health Insurance Portability and Accountability Act<sup>60</sup> enacted a complex framework for privacy protection that may be relevant to the research needs of population-based biobanks. The Health Insurance Portability and Accountability Act Privacy Rule stipulates specific protected health information that should be properly safeguarded but also allows for the flow of information needed for research that will advance the public health and well-being.<sup>61</sup> The Privacy Rule is pertinent to “covered entities” only and is not applicable when tissue cannot be linked to identifiable information about the individual,<sup>60</sup> as is the case with federal guidelines for institutional review board oversight.<sup>6</sup>

Federal legislation is under consideration in both the House of Representatives and Senate that would protect individuals from discrimination resulting from genetic testing.<sup>62,63</sup> All but 11 states in the United States have some form of state-specific legislation to protect privacy and/or confidentiality of its residents.<sup>64</sup> In our state of Connecticut, the General Statute 38a-999 decrees that insurance providers must protect against unauthorized disclosure of sensitive health information, including results from genetic testing or the mere fact that an individual has taken a test. The United States Indian Health Services group has also developed a privacy protection model for biobanks that they believe will address confidentiality issues specific to the American Indian population.<sup>65</sup> Collectively, these efforts serve as a strong foundation to ensure privacy protection associated with biobanks, but unique issues in privacy protection and confidentiality specifically for biobanks need further study.

### Informed consent models

Groundwork is being laid in the United States to address the issue of informed consent with respect to population-based genetic research, which, will be an extension of long-established protections for human subjects.<sup>66</sup> In response to the increasing need for a federal model that addresses informed consent in population-based genetic research, the National Office of Public Health Genomics at the Centers for Disease Control has proposed new language for informed consent forms.<sup>67</sup> The National Office of Public Health Genomics has also developed a supplemental brochure meant specifically for individuals considering participation in a population-based biobank, which details possible ELSI risks associated with genetics research.<sup>68</sup>

One of the provisions of the proposed United States Senate bill<sup>5</sup> calls for exploring the possibility of combining information from multiple biobanks, referred to as a distributed data-

**Table 3**

Potential knowledge requirements for state public health workforce involved in biobank research initiatives

- General knowledge of public health genomics
- Public consultation processes about biobanks
- Educational efforts focused on health care providers, academicians, and the public
- Informed consent models and opt-out policies
- Federal and state legislation about biobanks
- Civil court challenges to and rulings about existing biobanks
- Health Insurance Portability and Accountability Act and other federal privacy regulations related to public health research
- Current genetics research trends and institutional review board processes
- Establishment and management of tissue banks
- Oversight and regulatory mechanisms of biobanks

base. Recent court challenges indicate that a more uniform use of informed consent allowing for combinations of multiple biobanks may be needed. A lawsuit against the University of Arizona, for example, alleged that blood samples taken for a specific study were used for different research without proper consent.<sup>69</sup> A lawsuit directed at Miami Children's Hospital claimed that blood and tissue samples taken from patients with Canavan disease were used to develop a patented genetic test for the disease without proper informed consent.<sup>70</sup> New York State, however, allows for sharing of information derived from biological samples and genetic databases in its planned electronic health information infrastructure,<sup>71</sup> which may preempt legal challenges.

### Public consultation

There are no specific federal standards in place in the United States on how to conduct public consultation regarding the establishment of a government-sponsored population-based biobank. State public health departments may therefore need to develop proficiency in such matters as public health genomics,<sup>58,16</sup> DNA-banking practices,<sup>51</sup> and general issues in tissue-banking for research.<sup>51–53</sup> Another potential need involves management responsibility or governmental oversight of tissue biobanks. Establishing a biobank would require the capacity to collect, store, and catalogue specimens and to link information with other databases and, possibly, distribute specimens.<sup>53</sup> Diagnostic and screening laboratories operated by most state health departments for population-based programs such as newborn screening already play an administrative role in specimen-based activity. There is also precedent for enforcement of a multitude of health regulations, which could be modified for the monitoring of safe and lasting storage of tissue.

