Thomas Weaver shares his time between the MRC Geneservice, Babraham Bioincubator, Babraham, Cambridge, CB2 4AT, England and the MRC Rosalind Franklin Centre for Genomics Research, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SB, England.

Johannes Maurer is at the RZPD Deutsches Ressourcenzentrum für Genomforschung GmbH, Heubnerweg 6, D-14059, Berlin, Germany.

Yoshihide Hayashizaki is a member of the Laboratory for Genome Exploration Research Group, RIKEN Genomic Science Centre, RIKEN Yokohama Institute 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan and the DNA Book Team, Preventure Program, Office of Technology Transfer, Japan Science and Technology Agency (JST), 1-1-40, Suehiro-cho, Tsurumi-ku, Yokohama-shi, Kanagawa, 230-0045 Japan.

doi:10.1038/nrg1474

1. Cyranoski, D. Research materials: share and share alike? *Nature* **420**, 602–604 (2002).
2. Cech, T. R. *Sharing Publication-Related Data and Materials: Responsibilities of Authorship in the Life Sciences* (The National Academies Press, Washington, 2003).
3. Cozzarelli, N. R. UPSIDE: uniform principle for sharing integral data and materials expeditiously. *Proc. Natl Acad. Sci. USA* **101**, 3721–3722 (2004).
4. Collins, F. S., Green, E. D., Guttmacher, A. E. & Guyer, M. S. A blueprint for the genomic era. *Nature* **422**, 835–847 (2003).
5. Ishkanian A. S. *et al.* A tiling resolution DNA microarray with complete coverage of the human genome. *Nature Genet.* **36**, 299–303 (2004).
6. Baross A. *et al.* Systematic recovery and analysis of full-ORF Human cDNA clones. *Genome Res.* (in the press).
7. Lennon, G., Auffray, C., Polymeropoulos, M. & Soares, M. B. The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics* **33**, 151–152 (1996).
8. Cyranoski, D. Super-enzyme patents get their day in court. *Nature* **419**, 767 (2002).
9. Wellcome Trust *et al.* Summary of principles agreed at the first international strategy meeting on human genome sequencing. *Human genome project information*, [online] <http://www.ornl.gov/sci/techresources/Human_Genome/research/bermuda.shtml#1> (Bermuda, 25–28 Feb, 1996).
10. Food Standards Agency (UK) *et al.* Joint Code of Practice for Research [online] <http://www.foodstandards.gov.uk/multimedia/pdfs/QACOPRes.PDF> (2004).
11. De Meyer, B., Lurin, C., Small, I. & Hilson, P. MIAO, the minimum information about an ORF. *Genome Res.* (in the press).

Acknowledgements
We would like to thank M. Kanamori, Y. Shigemoto and M. Nishikawa for their assistance in preparing this perspective, and R. Drysdale for critically reading the manuscript.

Competing interests statement
The authors declare no competing financial interests.

Online links

FURTHER INFORMATION
DNA Data Bank of Japan: http://www.ddbj.nig.ac.jp/
Integrated Genomics: http://www.integratedgenomics.com/
International Organization for Standardization: http://www.iso.org/iso/en/ISOOnline.frontpage
Jackson Laboratories Molecular Probes and Clones Query Form: http://www.informatics.jax.org/searches/probe_form.shtml
Mammalian Gene Collection: http://mgc.nci.nih.gov
Access to this interactive links box is free online.

SCIENCE AND SOCIETY

The social and ethical issues of post-genomic human biobanks

Anne Cambon-Thomsen

Abstract | Biobanking — the organized collection of biological samples and associated data — ranges in scope from small collections of samples in academic or hospital settings to large-scale national repositories. Biobanks raise many ethical concerns, to which authorities are responding by introducing specific regulations. Genomics research, which thrives on the sharing of samples and information, is affected by two prominent ethical questions: do ethical principles prevent or promote the sharing of stored biological resources? How does the advent of large-scale biobanking alter the way in which ethical issues are addressed?

The term biobank refers to organized collections of biological samples and the data associated with them. Biobanks come in many different forms, according to the type of samples that are stored and the domain in which they are collected[1–4] (BOX 1; TABLES 1,2; online supplementary information S1,S2,S3 (tables)). For example, samples might derive from a clinical setting, from research projects or from the judiciary domain. The data might relate to a given individual (clinical data), to a family (genealogy information or ethnic origin) or to a group (geographical location of a population, or its language). Such data might be collected at the time of sampling or can be added to the database at a later date. The collected samples do not need to be physical, as in the case of mutation databases, which contain only the data plus information on the sample origin and its geographical provenance.

Biobanking is not a new concept, but its prevalence is increasing in all countries. This is for three main reasons: the growth of biomedical research has increased the number of people who might benefit from biobanks; the growing size of the collections increases their scientific value; and the range of applications of databanks has grown, especially in genomics and in population genomics[1,4–6]. As a consequence, and in line with present trends in science and society relations, the recognized status of biobanks is changing. From being clinical or academic research tools that were largely ignored by the general public, they have become a subject of societal debate. They have acquired the status of national resources; the term 'biological resource centre', as used in the OECD (Organisation for Economic Co-operation and Development), is representative of this trend[6]. Specific funds are allocated to biobanking in clinical or research settings, where it had gone unnoticed for years. In addition, the 'biobank manager' is emerging as a new profession, and new scientific societies such as the ISBER (International Society for Biological and Environmental Repositories) are being created. Some politicians concerned with biotechnology development regard biobanking as an important economic development, and large private investments are being introduced, sometimes in conjunction with public or private non-profit funding[5].

