

Community Engagement and Informed Consent in the International HapMap Project

Charles Rotimi^a Mark Leppert^b Ichiro Matsuda^e Changqing Zeng^h Houcan Zhangⁱ
Clement Adebamowo^j Ike Ajayi^j Toyin Aniagwu^j Missy Dixon^b Yoshimitsu Fukushima^f Darryl Macer^k
Patricia Marshall^d Chibuzor Nkwodimmah^j Andy Peiffer^b Charmaine Royal^a Eiko Suda^g Hui Zhao^h
Vivian Ota Wang^c Jean McEwen^c and The International HapMap Consortium

^aNational Human Genome Center, Howard University, Washington, D.C., ^bDepartment of Human Genetics, University of Utah, Salt Lake City, Utah, ^cNational Human Genome Research Institute, US National Institutes of Health, Bethesda, Md., and ^dDepartment of Bioethics, Case Western Reserve University School of Medicine, Cleveland, Ohio, USA; ^eHealth Sciences University of Hokkaido, Hokkaido, ^fDepartment of Medical Genetics, Shinshu University School of Medicine, Matsumoto, and ^gUniversity of Tsukuba, Eubios Ethics Institute, Tsukuba Science City, Japan; ^hBeijing Genomics Institute, Chinese Academy of Sciences, and ⁱBeijing Normal University, Beijing, China; ^jUniversity of Ibadan College of Medicine, Ibadan, Oyo State, Nigeria; ^kUnited Nations Educational, Scientific and Cultural Organization (UNESCO), Prakanong, Bangkok, Thailand

Key Words

Community engagement • Community consultation • Public consultation • Informed consent • International HapMap Project • Genetic variation research

Abstract

The International HapMap Consortium has developed the HapMap, a resource that describes the common patterns of human genetic variation (haplotypes). Processes of community/public consultation and individual informed consent were implemented in each locality where samples were collected to understand and attempt to address both individual and group concerns. Perceptions about the research varied, but we detected no critical opposition to the research. Incorporating community input and responding to concerns raised was challenging. However, the experience suggests that approaching genetic variation research in a spirit of openness can help investigators better appreciate the views of the communities whose samples they seek to study and help communities become more engaged in the science.

Copyright © 2007 S. Karger AG, Basel

Background

The International HapMap Consortium has completed the first phase of the International HapMap Project, an effort to develop a haplotype map of the human genome that describes the common patterns of DNA sequence variation [1]. The HapMap will be used as a resource to facilitate future studies that relate genetic variation to health, disease and drug response [2].

DNA samples from 4 populations were studied in the first phase: Yoruba from Ibadan, Nigeria; Japanese from Tokyo, Japan; Han Chinese from Beijing, China; CEPH (Utah, US residents with northern and western European ancestry; table 1). Investigators in Japan, the UK, Canada, China and the US analyzed all samples across the genome to determine their haplotype structure (the patterns of genetic variation). Based on the data generated, future investigators searching for genes that contribute to disease will be able to choose tag single nucleotide polymorphisms (SNPs, or sites in the DNA sequence where individuals vary) that they can use to conduct their studies much more efficiently.

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

© 2007 S. Karger AG, Basel
1422–2795/07/0103–0186\$23.50/0

Accessible online at:
www.karger.com/cm

Dr. Jean McEwen
US National Institutes of Health, National Human Genome Research Institute
5635 Fishers Lane, Suite 4076
Bethesda, MD 20892 (USA)
Tel. +1 301 402 4997, Fax +1 301 402 1950, E-Mail mcewenj@mail.nih.gov

Table 1. Populations included in developing the genome-wide HapMap

Population	Samples collected	Criteria for defining population membership
Yoruba in Ibadan, Nigeria (YRI)	parent-adult child trios	4 of 4 Yoruba grandparents
Japanese in Tokyo, Japan (JPT)	unrelated individuals	'aim to collect from individuals whose grandparents were all Japanese'
Han Chinese in Beijing, China (CHB)	unrelated individuals	at least 3 of 4 Han Chinese grandparents
CEPH (CEU) (Utah residents only; individuals with Northern and Western European ancestry)	parent-adult child trios	not stated*

* The original aim was to collect samples from large three-generation families suitable for the construction of genetic linkage maps.