In 1999, the government of Canada established the Canadian Biotechnology Advisory Committee, a body of external experts to provide the public comprehensive advice on current policy issues associated with the ethical, social, regulatory, economic, scientific, environmental, and health aspects of biotechnology.<sup>19</sup> Canadian Biotechnology Advisory Committee has prepared a comprehensive scholarly report about ELSI issues surrounding the establishment of a national biobank in Canada and arranges workshops and other opportunities for citizens and external experts to voice their views.

### Planning for the future

As with a growing number of states, the Department of Public Health in our state of Connecticut developed a Genomics Action Plan<sup>72</sup> to serve as a guide for the translation of research findings from the Human Genome Project into public health practice. Envisioning and documenting the possible role of a state health department in the development of population-based biobanks for genomics research was an outgrowth of this plan. Most of the Connecticut plan addresses translational issues in concert with recommendations by the National Office of Public Health Genomics<sup>58</sup> and includes anticipated activities such as: 1) educating the citizenry, health care community, and public health workforce; 2) exploring integration of health

information data; 3) developing germane health policies, legislation, programs, and services; and, 4) identifying future infrastructural needs. The first goal to be realized was the establishment of an internal working group, known as the Virtual Office of Genomics.<sup>73</sup> The Virtual Office of Genomics is a compendium of staff members from several sectors among the Connecticut Department of Public Health who meet on a regular basis to exchange information, conduct training workshops, and produce discussion papers with the goal of eventually constructing a fully operational Office of Genomics. Mindful that public health genomics is an ever-evolving enterprise,<sup>8</sup> the Virtual Office of Genomics formed an External Genomics Advisory Panel, which is a counseling body consisting of health care providers, clinical laboratory directors, genetic counselors, lawyers, medical ethicists, and academicians.

Appreciation of the research utility of a national population-based biobank or an integrated system of repositories is gaining momentum in the United States. Evidence-based guidelines derived from these ventures could dramatically change the nature of personal and public health medicine. Proper planning may avoid public disruption and thereby help to fulfill the promise of the Human Genome Project.

### ACKNOWLEDGMENTS

We are very grateful for support and guidance from the Department of Public Health, State of Connecticut. In particular, we thank Meg Hooper, MPA, Marianne Horn, JD, and Mary Lou Fleissner, DrPH, for significant contributions. We also extend gratitude to the members of the Virtual Office of Genomics and the External Genomics Advisory Panel for excellent commentary.

### References

1. Lander ES, Linton LM, Birren B, and the International Human Genome Mapping Consortium. Initial sequencing and analysis of the human genome. *Nature* 2001; 409:860–921.
2. Venter JC, Adams MD, Myers EW, and Celera Genomics. The sequence of the human genome. *Science* 2001;291:1304–1351.
3. National Human Genome Research Institute. All about the Human Genome Project. Available at: <http://www.genome.gov/10001772>. Accessed September 30, 2006.
4. Austin CP. The impact of the completed human genome sequence on the development of novel therapeutics for human disease. *Annu Rev Med* 2004;55:1–13.
5. United States Congress, S. 3822. Available at: [http://olpa.od.nih.gov/tracking/109/senate\\_bills/session2/s-3822.asp](http://olpa.od.nih.gov/tracking/109/senate_bills/session2/s-3822.asp). Accessed September 24, 2006.
6. Kapp MB. Ethical and legal issues in research involving human subjects: do you want a piece of me? *J Clin Pathol* 2006;59:335–339.
7. Kaiser J. NIH ponders massive biobank of Americans. *Science* 2004;304:1425.
8. Gwinn M, Khoury MJ. Genomics and public health in the United States: signposts on the translation highway. *Comm Genet* 2006;9:21–26.
9. Collins FS. The case for a prospective cohort study of genes and environment. *Nature* 2004;429:475–477.
10. Jin J. An evaluation of the ethical, legal and social implications program of the U.S. Human Genome Project. *Princeton J Bioeth* 2000;3:35–50.
11. Prematurity Research Expansion and Education for Mothers who deliver Early Act (known as "PREEMIE Act"). Available at: <http://www.govtrack.us/congress/billtext.xpd?bill=s109-707>. Accessed February 5, 2007.
12. Maschke KJ, Murray TH. Ethical issues in tissue banking for research: the prospects and pitfalls of setting international standards. *Theor Med Bioeth* 2004;25:143–155.
13. Corrigan O. Biobanks: can they overcome controversy and deliver on their promise to unravel the origins of common diseases? *Med Educ* 2006;40:500–502.
14. Austin MA, Harding S, McElroy C. Genebanks: a comparison of eight proposed international genetic gatabases. *Comm Genet* 2003;6:37–45.
15. Burke W, Khoury MJ, Stewart A, Zimmern RL. Bellagio Group. The path from genome-based research to population health: development of an international public health genomics network. *Genet Med* 2006;8:451–458.