The situation in the United States is different from that in Europe, where there are fewer commercial biobanks and a less liberal attitude towards the use of tissues and cells of human origin[3,7]. Nevertheless, biobanking in all countries has raised similar ethical issues[8,9]. The first is the tension that exists between the rights of individuals or groups and the routes towards research progress. The second concerns the need to provide adequate information to individuals before giving consent to deposit their samples as well as raising awareness of the unforeseen research studies that could be carried out using the samples or their associated data. The third is the difficulty of reconciling the non-commercial use of human body parts with the growing role of commercial biobanks. Finally, a debate continues over how best to ensure the optimal and transparent use of biobanks while defining the rights of priority of researchers and companies over samples and data. Regulatory bodies have generated an avalanche of guidelines and regulations related to biobanks[3,10–16]; entire conferences are dedicated to this field and many reports and opinions have been published. Therefore it is interesting to analyse the ethical developments that have been

Box 1 | **The variety of biobanks**

Biobanks vary in size, degree of access, the status of the institution(s) in charge of their constitution and/or management (public, private companies, private non-for-profit organizations), the range of possible uses and the extent to which samples can be traced back to their donor[1,3,7,41]. TABLES 1,2 list some examples of large biobanks and their main features (see also online supplementary information S1,S2,S3 (tables) for more examples).

Biobanks also vary according to the scientific sector in which the samples were collected:

Medical and academic research. Medical genetic studies of disease (especially if rare) have motivated collections that usually consist of small case- or family-based repositories. Population-based collections have long existed in the fields of genetic anthropology and history of world populations, although these are also usually small and have been used for academic research. Some large epidemiological studies have also involved the collection of a large number of samples.

Clinical studies. There are large collections of samples in hospitals, where they are primarily used for informing diagnosis, and for clinical or therapeutic follow up. Pathology departments, in particular, have collected huge numbers of tissue sections over the years. Transplantations using cells, tissues or organs from unrelated donors have also led to the development of tissue and cell banks for therapeutic use. The treatment of infertility has motivated collections of sperm and eggs. The development of this clinical sector results in collections of frozen embryos that in some cases are no longer required for their intended use. Their possible destruction or their alternative research use have been the matter of extensive discussions in many countries.

Biotechnology domain. Collections of reference cell lines that are well characterized for several relevant characteristics (such as cancer cell lines or antibody-producing cell lines), and stem cell lines of various origin, are mainly used in biotechnology research and development.

Judiciary domain. This sector hosts huge collections of different sources of biological material, data and DNA fingerprints, which have very restricted uses.

raised by biobanking, particularly as they apply to large-scale biobanks, and how biobank managers are responding to the explosion of regulations. Societal issues have such a large role in the construction of projects that they themselves become a driving force in this field.

In this article I describe the aforementioned driving forces and how they affect the main stakeholders. I limit my discussion to biobanks that contain human samples and data used for genetic and genomic studies. I also only consider repositories that offer some degree of accessibility, that are available for various research purposes and that permit the exchange of materials among users. Biobanks that are used only in diagnostic, therapeutic, forensic or judiciary settings are not considered.

Post-genomic and large-scale biobanks

Several factors have contributed to the recent shift from small, local biological repositories to large population-based collections[5,17,18]. These include technical and computational advances (such as high-throughput genomics techniques), new systematic approaches (including large-scale SNP genotyping to characterize genomic variation), and the growing level of exchange of biological material and information among researchers.

It is generally assumed that large samples will help to explore the genetic basis of common multifactorial diseases and the contribution of gene–environment interactions to disease[19]. The trend is therefore towards forming large population collections on the basis of the models of biological databases that have been set up in Iceland, Estonia, Latvia, Sweden and the United Kingdom, as well as other countries[18,19,20] (see TABLES 1,2 and online supplementary information S2,S3 (tables)).

Although large population-based collections represent only one type of biobank[1,3], these have caught the attention of the public and of the media. This has been the case particularly for the Icelandic DeCode initiative, which was the first of its kind (see TABLE 1, online supplementary information S2 (table) and Online links box).

Some existing biobanks have been created in response to a scientific need. For example, gathering many samples from individuals with a rare disease or from families with multiple cases of the same disease might be necessary to reach an appropriate sample size and to permit adequate statistical analysis. However, the usefulness of new large biobanks has been questioned by ethicists[5,17–21] and by scientists[22]. The accuracy and use (in terms of statistical power) of large, costly biobanks compared with smaller, more targeted ones is questioned. I argue that large population biobanks are indeed useful tools not only as a repository of genomic knowledge but also as a means of measuring non-genetic environmental factors. As such, they give epidemiologists and geneticists a new tool to explore complex gene–gene and gene–environment interactions at the population level. The trend towards larger biobanks also raises concerns about how to ensure the ethical use of human samples and the associated information. Most researchers agree on the great potential of biobanks, but realize that the principal obstacle to their success depends on their acceptance by the public. Even lawmakers are hesitant to proceed as long as the ethical environment and public acceptance remain unclear. On the one hand, large population-based biobanks raise new ethical challenges[4,18], and on the other hand, societal issues influence the way projects are constructed and presented, as is analysed in the following sections.

Ethical and social issues

The general ethical principles of autonomy, beneficence/non-maleficence and justice are usually translated into actions through informed consent, protection of confidentiality and private life, and non-discrimination measures. Regulation is expressed by international declarations or conventions, professional guidelines or by national legislation[13,16,23]. 'Large-scale biobanking' or 'biobank-omics'[4,8] has had to adapt the ethical frameworks that were developed for smaller biobanks, while retaining the ethical principles themselves. Given that a large-scale biobank might include information on all or a good proportion of a population, new societal dimensions also need to be considered.