In a later phase, samples from several additional populations will be analyzed across a subset of genomic regions to assess how well the tag SNPs based on the data from the 4 initial populations will work in other groups. If these tag SNPs do not adequately capture the haplotype patterns in those populations' samples, additional tag SNPs based on the data from these populations may need to be identified and added to the database. Eventually, other investigators will likely provide data from other populations.

The blood samples collected to develop the HapMap were transformed into cell lines at the National Institute of General Medical Sciences (NIGMS) Human Genetic Cell Repository at the not-for-profit Coriell Institute for Medical Research in Camden, New Jersey. The Coriell Institute makes the cell lines and DNA available both to Project investigators and to investigators conducting other genetic variation research approved by the institutional review board (IRB). The samples have no individual identifiers or associated medical information. However, each sample set is identified with a population label, so that investigators can compare patterns of genetic variation not only among individuals, but among groups. This facilitates the selection of optimal tag SNPs for use in genetic association studies in specific study populations. However, labeling the populations raises complex ethical and social issues, because all members of a population and of closely related populations may be affected by the research, regardless of whether they personally donated samples.

For this reason, and because of the wide range of cross-cultural issues raised in a large, complex international study of this type, the Project devoted considerable time

and resources to addressing the ethical and social issues. Bioethicists as well as social and behavioral scientists worked alongside genomics researchers in formulating important aspects of the study design. The Project also implemented processes to engage a range of people in each of the communities approached for the donation of new samples as well as a rigorous informed consent process, so that both individual and group concerns could be anticipated, understood and, to the extent possible, addressed. In this paper, we describe these processes, review some preliminary impressionistic findings and explain how the Project responded to some issues raised. An earlier paper provided a general discussion of the ethical, social and cultural issues the Project raises [3]. Later articles will describe in detail the methods and findings from the individual sites discussed in this paper and from the additional communities approached for later participation.

Populations and Communities

While recognizing the considerable complexities inherent in defining such terms as 'population' and 'community', for the purposes of the Project, a 'population' was defined as a group of individuals who have a common geographic ancestry, while a 'community' was defined as a group within a population with many local units of social organization [4]. The scientific, ethical and practical rationales for the decisions about which populations to include and which specific communities to engage, in the initial iteration of the HapMap, have been previously described [1, 2].

The CEPH samples studied for the Project were collected in 1980 from residents of Utah whose recent ancestors came primarily from northern and western Europe. The label CEPH is an acronym for the Centre d'Etude Polymorphisme Humain, the organization that originally collected the samples. Cell lines for the CEPH samples have been publicly available since 1992 from the Coriell Institute, and the samples have been used in numerous previous genetic studies, including the development of the human linkage map. HapMap investigators wanted to build on this valuable body of existing data, and thus chose to use a subset of these same samples to develop the HapMap. Although the original complete CEPH sample set includes some samples from individuals in France and Venezuela and individuals identified as Amish from Pennsylvania, only samples from the Utah CEPH donors were used for the Project. Although the CEPH population is the only one whose samples are included in the HapMap to be known by an acronym, investigators chose to retain this population label for the Project so that the HapMap data could be readily integrated with preexisting data from other studies of these samples without engendering confusion within the scientific community.

The consent process originally used to collect the CEPH samples, while quite comprehensive, did not meet the stringent standards established for the Project. However, because the investigators who had collected the CEPH samples had retained links to the donors' identities and had developed trust relationships with many donors, it was feasible for them to recontact most of the still living donors to seek new consent for their samples to be used specifically to develop the HapMap. The local IRB gave permission for deceased donors' samples to be used. Because the local IRB required the maintenance of absolute confidentiality with respect to the identities of the CEPH donors (who are known to others in their own families, but not to any other CEPH donors), it was not feasible to reconvene the CEPH donors as a group for a formal community engagement process analogous to that conducted for the Yoruba, Japanese and Han Chinese.