16. Khoury MJ, Burke W, Thomson EJ, editors. *Genetics and public health in the 21st century*. Oxford: University Press, 2000.
17. Abbott A. Icelandic database shelved as court judges privacy in peril. *Nature* 2004; 429:118.
18. CARTaGENE. Available at: <http://www.cartagene.qc.ca>. Accessed December 21, 2006.
19. Sheremeta L. Population biobanking in Canada: ethical, legal, and social issues. Canadian Biotechnology Advisory Committee. Available at: <http://cbac-cccb.ca/epic/internet/incbac-cccb.nsf/en/ah00482e.html>. Accessed July 12, 2005.
20. Bjarnason O, Bjarnason V, Edwards JH, Fridriksson S, Magnusson M, Mourant AE, et al. The blood group of Icelanders. *Ann Hum Genet* 1973;36:425–458.
21. Gulcher JR, Kong A, Stefansson K. The role of linkage studies for common diseases. *Curr Opin Genet Dev* 2001;11:264–267.
22. UK Biobank. Available at: <http://www.ukbiobank.ac.uk>. Accessed July 12, 2005.
23. Kaiser JUK. Biobank taking deposits. *Science* 2006;311:1535.
24. Government response to the work of the Medical Research Council by the House of Commons Science and Technology select committee (HC 132). In: Sheremeta L. Population biobanking in Canada: ethical, legal, and social issues. Canadian Biotechnology Advisory Committee. Available at: <http://cbac-cccb.ca/epic/internet/incbac-cccb.nsf/en/ah00482e.html>. Accessed December 20, 2006.
25. Godard B, Marshall J, Laberge C, Knoppers BM. Strategies for consulting with the community: the cases of four large-scale genetic databases. *Sci Eng Ethics* 2004;10: 457–477.
26. Gulcher JR, Stefansson K. The Icelandic Healthcare Database and informed consent. *N Engl J Med* 2000;342:1827–1830.
27. Winickoff DE. The Icelandic Healthcare Database. *N Engl J Med* 2000;343:1734–1735.
28. Your questions answered. Available at: <http://www.ukbiobank.ac.uk/about/faqs.php>. Accessed July 15, 2005.
29. Ethics and governance council formed to oversee UK Biobank. Available at: [http://www.wellcome.ac.uk/doc\\_WTX022802.html](http://www.wellcome.ac.uk/doc_WTX022802.html). Accessed July 15, 2005.
30. Barker JH. Common-pool resources and population genomics in Iceland, Estonia, and Tonga. *Med Health Care Philos* 2003;6:133–144.
31. Frank L. Biotechnology in the Baltic. *Nat Biotechnol* 2001;19:513–515.
32. United States counsel for war crimes trials of war criminals before the Nuremberg Military Tribunals under control council law. Washington, DC: US Government Printing Office. Available at: <http://ohsr.od.nih.gov/guidelines/nuremberg.html>. Accessed August 7, 2005.
33. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects, 1964. Available at: <http://www.wma.net/epolicy/b3.htm>. Accessed August 7, 2005.
34. Hansson MG, Dillner J, Bartram CR, Carlson JA, Helgesson G. Should donors be allowed to give broad consent to future biobank research? *Lancet Oncol* 2006;7:266–269.
35. IRB Guidebook, Office for Human Research Protections, US Department of Health and Human Services. Available at: [http://www.hhs.gov/ohrp/irb/irb\\_guidebook.htm](http://www.hhs.gov/ohrp/irb/irb_guidebook.htm). Accessed April 19, 2006.
36. General Requirement for Informed Consent. Office of Human Research Protection, Title 46.116, United States Department of Health and Human Services. Available at: <http://www.hhs.gov/humansubjects/guidance/45cfr46.htm#46.116>. Accessed August 16, 2005.