What are the relevant ethical and societal issues? The amount of information that can be extracted from a sample has grown exponentially: are individuals aware of such progress? Genetic testing can use samples that were not collected initially for genetic studies. This raises the issue of secondary non-planned uses. Which criteria should be considered in such decisions, who should be consulted and who should decide? Population biobanks that were previously used in epidemiology or in anthropology in an academic context are now of utmost interest to industry for pharmacogenetic applications[4]. Similarly, collections of tissue biopsies that were of no other use than for individual diagnosis or clinical follow-up are now the source of new information for gene expression studies[9]. How should relationship between public and private

Table 1 | **Human post-genomic biobanks in Europe**

Biobank properties	Iceland DeCode Biobank	Estonian Genome Project*	UK Biobank	GenomEUtwin* (Finland)
Institution	DeCode Genetics (private company)	Estonian Genome Project Foundation (non-profit)	UK Biobank Limited (charitable company limited under guarantee, under a joint venture agreement)	National Public Health Institute and University of Helsinki
Type	National. Population based. Private license	National. Population based; three-quarters of the country's population	National. Population based; aged between 45 and 69 yrs old	International. Population-based twin cohorts in six EU countries and Australia
Aims	Identify genetic component of common diseases	Genetic research; public health surveillance involving genetic components to be developed	Mainly gene–environment interactions in complex diseases	Influence of genetic and non-genetic factors on five complex traits: obesity, stature, coronary heart diseases, stroke and longevity
Notable features	The first of the large population-based national studies to obtain genealogy, health data and DNA samples. Most studied and debated	The largest project envisaged so far	The first project of its kind in an EU country with a population size over 50 million. An experience in population consultation	A large multinational twin study building on existing registries
Total number of individuals	270,000	1,065 million	500,000 aged between 45 and 69 yrs old	80,000
Budget	NS; US$800,000 granted annually to Icelandic government by DeCode‡	NS	GB£61 million (US$109 million)	NS
Year started (stopped or planned completion)	1998 (Act on biobanks in May 2000)	1999 (In 2000, the Human Genes Research Act was passed; project is planned over 10 yrs)	Plans started 1999; pilot study planned 2004; recruitment from 2005 (planned for 10 yrs)	2002 (>4 yrs)
Follow-up of individuals (prospective)	Yes, through health database	Yes	Partial, through medical records; re-contact for a sub-group	Yes
Current progress	>80,000 samples obtained	Pilot project of 10,000 in 3 counties done; more recruitment undergone	Pilot study planned for 2004	Standardization of methods. First phase: genotype of 10,000 samples with genome-wide markers and specific tag SNPs
Further description	Three linked databases: genealogical, health and DNA	Includes public health development; public foundation and private companies (Egeen Inc. agreement).	No individual results provided to subjects/doctors; open to scientific community (protocols might vary). No DNA sent abroad. P3G membership envisaged	The data and sample base (biobank) are only part of the study. The project mostly uses existing samples and acquires new ones for part of the study
Commercial aspects	Yes, major; 12 yrs exclusive rights	Plans are to market products and access to information through Egeen to pharmaceutical companies	Yes, but not precisely described. No exclusive rights will be granted	None
Ethical governance, informed consent and confidentiality	Opt out for health and genealogical data; informed specific consent for biological samples; complex structure for protection of data; supervision by national ethics committee	*Ad hoc* specific ethics committee. Volunteer (opt in) participation, with a restricted opt out option, coded and encrypted data; genetic results to be provided to participants and/or their doctor	Ethics and governance protocol publicly available in 2003, under revision in 2004. Ethics and Governance Council appointed (2004). Extensive information: written consent for many uses; re-consent if other uses; possibility to withdraw; coded data (link accessible to a few people); different kinds of data in different unconnected machines	Major component. Protocol of governance under development; education in ethics for scientists included in protocol; new consent required in case of existing samples. Coded information; no link with any identifying information in the database
Public debate organized	Not initially, which has been a problem. Yes (secondarily)	Yes. Press coverage	Yes. Public consultation in 2002, industry consultation in 2003; documents available on the internet (also ethics-governance)	No
Group opposition	Yes (>20,000 people opted out). Strong debate around consent, true anonymity, children being involved, exclusive rights and commercial issues	Not strong. Discussions about priorities in public health	Some: questions on scientific validity; on security of data; various points underlined at public consultation	No; existing accepted cohorts are used

*Belongs to P3G (Public Population Project in Genomics, http://www.p3gconsortium.org) — a consortium in which four human genome research database initiatives collaborate for the creation of a public accessible database (CARTaGENE, Estonian Genome Project, UK biobank and GenomeEUtwin). In 2003, the possibility of collaboration between COGENE, a project funded by the European Union (EU) to coordinate National Genome Initiatives at EU level and the P3G consortium was envisaged. ‡See ref. 51. NS, not specified; yrs, years. A more comprehensive version of this table is available online (supplementary information S2 (table)).

sectors, which traditionally only involved academics or clinicians, be organized[7]? Bioinformatics allows large-scale data storage and management in a refined way; is this a danger for personal data protection or a help in protecting such data[24]?

The methodological, technical and conceptual developments in genetics and genomics therefore generate concerns, although several of them are not restricted to genetics[4,17]. These can be addressed by improving the means of communication and debate among scientists, regulatory bodies and the public, and by increasing societal control over the economical exploitation of biobanks, their content and their use. How public concerns about ethical and societal issues can be integrated with those of the scientific community is discussed for each issue in the following sections.

Informed consent. Protecting autonomy through a consent procedure is extremely important as it shows respect for the individual. Although the principle of informed consent[14,25] is recognized by all, its translation into practice encounters difficulties in the case of large-scale biobanks, long-term use of samples or data, or numerous exchanges[11,13,26]. Although it does not in itself protect an individual, informed consent allows individuals to exercise their fundamental right to decide whether and how their body, its parts and the associated data will be used in research[27]. Some policies for the use of biobanks, such as a presumed consent to all possible future uses of samples and/or data, limit this right, especially when applied to already stored samples. Given the importance of individual autonomy, these limits cannot be declared purely for practical reasons. Instead, they need to rely on other principles — these include solidarity or the recognition that individuals might have to relinquish some control over the use of their own samples and data if it is for the common good.

New ethical frameworks have been created on the basis of such proposals[9,17,28,29]. Consent increasingly includes several alternative modalities of involvement for the sample donor. True 'informed consent' is strictly defined as specific consent given for well-defined uses; the donor is given transparent information, the possibility of dialogue with a professional, and time to think about the implications before a decision is taken[11,20,26,27]. Other forms of consent or modalities of involvement have been used including enlarged consent, consent with several options for research use, presumed consent and blanket consent (although these last two can hardly be considered consent at all)[13,19,30].