Unlike the case with the other 3 populations whose samples were studied, the criteria used to assess ancestry to determine eligibility for sample donation were not specified in the case of the CEPH donors. This again, however, reflects merely the historical reality that it was not the norm to explicitly define the criteria for population membership when the CEPH samples were collected. What is known is that all the donors resided in Utah and that most of their recent ancestors (like the recent ancestors of most Utah residents in the areas where the

samples were collected) came from northern or western Europe [5].

The Yoruba samples were collected in Ibadan, the second largest city in Nigeria, with a population of nearly 2 million. The Yoruba are predominantly urban dwellers with a complex population history and a complex political and social organization. The group constitutes the majority population in Ibadan and approximately 30% of Nigeria's total population. Around 40 million individuals throughout West Africa self-identify as Yoruba [6]. Through previous research collaborations, the investigators enlisted to collect the samples had already developed a close working relationship with a Yoruba community in a particular area of metropolitan Ibadan. A robust approach to community engagement was thus designed to include as many community members as possible. To be eligible for sample donation, individuals were required to have 4 of 4 grandparents who self-identified as Yoruba.

In Japan, because of the population's relative ancestral homogeneity, it would have been possible to approach people for participation almost anywhere in the country. However, the sample collection took place in Tokyo, which draws people from all geographic areas of Japan. It included individuals who were, for the most part, already accustomed to participating in research. Individuals wishing to donate were simply told that the aim was to include samples from persons whose grandparents were all from Japan; donors were not asked whether they had a certain number of grandparents 'born in Japan' because it was thought that some people might find this question culturally insensitive. People from many parts of Japan, and especially from the Kanto area surrounding Tokyo, participated in the community engagement activities. Thus, people from a wide range of backgrounds were consulted. In addition, some input was obtained through conference presentations in several other countries.

The Han Chinese population is the largest of 56 ethnic groups in China; about 90% of all Chinese people self-identify as Han [7]. The specific community involved encompassed the entire residential community at Beijing Normal University (BNU), which includes almost 35,000 people, nearly all of Han ethnicity. Due to the wide geographic area and range of backgrounds from which BNU draws its residents, the community engagement, while situated in an academic environment, drew individuals with a range of backgrounds and ages. These individuals came originally from 22 of 34 Chinese provinces, autonomous regions, municipalities and special regions. For reasons of practicality, individuals who were approached to donate

samples were told that they should have at least 3 out of 4 grandparents born in China who self-identified as Han. However, given the BNU community demographics, all 4 grandparents of most donors were presumably Han.

Goals and General Approach

The goals of the individual informed consent process were to provide prospective sample donors with the information needed to ensure that their decision to donate was voluntary and informed. The goal of the community engagement processes was to give a broad range of members of the communities approached for participation an opportunity:

- to share their views about the ethical, social and cultural issues the Project raised for them, their immediate communities, and the broader communities and populations of which they are a part
- to provide input into such matters as how the samples from their locality would be collected and described
- to obtain extensive information about the Project so that the decisions of individuals whether to donate would be better informed
- to remain informed about how the HapMap and the samples are being used and about findings from future studies based on the HapMap or the samples.

Because it would have been impossible to seek input from all, or even most, people around the world who shared the relevant population labels, we focused our efforts primarily on the level of the specific localities from which we hoped to recruit donors. We recognized the significant limitations inherent in this approach, but reasoned that through in-depth, detailed inquiries in these communities, using a range of methodologies, we could reach a reasonably large and diverse range of individuals who, by virtue of sharing the same population labels with the actual sample donors, would most likely be affected by the research.

The aim of the engagement processes was not to achieve consensus or 'community consent,' even within these selected localities, nor to seek lay input into the advisability, as a matter of science policy, of launching a project of this type. In this respect, the approach differed from the 'participatory action' research model used in some studies of public health interventions, in which communities advise investigators on research priorities and have considerable input into major aspects of study design. The approach also differed from the formal community consultation processes required when conduct-

ing research in communities with sovereign status or highly organized political structures, such as American Indian tribal communities.

The specific approaches employed at each site to engage communities and to elicit individual informed consent were informed by relevant sets of then-existing local, national and international guidelines [8–17]. The activities were conducted under the auspices of local ethics committees.