37. Kaye J, Martin P. Safeguards for research using large scale DNA collections. *BMJ* 2000;321:1146.
38. Maschke KJ. Navigating an ethical patchwork: human gene banks. *Nat Biotechnol* 2005;23:539–545.
39. Althingi. Act on a Health Sector Database No. 139/1998. Available at: <http://www.informatica-juridica.com/anexos/anexo610.asp>. Accessed August 10, 2005.
40. McInnis M. The assent of a nation. Available at: <http://www.mannvernd.is/greinar/assent10.htm>. Accessed August 10, 2005.
41. Winickoff DE, Neumann LB. Towards a social contract for genomics: property and the public in the “Biitrust Model.” *Genom Soc Policy* 2005;1:8–21.
42. Canadian Institutes of Health Research. The Canadian Lifelong Health Initiative: a CIHR cross-cutting initiative. Available at: <http://www.cihr-irsc.gc.ca/e/18542.html>. Accessed August 12, 2005.
43. Normile D. Japan guidelines under fire after protest halts study. *Science* 2003;301: 1039.
44. Tilstone C. Further plans announced for national biobanks. *Lancet Oncol* 2006;7: 195–196.
45. Marshfield clinic launches large-scale population-based genetic research project. Available at: <http://www2.marshfieldclinic.org/cattails/02/novdec/home.asp>. Accessed March 21, 2006.
46. New Mexico State Legislature, HJM 79: A joint memorandum requesting the New Mexico legislative council to direct appropriate interim committee to study feasibility and desirability of requiring every New Mexican to provide a DNA sample to the state. Available at: <http://www.legis.state.nm.us/Sessions/06%20Regular/memorials/house/HJM079.html>. Accessed March 27, 2006.
47. NUGene. Northwest University, Chicago, IL. Available at: <http://www.nugene.org/about.htm>. Accessed March 24, 2006.
48. Duke Center for Human Genetics DNA bank and tissue repository. Duke University, Durham, NC. Available at: <http://www.chg.duke.edu/research/dnabank.html>. Accessed March 24, 2006.
49. Kaiser J. African-American population biobank proposed. *Science* 2003;300:1485.
50. Couzin J. Gene bank proposal draws support—and a competitor. *Science* 2005;309: 684–685.
51. Steinberg K, Beck J, Nickerson D, Garcia-Closas M, Gallagher M, Caggana M, et al. DNA banking for epidemiologic studies: a review of current practices. *Epidemiology* 2002;13:246–254.
52. Goodman GE, Thornquist MD, Edelstein C, Omenn GS. Biorepositories: let’s not lose what we have so carefully gathered! *Cancer Epidemiol Biomarkers Prev* 2006;15: 599–601.
53. Goodman MT, Hernandez BY, Hewitt S, Lynch CF, Cote TR, Frierson HF, et al. Tissues from population-based cancer registries: a novel approach to increasing research potential. *Hum Pathol* 2005;36:812–820.
54. National Heart, Lung, and Blood Institute. Framingham Heart Study Cohort. Available at: <http://www.nhlbi.nih.gov/resources/deca/descriptions/framcohort.htm>. Accessed August 19, 2005.
55. Kaiser J. NIH goes after whole genome in search of disease genes. *Science* 2006;311: 933.
56. Guengerich FP. The Environmental Genome Project: functional analysis of polymorphisms. *Environ Health Perspect* 1998;106:365–368.