However, informed consent is far from being a magical solution to ethical preoccupations, as the level of perceived information is variable[20] and the limitations of this process are well known, as described next[11,16].

Depending on how samples are stored (see BOX 2) individuals might be able to find out how their samples have been or will be used, with the possibility of selectively or definitively opting out of further use of samples or data. However, the more sophisticated the encoding or encryption system, and the more exchanges of data and samples occur, the more difficult it becomes to destroy samples and generated data. In addition, it is sometimes necessary to keep data for further follow-up, and to use the original sample as a control. In such cases, it is important to make clear to the person at the time of consent that withdrawal at a given date will ensure that no new results will be generated and that the remaining sample is destroyed, even though this does not guarantee the destruction of existing data. This is a matter of controversy. I believe that when scientific data have been produced with the consent of a person, this person should not have the right to ask for their destruction, but only for their anonymization. Such a view is shared by many scientists, but not by all ethics committees.

It has also been reported that an individual who is frequently asked for consent to carry out detailed uses develops 'consent fatigue'[21], and that some individuals do not wish to receive extensive information[31]. Furthermore, although the consent process and the whole ethical approval procedure protects both the researchers and the research subjects, the latter sometimes perceive the consent to be a contract devised to protect the professionals[32].

Issues that relate to group consent have arisen particularly in population genetics. More than 10 years ago, the debate[33,34] surrounding the proposed Human Genome Diversity Project (see Online links box) led to an ethical framework[35] that took into account not only individual rights but also the cultural differences among communities or groups, as well as their identity, autonomy, opinion and rights. As a result, the group or community consent concept was put forward[36]. This now classical, ethical framework for population genetics has recently been deepened[30] and formalized[37], and has been adapted to large population-genetic studies involving large-scale biobanking[21,28]. Collective consent is used in small communities but is not adapted to large heterogeneous groups. Instead, a collective debate is held before a project begins and before individual consents are pursued: a person can then take a decision that remains at the individual level, but takes into account the dimensions underlined through this collective debate. Some bioethics committees now recommend, rather than merely advise, this kind of debate[11,16,38]. Organizing such debates or population consultations is challenging[11] and can influence the research protocol itself by taking into account the views of individuals or groups that are consulted. For example, the involvement of new considerations such as a fair sharing of data for a common good and a new form of solidarity have been put forward following debates on large biobanks[11,17,28].

Secondary or extended uses: under which conditions? Secondary uses for stored human samples are nearly always possible even though they are usually not foreseeable at the time of sampling[2,11,13,20,30,38,39]. The main ethical issues relate to the level of completeness of the information given, the necessity or not of obtaining a new individual consent for each use, and who is going to decide on the issue. Several views have been expressed, ranging from denying any use, other than that initially

Box 2 | Degree of possible identification of a sample donor

The possibility of tracing the person from whom sample and data were derived varies according to how the samples are linked to their donor identity in the database. Samples and associated information can be:

Identifiable. **The identity (or personal and unique social security number) of individuals is directly attached or linked to the samples or data.**

Traceable or coded. **A code is attached to them and the correspondence between code and identity is physically separated from sample and data. A limited number of people can connect the code to the identity.**

Encrypted. **There is a further level of protection through encryption (that is, the code is transformed into several characters that are linked to the code with the intervention of a third party). This third party intervention will then be required to trace individual identity.**

Anonymized. **The link has been irreversibly cut between sample/data and the individual identity.**

Anonymous. **There has never been any possibility to link the sample and the attached data to a given person.**

Table 2 | **Human post-genomic biobanks outside of Europe**

Biobank properties	Biobank Japan	CARTaGENE* (Québec, Canada)	Personalized Medicine Research Project	GRAD§ (USA)
Institution	University of Tokyo	RMGA (Network of Applied Genetic Medicine), planned creation of an Institute for Populations and Genetics (non-profit)	Marshfield Medical Research Foundation	Howard University National Human Genome Center (a private, non-profit institution, in partnership with First Genetic Trust)
Type	National. Hospital-patient based	Regional. Population based (1.5% of Québec population)	Regional. Patient based, recruitment among patients attending Marshfield clinic centres	Ethnicity based: African diaspora biobank
Aims	Forty diseases targeted. Mainly pharmacogenetics research and towards its clinical application	Québec population genetic variation of medical relevance, genetic mechanism of diseases and genomic research towards clinical applications	Study of the effect of genetic variation on specific diseases and medication	Study of common genes associated with diseases prevalent in the African diaspora
Notable features	Example of hospital-based recruitment	Detailed protocol; strategy of communication and public consultation refined	A regional US project based on one clinical centre	African diaspora targeted
Total number of individuals	300,000	60,000–65,000 individuals aged between 24 and 75 yrs old	400,000 over 18 yrs old	25,000 individuals of African descent (planned overseas recruitment evoked)
Budget	US$218 million for 5 yrs	NS. Still searching for funding	US$4 million	US$18 million
Year started (stopped or planned completion)	2004	Protocol development started in 1999. When the budget is secure, recruitment will be over 4 yrs	2002 (2005)	2004 (planned for 5 yrs)
Follow-up of individuals (prospective)	NS	Semi-longitudinal study	No	No
Current progress	Protocol completed, research started	Protocol almost completed after several workshops and consultations	10,000 individuals already recruited (September 2003)	Fund raising. Ethics protocol is being finalized
Further description	Launched with press conference in 2003, planned as a resource for several projects/users. Part of several large-scale projects to come under the MoE	Study population will be randomly selected from the social security registry	Homogeneous population mainly from central and northern German-European origin	Assumes that genetic susceptibility can be explained by ethnic genetic factors
Commercial aspects	Yes	No private ownership — possible access for private sector under specific contracts	All profits will go back to the Foundation	None so far
Ethical governance, informed consent and confidentiality	Specific framework, individual explicit consent, coded data, restricted access to code	Detailed; available and discussed through open workshops. Informed consent and multi-layer options. Double-coded information. Specific ethics committee	Consent asked for participation. Genetic results not included in patient medical record, for confidentiality reasons and because they are considered as research results. Coded information. Medical data in the database not connected to any other database or internet	Project included in competency of newly created GenEthics core of Howard University. Ethics is made a priority in this 'sensitive issues' project
Public debate organized	Press releases; information of Japanese public and Japanese industry in 2003; information available through Asian technology information programme; little debate	Yes. Central to the project: protocol modified to follow public perceptions and suggestions by switching from anonymization to double-coded information	Focus groups on the subject influence the protocol (for example, publication policy)	Press releases, debate on possible discrimination; active ethics core and experienced University in African diaspora studies contributes to building trust; but no plan of structured public consultation available. Several procedures evoked
Group opposition	None at this stage	None. Discussion favoured. Acceptance by the Québec population is one of the conditions of the project	No	None so far