Investigators were also guided by the community consultation policy of the NIGMS Human Genetic Cell Repository at the Coriell Institute (<http://ccr.coriell.org/nigms/comm/submit/collpolicy.html>). This policy requires some form of community consultation or engagement, tailored to local cultural norms, before the repository will accept any new samples with population identifiers for its genetic variation panels. It also provides for the establishment of a community advisory group (CAG) in each community where new samples are collected, to serve as a liaison between the community and the Coriell Institute to ensure that all research using the samples is consistent with the terms of informed consent (see Appendix 1).

In China and Japan, the community engagement/public consultation and sample collection activities were funded by those countries' participating genotyping centers, and local investigators conducted the work. In Nigeria, the work was funded by the US National Institutes of Health (NIH) and conducted by US investigators in collaboration with local investigators. The NIH also funded the process of obtaining new consent from the CEPH donors and the collection of samples from the additional populations whose samples will be studied in a later phase. At each site where new samples were collected, the community engagement and sample collection teams included individuals trained in genetics, individuals with background or training in bioethics or social science, and others (table 2).

Methodologies for Engaging the Communities and Obtaining Informed Consent

At all 3 sites where participants were approached to donate new samples, protocols were developed to engage the communities; these protocols were separate from those used to obtain individual informed consent. As noted earlier, community engagement of the CEPH donors, at least in a form analogous to that used with members of the other 3 populations, was not feasible because of IRB-imposed constraints related to individual donor

Table 2. Community engagement/public consultation teams

Population	Investigators	Institutions	Backgrounds
Yoruba	Charles Rotimi Clement Adebamowo Patricia Marshall Charmaine D.M. Royal Ike Ajayi Toyin Aniagwu Chibuzor Nkwodimmah	Howard University University of Ibadan Case Western Reserve University Howard University University of Ibadan University of Ibadan University of Ibadan	genetic epidemiology epidemiology anthropology/bioethics genetics/bioethics epidemiology nursing nursing
Japanese	Ichiro Matsuda Darryl Macer Eiko Suda Yoshimitsu Fukushima	Health Science University of Hokkaido Eubios Ethics Institute Eubios Ethics Institute Shinshu University	genetics medicine bioethics bioethics genetics
Han Chinese	Houcan Zhang Changqing Zeng Hui Zhao	Beijing Normal University Beijing Genomics Institute Beijing Genomics Institute	psychology molecular biology genetics
CEPH	Mark Leppert Missy Dixon Andy Peiffer	University of Utah University of Utah University of Utah	genetics psychology medicine

privacy. At the other 3 sites, however, no individual was approached to donate a sample until the process of community engagement was already well underway. Copies of the consent forms were distributed widely in each community from the Project's inception, however, to introduce the study and initiate discussion about its potential risks and benefits.

Templates for an informed consent form for sample donation and to obtain new consent from the living CEPH donors were developed by an initial planning group, with input from bioethicists, social and behavioral scientists, and geneticists. Each team of investigators responsible for community engagement (or, in the case of the CEPH donors, the team responsible for obtaining new consent) modified the consent documents as needed to make them culturally appropriate for their locality. Individuals in the communities where new samples were collected were subsequently given an opportunity to provide direct input into the consent form, although that process in most cases did not lead to substantive modifications. Both the modified informed consent forms and the sample collection protocols were reviewed by IRBs or ethics committees at all the institutions involved.

The specific approaches to engaging the communities and obtaining individual informed consent varied among the sites because of the vastly different community struc-

tures and cultural norms. In the Yoruba community approached for participation, it was necessary formally to consult a community leader (the Baale) before any individuals were approached. In China, investigators secured cooperation from the BNU administration and several academic departments before beginning their work. In Japan, where most of the work was carried out in a large urban area and the community was more loosely organized, the approach was more open ended.