57. Foundation for the National Institutes of Health. Genetic Association Network Launched. Available at: <http://www.genome.gov/17516722>. Accessed September 20, 2006.
58. National Office of Public Health Genomics, Centers for Disease Control and Prevention. Available at: <http://www.cdc.gov/genomics/>. Accessed October 12, 2006.
59. Centers for Genomics, National Office of Public Health Genomics, Centers for Disease Control and Prevention. Available at: <http://www.cdc.gov/genomics/phpractice.htm>. Accessed October 12, 2006.
60. Office for Civil Rights, United States Department of Health and Human Services. Medical privacy: national standards to protect the privacy of personal health information. Available at: <http://www.hhs.gov/ocr/hipaa/finalreg.html>. Accessed August 16, 2005.
61. Office of Civil Rights, United States Department of Health and Human Services. Privacy brief. Summary of the HIPAA Privacy Rule. Available at: <http://www.hhs.gov/ocr/hipaa/>. Accessed August 16, 2005.
62. United States Congress, H.R. 1227: Genetic Information Nondiscrimination Act of 2005. Available at: <http://www.govtrack.us/congress/bill.xpd?bill=h109-1227>. Accessed April 11, 2006.
63. United States Congress, S. 306: Genetic Information Nondiscrimination Act of 2005. Available at: <http://www.govtrack.us/congress/bill.xpd?bill=s109-306>. Accessed April 11, 2006.
64. Council for Responsible Genetics. Genetics and the law database. Available at: <http://www.genelaw.info/pages/legislationresults.asp?area=3>. Accessed March 27, 2006.
65. Ueda A. Moving from concerns to considerations in population-based genetic research. The 131st annual meeting of the American Public Health Association. Available at: [http://www.apha.confex.com/apha/131am/techprogram/paper\\_64777.htm](http://www.apha.confex.com/apha/131am/techprogram/paper_64777.htm). Accessed March 21, 2006.
66. Office of Human Research Protection: Code of Federal Regulations. Available at: <http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm>. Accessed April 19, 2006.
67. Beskow LAM, Burke W, Merz J, Barr PA, Terry S, Pechaszadeh VB, et al. Informed consent for population-based research involving genetics. *JAMA* 2001;286:2315–2321.
68. Informed consent: taking part in population-based genetic research. National Office of Public Health Genomics, Centers for Disease Control and Prevention. Available at: <http://www.cdc.gov/genomics/info/reports/policy/brochure.htm>. Accessed March 21, 2006.
69. Dalton R. When two tribes go to war. *Nature* 2004;430:500–502.
70. U.S. District Court for Southern Florida. Greenberg *versus* Miami Children’s Hospital. Available at: <http://www.genelaw.info/pages/casesresults.asp?area=8>. Accessed March 27, 2006.
71. New York State Assembly, S6859: Health E-Links New York Act. Available at: <http://www.public.leginfo.state.ny.us/menuegetf.cgi>. Accessed March 27, 2006.
72. Connecticut Department of Public Health. Connecticut Genomics Action Plan, 2005. Available at: <http://www.dph.state.ct.us/Genomics/publications.htm#GAP>. Accessed October 10, 2006.
73. Virtual Office of Genomics, Department of Public Health, State of Connecticut. Available at: <http://www.dph.state.ct.us/Genomics/index.htm>. Accessed October 10, 2006.