*Belongs to P3G (Public Population Project in Genomics, http://www.p3gconsortium.org). ‡Wisconsin, USA. §Genomic research in the African diaspora. MoE, Ministry of Education; NIH, National Institutes of Health; NS, not specified; yrs, years. A more comprehensive version of this table is available online (supplementary information S3 (table)).

Box 3 | **The extent of biobank use, organization and quality**

Are new uses a real possibility for most biobanks? In many cases, the quality and the practical organization of the collections or their economical support make it difficult to envisage further use: the quality and storage conditions of many samples might not allow their multiple use, and the way in which the attached database is constructed might make it difficult to use the data. The insufficient quality standards of many public health sector or traditional research biobanks are one of the reasons that stimulated the creation of *de novo* commercial biobanks[7]. The OECD (Organisation for Economic Co-operation and Development) has dedicated a task force to design recommendations to regulate the quality and management standards of all kinds of biological resources centres[6].

Whereas the occurrence of further possible uses has not always been predictable, it is now a foreseeable event even if future uses cannot be defined in detail[11]. At the start of a new collection it is wise to set a policy for future uses in the protocol; the policy should outline the consent process, the necessary information to be given to participants and should address the issue systematically. This should also be one of the assessment criteria of the ethical review committee. Not mentioning this possibility at the time of first information and consent is deemed to constitute lack of transparency. Protecting the possibility of long-term broader use, within a framework that respects a subject's rights and that is organized according to approved ethical standards, is an important challenge for the ethical management of precious human samples and data.

stated, to more flexible attitudes. The latter take into account the traceability or not of the individual identity, the kind of further uses that are envisaged in relation to the original one, the implications of the research for the individual (so-called 'minimal risk' research being more easily allowed), how precisely the use was described at the time of sampling and, finally, the kind of consent that was originally granted, as described in the previous section. A common feature of all recommendations and regulations on this issue is that any unplanned use requires an authorization, with or without a new consent, following the consultation of an independent research ethics committee or institutional review board[11,14,23,25,38]. This body can itself make authoritative decisions, or can be only consultative with another administrative authority being in charge of the final decision.

Despite the unprecedented number of recommendations, bounding regulations and new laws on biobanks and genetic databases, new voices and new views are adding to the debate about existing and further uses[40]. Indeed, sharing or at least exchanging views in a large forum of stakeholders is one of the characteristics of the debate on the legitimate uses of biobanks (see also BOX 3). The public or the patients themselves have been consulted[29,39], and debates have involved the direct intervention of individuals or associations of patients rather than only ethical bodies and medical/scientific professionals. There has also been an increasing involvement of industrial interest in establishing and developing biobanks[19], in establishing commercial biobanks[7] and in promoting the use of biobanks for a range of applications rather than as part of a defined project[1,41]. A sample is now viewed as an evolving concept: when does a sample become a product and when does a sample become data, especially in the world of genetics[5]?

Protection of the person. In the context of biobanks, protection of the person is practically synonymous with controlling access to the data and use of such data. This operation should ensure that individuals or groups are not discriminated against and that medical and personal information is not disclosed to third parties (such as other family or community members, colleagues, employer or insurance companies). Absolute protection is a central issue in ethical analyses related to biobanks[10,11,13,16,20,23,38,42] and is best achieved through anonymized data[3,11,13] (BOX 2); however, absolute control of use by the sample donor is only possible when the link between a sample and its donor is maintained and is somewhat accessible. Those two aims are mutually exclusive. Information technology (IT) was once (rightly) seen as a source of possible danger to a person's rights; today, sophisticated IT tools are (ironically) one of the best ways to protect confidentiality[24]. Indeed, one argument for setting up larger biobanks is the greater ease with which one can access and develop sophisticated IT tools in such a context, which might be too difficult or too expensive to set up for small biobanks.

Nearly all biobanks can be used for genetic studies, primarily or secondarily. Genetic data are often considered as being particularly sensitive. In the recent 'UNESCO Declaration on Human Genetic Data' (adopted 16 Oct 2003)[16] one finds: "Human genetic data have a special status because: (i) they can be predictive of genetic predispositions concerning individuals; (ii) they may have a significant impact on the family, including offspring, extending over generations, and in some instances on the whole group to which the person concerned belongs; (iii) they may contain information the significance of which is not necessarily known at the time of the collection of the biological samples; (iv) they may have cultural significance for persons or groups. Due consideration should be given, and where appropriate special protection should be afforded to human genetic data and to the biological samples." The belief that genetic tests are unique and therefore justify special consideration with regard to informed consent and privacy is known as 'genetic exceptionalism'[43,44]. But not all genetic data have the same potential consequences for the person or its family (for example, tests for multifactorial diseases with low predictive value). There is therefore an ongoing debate about whether it is necessary to develop specific legislation for genetic data or for biobanks used in genetics, or, as I believe should be the case, whether both biobanks and genetic data should be regulated by existing or more general legislation.