For the CEPH, where the samples had already been collected and where thus, for historical reasons, the donors could not technically be engaged as a 'community' but instead only as individuals or families, the donors were merely approached to give new consent, using an individualized or family-based approach instead of a group-based process. Because of the continued interactions with the donor families, investigators were able to locate 44 of the original 47 families. Many were already involved in a separate ongoing genetics research project that required them to return periodically for follow-up. This gave investigators an opportunity to discuss the HapMap Project in person with them. Those donors who had already revisited the investigators were contacted by mail, with telephone follow-up. The remaining donors who had not been in recent contact or who did not initially reply were visited at home or called by a study co-

Table 3. Methodologies

Population	Individual interviews	Focus groups	Public meetings	Attitudinal surveys	Other
Yoruba	7	1	3	231	initial working group
Japanese	20	8	5	377	5 conference presentations; 10 explanatory meetings
Han Chinese	100	6	3	130	production of 9 min video compact disc (VCD) used to introduce the project to interviewees
CEPH (new consent)					personal visits, mail contact, telephone follow-up

ordinator, who explained the Project in detail and gave them a chance to discuss it and ask questions.

Language and comprehension issues presented a major challenge during the community engagement and informed consent processes (see Appendix 2). Thus, open-ended discussions about the Project and the issues that it raises were encouraged. Major points contained in the consent form were explained orally and individuals seeking additional clarification had an opportunity to ask questions. In Nigeria, depending on their preference, participants were administered the informed consent procedures orally or in writing, either in English or Yoruba.

A detailed description of the methodologies employed for the community engagement activities at the 3 sites where new samples were collected is beyond the scope of this paper, but will be outlined in separate papers by the investigators from each site. Methodologies ranged from the use of extended, semistructured individual interviews and focus groups to large public meetings or lectures (followed by discussion) and public attitudinal surveys (table 3). The extensiveness of the processes employed varied considerably among the sites due to differences in the level of funding available for these activities. The empirical rigor of the processes for collecting and analyzing the data also varied somewhat from site to site. However, the processes were not specifically designed to provide data that were explicitly comparable across sites. The processes were rather designed simply to elicit the views of a range of people within each community, including those skeptical about genetics research, to glean general impressions (which in some instances could be little more than anecdotal) about the acceptability of the goals of the Project and other pertinent issues. Open discussion was encouraged at each site so that investigators could be alerted to any specific concerns and address

them to the extent possible. Thus, participants at each site were asked about their attitudes toward genetic research in general and genetic variation research in particular (including research like the HapMap, which would have no immediate health benefits). Participants were given an opportunity to raise concerns about proposed methods of recruitment, privacy and confidentiality risks, risks of discrimination or group stigmatization, issues relating to commercialization and intellectual property, and any other pertinent matters.

Specific Elements of the Informed Consent Process

The consent process explained how the HapMap would be developed and used in future studies to find genes that affect diseases, drug response and other traits. Prospective donors were told that neither names nor medical information would be taken – only the name of the population from which the donor came – and that more samples would be collected than would be used, as an additional protection of individual privacy. For the CEPH donors, links to the donors' identities exist, but only the investigators who collected the samples – not the repository, HapMap investigators or any other investigators who order the samples – have access to this information.

Prospective donors were also told during the consent process that the samples would be sent to the Coriell Institute, transformed into cell lines and made available to Project investigators and other investigators worldwide for use in future genetic variation studies. All such future studies, however, need to be approved by the repository's IRB (and any other relevant ethics committees) to ensure that the proposed research is consistent with the terms of the consent form.

Because no individual identifiers are available for any of the newly-collected samples, it will not be possible to recontact donors in the future to seek their consent to other studies. However, the consent form described the general nature of future studies for which the samples and the HapMap may be used, such as studies of the biology of DNA, how new variations arise, the genetic history of human groups, and how people from different parts of the world are related. The consent form also expressly authorized use of the samples for gene expression studies, but forbade their use for reproductive cloning. The reference to cloning was included because of concerns expressed in some communities about this possibility. It was explained that the risks raised by the types of future studies authorized in the consent form are unlikely to be different in kind from the risks raised by the Project itself.