Box 4 | **Assessing the impact of collaborative biobanks**

It is difficult to evaluate the economic and scientific impact of biobanks, as there is no standardized tool for this purpose. One proposed tool involves setting up indexes such as a Biobank Impact factor (BIF)[49] that would take into account the citation index of a given resource in literature and patent files (provided that a standardized way of citing the resources is set up) as well as the activity index of sample/data access or access authorization when required.

In the long term, such an index could help to evaluate the actual use of a biobank and the results derived from its uses. Such an index could help to recognize a biobank's activity and would encourage it to share its resources more widely. Although free access is often claimed to occur in the academic world, sharing is not necessarily the rule outside a given professional circle. A BIF would also provide ways of recognizing biobanking in evaluating researchers' activity, which is not rationally done at the moment; scientific evaluators currently have only empirical data to judge the scientific contribution of a given biobank. Resource sharing is not a trivial task, as organizing the optimal and transparent sharing of biological resources in the context of an economical development is one of the challenges of biotechnology.

Public–private collaborations and benefit sharing. Biobanks have created a host of new professions (such as the 'biobanker') and expertise, which need to be regulated by specific normative frameworks. New expertise is required to ensure high quality and highly organized biobanks, to manage the cost of banking activities, to be aware of the economical value of collections even in academic settings, and to manage the interest of pharmaceutical companies in collections intended for academic research[1,19]. On the one hand, commercial biobanks are being set up[7], whereas on the other hand, companies need to negotiate access to samples and data collected in the public research or health sectors for their assays.

Ownership, benefit sharing and return of results have been the subject of debates and proposals[13,45,46]. The non-commercial use of human body elements is respected in many countries, but the way in which the principle is applied varies; for example, there is generous compensation for using parts of the body in some countries but not in others, and countries vary in the degree of *in vitro* manipulation that is required to transform a human body sample into a manufactured product[47]. This non-proprietary treatment of body parts is underlined in a recent European Union Directive on setting standards for quality and safety for the donation, procurement, testing, processing, preservation, storage and distribution of human tissues and cells that are intended for human application[12]. However, this directive only applies to *in vivo* therapeutic use[12]. The legal frameworks that regulate samples, products and data are different because body parts, personal data and manufactured products of human origin are under different directives or laws[5,15]; a unified status that encompasses samples and data would simplify practical implementation for biobanks.

The ownership of samples, data and databases is a complex matter and source of tension. There is a conflict between the need for open exchange and intellectual property rights over the content of the databank[48]. The European Union Database Protection Directive of October 1995, which regulates data ownership, could lead to difficult situations owing to the complexity of biological databases. This might result in considerably slower access to data, which could discourage their use. This is such a pressing issue that a commission set up by the OECD, Health Canada and the Japanese Association of Legal Medicine conducted a survey in 2003 on 'Human genetic research databases: links with commercial organisations and the impact of the European database legislation'. It is my view that many restricted uses and opposition to sharing bioresources are a result of intellectual property rights or the control that scientists want to exert on the biobanks they have established with great effort, rather than ethical issues related to respect for the individual rights of donors.

Some specific minorities or developing countries have complained about the exploitation of human genetic data and resources that brings no return or benefit to the population. Recommendations exist[13,38,45] to regulate benefit sharing, although these issues have not been developed as extensively as the consent process. One of the many reported views recommends that benefits be shared not on an individual basis but at the population or group level. This can be done in various ways: a percentage of benefits could be contributed to health sector organizations, or donated to humanitarian or educational programmes; alternatively, benefit could be gained by obtaining free access to treatment developed through biobanks[46].

Although the principles of benefit sharing have been extensively discussed, they have yet to be translated into practical guidelines. Besides economic benefits, the access to medical or scientific results by participants is also increasingly seen as an issue. The way in which participants can access this information and the level of details that need to be provided are emerging questions in relation to biobanking and, so far, they have not received a harmonized response[11,38].

Conclusion

Biobanks represent a lively and changing area of scientific development that raises many societal and ethical challenges. These will need to be confronted through multidisciplinary approaches that engage not only academic scientists, but also specialists in the social and human sciences, the medical sector, the economical sector and society at large. Ethical issues need to balance resistance and willingness to share biological resources, benefits, views and concerns between stakeholders. Ethics should be seen as promoting the sharing of bioresources with increased transparency (BOX 3). The growing involvement of society is shaping not only the regulatory framework but the scientific uses of biobanks themselves. In future we will probably witness the development of international guidelines for managing human biobanks, a clarification of ownership issues and detailed conditions for use of biobank material (BOX 4). We will also probably develop imaginative ways of sharing benefits, and will see an increased involvement of patients or entire populations not only as donors but in the process of defining the uses and management of large biobanks. The consideration of ethical issues in such contexts is of primary importance as it is related not only to individual decisions or projects, but also to national policies in the international context and to the organization of the democratic debate about the place of science in society.

Anne Cambon-Thomsen is at the Institut National de la Santé et de la Recherche Médicale, Inserm U 558, Faculté de médecine, 37 allées Jules Guesde, F-31073 Toulouse Cedex, France.
e-mail: cambon@cict.fr

doi:10.1038/nrg1473

1. Hirtzlin, I. *et al.* An empirical survey on biobanking of human genetic material and data in six EU countries. *Eur. J. Hum. Genet.* **11**, 475–488 (2003).
2. Steinberg, K. *et al.* DNA banking for epidemiologic studies: a review of current practices. *Epidemiology* **13**, 246–254 (2002).
3. National Bioethics Advisory Commission USA. The use of human biological materials in research <http://bioethics.gov/briefings/index.htm#jan99> (1999).
4. Cambon-Thomsen, A., Ducournau, P., Gourraud, P. A. & Pontille, D. Biobanks for genomics and genomics for biobanks. *Comp. Funct. Genomics* **4**, 628–634 (2003).
5. Hansson, M. G. & Levin, M. (eds) *Biobanks as Resources for Health* (Uppsala Univ., Uppsala, 2003).
6. OECD. *Biological Resource Centres: Underpinning the Future of Life Sciences and Biotechnology* OECD code 932001041E1 (OECD, Paris, 2001).
7. Anderlik, M. Commercial biobanks and genetic research: ethical and legal issues. *Am. J. Pharmacogenomics* **3**, 203–215 (2003).
8. Cambon-Thomsen, A. & Rial-Sebbag, E. Aspects éthiques des banques d'échantillons biologiques. *Rev. Epidemiol. Sante Publique* **51**, 101–110 (2003) (in French).
9. Reymond, M. A., Steinert, R., Escourrou, J. & Fourtanier, G. Ethical, legal and economic issues raised by the use of human tissue in postgenomic research. *Dig. Dis.* **20**, 257–265 (2002).
10. Commission de l'éthique et de la technologie. *Les Enjeux Éthiques des Banques d'Information Génétique: pour un Encadrement Démocratique et Responsable* <http://www.ethique.gouv.qc.ca/fr/ftp/AvisBanquesGen.pdf> (Gouvernement du Québec, Montréal 2003).
11. CCNE. *Ethical Problems Raised by the Collected Biological Material and Associated Information Data: 'Biobanks', 'Biolibraries'* <http://www.ccne-ethique.fr/english/pdf/avis077.pdf> (French National Advisory Bioethics Committee, Paris, 2003).
12. The European Parliament and The Council of the European Union. Directive 2004/23/EC of the European Parliament and of the Council of 31 March 2004 on setting standards of quality and safety for the donation, procurement, testing, processing, storage and distribution of human tissues and cells. *Off. J. Eur. Union* **47**, L102/48–L102/58 (2004).
13. Godard, B., Schmidtke, J., Cassiman, J. J. & Ayme, S. Data storage and DNA banking for biomedical research: informed consent, confidentiality, quality issues, ownership, return of benefits. A professional perspective. *Eur. J. Hum. Genet.* **11** Suppl. 2, 88–122 (2003).
14. HUGO Ethics Committee. Statement on DNA sampling control and access. *Genome Digest* **6**, 8–9 (1999).
15. Matthiessen, L. Survey on opinions from National Ethics Committees or similar bodies, public debate and national legislation in relation to human biobanks (European Commission Research Directorate-General, Brussels, 2002).
16. UNESCO, Bioethics Committee. International Declaration on Human Genetic Data. <http://unesdoc.unesco.org/images/0013/001312/131204e.pdf#page=27> (2003).
17. Knoppers, B. M. & Fecteau, C. Human genomic databases: a global public good? *Eur. J. Health Law* **10**, 27–41 (2003).
18. Sallée, C. in *OECD Workshop on Human Genetic Research Databases: Issues or Privacy and Security* (ed. OECD) 47 (OECD, Tokyo, 2004).

19. Austin, M. A., Harding, S. & McElroy, C. Genebanks: a comparison of eight proposed international genetic databases. *Community Genet.* **6**, 37–45 (2003).
20. Caze de Montgolfier, S. *Collecte, Stockage et Utilisation des Produits du Corps Humain dans le Cadre des Recherches en Génétique: État des Lieux Historique, Éthique et Juridique; Analyse des Pratiques au Sein des Biothèques.* Thesis, Univ. René Descartes, Paris (2002).
21. Knoppers, B. M. (ed.) *Populations and Genetics. Legal and Socio-ethical Perspectives* (Martinus Nijhoff, Leiden, 2003).
22. Barbour, V. UK Biobank: a project in search of a protocol? *Lancet* **361**, 1734–1738 (2003).
23. Knoppers, B. M. (ed.) *Human DNA: Law and Policy. International and Comparative Perspectives* (Kluwer Law International, Boston, 1997).
24. Gulcher J. R., Kristjansson K., Gudbjartsson H. & Stefansson K. Protection of privacy by third-party encryption in genetic research in Iceland. *Eur. J. Hum. Genet.* **8**, 739–742 (2000).
25. White, M. T. & Gamm, J. Informed consent for research on stored blood and tissue samples: a survey of institutional review board practices. *Account Res.* **9**, 1–16 (2002).
26. Deschenes, M., Cardinal, G., Knoppers, B. M. & Glass, K. C. Human genetic research, DNA banking and consent: a question of 'form'? *Clin. Genet.* **59**, 221–239 (2001).
27. Sade, R. M. Research on stored biological samples is still research. *Arch. Intern. Med.* **162**, 1439–1440 (2002).
28. Chadwick, R. & Berg, K. Solidarity and equity: new ethical frameworks for genetic databases. *Nature Rev. Genet.* **2**, 318–321 (2001).
29. Kaye, J. Genetic research on the UK population — do new principles need to be developed? *Trends Mol. Med.* **7**, 528–530 (2001).
30. Greely, H. T. Informed consent and other ethical issues in human population genetics. *Annu. Rev. Genet.* **35**, 785–800 (2001).
31. Hoeyer, K. 'Science is really needed — that's all I know': informed consent and the non-verbal practices of collecting blood for genetic research in northern Sweden. *New Genet. Soc.* **22**, 229–244 (2003).
32. Ducournau, P. in *Normes et Valeurs dans le Champ de la Santé* (eds Cresson, G., Penec, S. & Schweyer, F.) 281–289 (Collection 'Recherche, santé, social', Rennes, 2004).
33. Harkin, M. The devil, the details, and the HGDP. *Politics Life Sciences* **18**, 301–303 (1999).
34. Lock, M. Interrogating the Human Diversity Genome Project. *Soc. Sci. Med.* **39**, 603–606 (1994).
35. Weiss, K. M. *et al.* Proposed model ethical protocol for collecting DNA samples. *Houst. Law Rev.* **33**, 1431–1474 (1997).
36. Foster, M. W., Bernsten, D. & Carter, T. H. A model agreement for genetic research in socially identifiable populations. *Am. J. Hum. Genet.* **63**, 696–702 (1998).
37. Quebec Network of Applied Genetic Medicine (RMGA). Statement on the ethical conduct of genetic research involving populations. <http://www.rmga.qc.ca/doc/ENONCE2002.ENG.pdf> (2002).
38. Nationaler Ethikrat. Biobanks for research. <http://www.nationalerethikrat.de/_english/publications/Opinion_Biobanks-for-research.pdf> (German National Ethics Council, Berlin, 2004)
39. Wendler, D. & Emanuel, E. The debate over research on stored biological samples: what do sources think? *Arch. Intern. Med.* **162**, 1457–1462 (2002).
40. Hannig, V. L., Clayton, E. W. & Edwards, K. M. Whose DNA is it anyway? Relationships between families and researchers. *Am. J. Med. Genet.* **47**, 257–260 (1993).
41. Blatt, R. J. Banking biological collections: data warehousing, data mining, and data dilemmas in genomics and global health policy. *Community Genet.* **3**, 204–211 (2000).
42. Anderlik, M. R. & Rothstein, M. A. Privacy and confidentiality of genetic information: what rules for the new science? *Annu. Rev. Genomics Hum. Genet.* **2**, 401–433 (2001).
43. Green, M. J. & Botkin, J. R. 'Genetic exceptionalism' in medicine: clarifying the differences between genetic and nongenetic tests. *Ann. Intern. Med.* **138**, 571–575 (2003).
44. McGee, G. Foreward: genetic exceptionalism. *Harv. J. Law Technol.* **11**, 565–570 (1998).
45. HUGO Ethics Committee. Statement on benefit sharing. <http://www.hugo-international.org/hugo/benefit.html> (2000).
46. Knoppers, B. M. Population genetics and benefit sharing. *Community Genet.* **3**, 212–214 (2000).
47. Nelkin, D. & Andrews, L. Homo economicus: commercialization of body tissue in the age of biotechnology. *Hastings Cent. Rep.* **28**, 30–39 (1998).
48. Bovenberg, J. Should companies set up databases in Europe? *Nature Biotechnol.* **18**, 907 (2000).
49. Cambon-Thomsen, A. Assessing the impact of biobanks. *Nature Genet.* **34**, 25–26 (2003).
50. Ohayon, E. & Cambon-Thomsen, A. (eds) *Génétique des Populations Humaines.* (Inserm, Paris, 1986).
51. Merz, J. F., McGee, G. E. & Sankar, P. 'Iceland Inc.'?: on the ethics of commercial population genomics. *Soc. Sci. Med.* **58**, 1201–1209 (2004).
52. Collins, F. The case for a US prospective cohort study of genes and environment. *Nature* **429**, 475–477 (2004)