Prospective donors were also told that because of the absence of individual identifiers, donors could neither receive individual feedback on the research findings nor individually withdraw their samples or data. However, in each community where new samples were collected, it was explained that a CAG would be established, through which the community would be able to stay informed about general findings and how the HapMap and samples are being used. Prospective donors were also informed that a community could request, through its CAG, that all of its samples be withdrawn from distribution in the unlikely event that a serious disagreement about future uses of the samples arises that cannot otherwise be resolved.

Prospective donors were informed that the HapMap would be publicly available in a database on the Internet. It was explained that the genetic information available about each donor would be quite extensive, but that it was very unlikely that any information could be linked to a specific donor, at least without having another sample from the donor for comparison.

Prospective donors were also told that they would receive no immediate health benefits from donating samples; any benefits would likely come only in the future, as investigators use the resource to find genes related to disease and then gradually develop improved methods of prevention, diagnosis and treatment. Donors were also informed that they would receive no financial benefits from participation, except for nominal compensation for their time and travel. They were further advised that while the Project itself would generate no commercial products, such products might be developed from other studies based on the stored samples or information in the

HapMap, and donors would not be able to share in any such profits.

The consent form specifically mentioned the potential group risks associated with genetic variation research, such as risks of group stigmatization or discrimination (if investigators in future studies were to find that genetic variants associated with a particular disease were more frequent in people from their group and this information were overgeneralized to all or most members of the group or to related groups). It was also explained that focus on group differences might 'reify' notions of race, thus potentially exacerbating societal prejudices.

In the Yoruba community, where investigators collected samples from parent-adult child trios, the procedures for handling findings of misattributed paternity or undisclosed adoption were also described; these procedures had been similarly described to the members of the CEPH donor families at the time the CEPH samples were originally collected. Prospective donors who were concerned that someone in the family might not be biologically related were advised that they could, but did not have to, disclose this information. The Coriell Institute would test all samples in the trios for relatedness and if it were to be found that not all family members in a trio were biologically related, no one would be told and the samples simply would not be used for the HapMap.

Although the Coriell Institute routinely tests all samples it obtains for the presence of HIV (and destroys any found to be infected), the Yoruba sample collection team decided to require HIV testing of all prospective donors prior to donation (with the opportunity for follow-up referrals and treatment where indicated). All the relevant IRBs approved this procedure. Yoruba community members viewed the opportunity for free HIV testing as a form of benefit associated with Project participation (although no one who underwent the testing had a positive test result). A separate process was used to obtain informed consent for the HIV testing.

Perceptions of Risks and Benefits

It is impossible to generalize from the perceptions of potential risks and benefits expressed by a small subset of individuals in the few specific localities where we did our work to everyone in these communities or to other communities. As noted earlier, most community engagement activities were designed to glean only general impressions about reactions to and concerns about the Project, and not as rigorous empirical data gathering exercises. The

general responses from those who took part in the community engagement activities are, however, instructive in demonstrating a range of perceptions about genetic variation research.

While the scientific details of the Project seemed difficult for many people to understand, most individuals – even in instances where substantial linguistic and educational barriers were present – appeared to comprehend the Project’s general purpose, as judged (albeit imperfectly) by what they communicated to investigators when asked to explain their understanding of why they were being asked to give samples. They also appeared to be able to understand that they would receive no immediate personal health benefits by participating but that future generations might benefit.

People consulted in Japan expressed a diversity of views, ranging from skepticism, through indifference, to a generally positive attitude about the potential of genetic variation research. A similar range of views was expressed in the BNU Han Chinese residential community, but among most people there, the Project was quite favorably received.

The Yoruba community as a whole demonstrated considerable enthusiasm about the Project. Many individuals expressed a strong sense of pride that their community had been selected as a possible sampling site for a major international biomedical research effort, especially when the ‘Out of Africa’ theory of human population history was discussed. Several people there commented that genetic variation research, by demonstrating people’s biological relatedness, might in some way help bring the world’s people together – especially Yoruba or other people with African ancestry separated from their roots through slavery. Few people, even when probed, expressed concern that the Project could exacerbate racial or ethnic divisions.

Among the CEPH donors approached for new consent, stated reactions to the Project were generally quite positive, as reflected in the fact that to date, 367 living individuals including third-generation offspring (far more than the number whose samples were needed) have consented to their samples being used, either for the HapMap or for other genetic variation research. Where investigators had the opportunity to discuss the Project in person, 95 out of 95 individuals agreed to participate. This very high rate of acceptance of the Project among the CEPH donors may, of course, be primarily a reflection of the relatively high socioeconomic status of this particular group of donors, coupled with their long, successful history of collaboration in other genetic studies; it is unclear

how generalizable this finding would be to other groups.

Some specific concerns about the Project were expressed. Although the data from each site have not yet been fully analyzed, a preliminary review suggests the emergence of a few predominant themes in each community.

Among the Yoruba, where most of the participants in the community engagement process were lay individuals with no background or training in biomedical research or related issues, the most frequent concerns raised were about:

- the physical process of blood drawing
- how the blood samples would be handled
- the disposal plans for the blood samples that were not used.

In Japan, where many though by no means all of the individuals approached had some sophistication about biomedical research, the main concerns expressed were about:

- privacy and confidentiality
- how the samples would be labeled
- whether the HapMap would somehow be used to try to define genetically who is a ‘real’ Japanese person
- the potential for discrimination against minority groups in Japan and against Japanese people living as minorities in other countries
- the potential for commercial use of the stored samples, especially by US biotechnology companies
- how adequate oversight over the samples would be ensured once the samples had been sent to a US repository.

In China, where the community engagement process involved a broad range of individuals, with varying degrees of sophistication regarding biomedical research, the main concerns (although voiced by only a few participants) were that:

- the Project would not include samples from all ethnic groups in China
- the samples might be used for reproductive cloning
- information from the samples might be used for forensic purposes
- the blood drawing might cause infection
- personal genetic information might be ‘leaked’ outside the Project
- knowledge derived from the Project could lead to discrimination (although few people reported that they thought the Project itself would exacerbate racial or ethnic tensions).

A few CEPH donors asked to give new consent to have their samples used for the Project questioned whether by allowing this, their privacy and the strict controls on access to phenotypic information would continue to be maintained rigorously. These donors were reassured, however, when it was explained that the HapMap would include no individual identifiers and that the links to the identities kept by the local investigators would never be shared with other Project investigators, the repository or any future investigators who may use the samples (except for collaborators of the sample collectors).

General Reactions to Genetic Variation Research

Despite the occasional expression of the concerns outlined above, we detected no critical opposition to the Project or to genetic variation research in general, at any of the sites. In particular, we found what appeared to be an absence of widespread concern about the potential for group stigmatization that might result from studies that use the HapMap. While some critics of genetic variation research might interpret this as indicative of people's naiveté or inability to understand the Project, some alternative explanations are plausible.

One explanation is that most of the people approached for participation were living in societies that were either racially or ethnically quite homogeneous (in the case of the Japanese) or in nonhomogeneous societies where they were members of the majority populations (in the case of the Yoruba and Han Chinese), and this circumstance may have considerably influenced their viewpoints. For example, in Nigeria, because most individuals' ancestors come from the African continent, considerably less 'race consciousness' can be observed than exists in the US, for example. Many Nigerians do have a strong sense of ethnic identification, but few people expressed concern that genetic variation research alone was likely to have much impact on ethnic divisions within the country or more broadly. Japan is a relatively ethnically homogeneous society, and in China the specific community engaged was composed almost exclusively of individuals from the majority (Han) population. Also, in both Japan and China, as in Nigeria, racial issues do not figure nearly as prominently in most people's thinking as they do in the thinking of people from Western countries.

Individuals' and communities' conceptions of the risks associated with genetic variation research can be expected to vary, depending on how they construct their social identities and how those identities are perceived by

others. For this reason, we must acknowledge that we may well have encountered many more expressions of skepticism about the Project had we actively sought out the views of minority group members (for example, Korean or Chinese individuals living in Japan, members of some of the non-Han ethnic groups in China, or members of some of the numerous ethnic groups in Nigeria that are smaller and thus potentially more vulnerable than the Yoruba). Indeed, members of some such groups might have wondered why samples from their populations were being 'excluded' from study – a circumstance that underscores the need for sustained engagement activities with more broadly based participation. We also do not know how similar groups in other countries, or groups perceived by others as similar to those participating in the Project, conceive risks associated with this research. This must be regarded as a significant limitation of these preliminary findings from the community engagement activities.