Acknowledgements

The support of Toulouse-Midi-Pyrénées-Genopole, Réseau National des Genopole (RNG) and DRESS (Direction de la Recherche, des Etudes, de l'Evaluation et des Statistiques — Mission de Recherche du Ministère de l'Emploi, du Travail et de la Cohésion Sociale et du Ministère de la Santé et de la protection sociale) in connection with the 2004 programme 'Biomedical sciences, health, society' is gratefully acknowledged.

Competing interests statement

The author declares no competing financial interests.

Online links

FURTHER INFORMATION

Cambon-Thomsen's homepage: http://www-toulouse.inserm.fr/srv/bmip/fdrech/unites/U558/cambon.html
Human Genome Diversity Project: http://www.hgalert.org/topics/personalInfo/hgdp.htm
ISBER: http://www.isber.org/
Mannvernd (Association of Icelanders for Ethics in Science and Medicine): http://www.mannvernd.is/english/
OECD: http://www.oecd.org/home/

SUPPLEMENTARY INFORMATION

See online article: S1 (table) | S2 (table) | S3 (table)
Access to this links box is available online.

OPINION

The nature of stem cells: state rather than entity

Dov Zipori

Abstract | Stem cells are endowed with self-renewal and multipotential differentiation capacities. Contrary to the expectation that stem cells would selectively express specific genes, these cells have a highly promiscuous gene-expression pattern. Here, I suggest that the transient stem cell state, termed the 'stem state', may be assumed by any cell and that the search for specific genes expressed by all stem cells, which would characterize the stem cell as a cell type, might be futile.

Stem cells are defined by their ability to give rise to various mature progeny while maintaining the capacity to self-renew. The development of organs during embryogenesis depends on these cells and, in the adult, frequent cell loss is compensated for by the activity of stem cells. Stem cells are therefore indispensable for the integrity of complex and long-lived organisms.

Although embryonic stem (ES) cells were known to give rise to the complete range of cells in the organism, adult stem cells were initially thought to have a differentiation potential restricted to their tissue of origin. Recent studies have revealed that adult stem cells are unexpectedly common and indicate that they might be more plastic in their ablility to differentiate into cell types of all the three germ layers than previously appreciated[1]. The differences that previously seemed to exist between embryonic and adult stem cells were reduced to the point that it is questionable whether they exist at all.

The presence of multipotent stem cells in the adult might open up new therapeutic opportunities on the basis of tissue and organ replacement. Therefore, the exact definition of stem cells and the ability to isolate them are matters of supreme importance. However, despite the efforts of many investigators who strive to determine their nature, a definitive stem-cell 'portrait' is lacking. In recent years, two independent studies[2–4] claimed to have identified a stem-cell-specific group of genes that form a 'stem-cell signature'. In fact, these studies have defined two different and unrelated groups of genes; the conclusion that these signatures characterize stem cells is therefore premature. Experimental and/or technical reasons might explain the disparity of the results from these independent studies, and alternative approaches that might lead to identification of the 'correct' gene-expression profile of stem cells were suggested[5,6]. But should one expect to find a stem-cell-specific signature using an approach based on the analysis of gene expression? Below, I argue that the opposite is expected. In as much as stem cells that originate from various tissues and organs are alike, this resemblance relates to their proliferation and differentiation