It is unclear whether the generally positive views about this type of research expressed by those with whom the Project has been discussed so far will turn out to be shared by most of the rest of the public, as awareness of the Project around the world grows. It is, however, worth observing that historically, concerns about the potential of genetic variation research to exacerbate forms of racism or ethnic tension, and about the tendency of such research to overfocus on genetics instead of the environment as a major causative factor in disease, have often emerged not so much from grassroots lay communities as from segments of the bioethics, social scientific and other professional communities. This is not to suggest that these concerns are without any basis, but merely to note that they may not be as widely shared by members of the public as some might assume. It will be instructive to learn how members of the additional populations approached for participation in a later phase of the Project – some of which are racial and ethnic minorities in the US – will react to the research.

Incorporating Community Input

While many scientific aspects of the Project design were essentially fixed for scientific reasons, limiting the extent to which communities could alter the study design, certain practical aspects of the recruitment and sample collection processes were modified in direct response to information obtained in community discussions. One especially important matter on which com-

munity input was weighed was how to label each population's samples and associated data. An exception to this were the CEPH donors, where, as discussed earlier, a formal process of community engagement was not feasible, where the acronym 'CEPH' had already been chosen when the samples were first collected, and where adopting a different label for these samples for this Project would likely only have engendered confusion in the scientific community.

In the places where new samples were collected, however, deciding how to label the samples required the disentanglement of complex notions of socially and genetically defined identity. While the selection of populations to be included in the Project was based on ancestral geography, the communities engaged in or consulted for the Project were, for the most part, composed of people with a range of *socially and culturally* (not genetically) defined identities. For example, some individuals in the Ibadan, Nigeria community engaged for participation considered themselves socially or culturally Yoruba even though they were not technically eligible to donate because not all of their grandparents were members of that specific ethnic group. Some individuals in the BNU residential community, while identifying as Han Chinese, may have had recent ancestors of other Chinese ethnicities. A small number of people in Japan who identify as, and culturally are, Japanese have some recent ancestry from Korea or other parts of Asia. In many of these cases, moreover, the basis for these individuals' constructions of their social identities is unknown by others. Many individuals also construct their social identities primarily around religious, political or other affiliations, not around ethnicity or ancestry. In addition, some individuals – perhaps most – view themselves simultaneously as belonging to several groups.

Thus, while each community provided input into how its population's samples should be labeled, the final decisions about labeling were based on a mix of community input and scientific, ethical and practical considerations. For example, in Japan, where this issue was discussed extensively, the names 'Asian' and 'East Asian' were rejected because these broad geographic areas include much greater ancestral diversity than just Japanese. Ultimately, the more specific descriptor 'Japanese in Tokyo, Japan' (JPT) was chosen. Likewise, because Han is only one of many Chinese ethnicities, the label 'Han Chinese in Beijing, China' (CHB) was chosen over the more general descriptor 'Chinese', and 'Yoruba in Ibadan, Nigeria' (YRI) was chosen over such terms as 'African', 'Sub-Saharan African', 'West African' or 'Nigerian'. To avoid overgen-

eralizing the results from any studies of these samples, users of the HapMap and of the samples are directed to a page on the Project website that explains the importance of using these specific terms, and not, for example, using 'African' for the Yoruba samples (<http://www.hapmap.org/citinghapmap.html.en>).

Our experience suggests that discussing the pros and cons of particular population identifiers with communities can be instructive – both to help communities understand the rationale for the study of genetic variation and to help investigators understand how people's own socially constructed notions of identity may differ from the identities geneticists seek to ascribe to them. This ultimately contributes greater rigor to the way the data are interpreted. Such discussions can also help elucidate other community concerns. For example, investigators may learn that a particular locality does not want its specific name used to better preserve its privacy.

Responding to Community Concerns

During the course of engaging the communities, several issues arose that necessitated a considered response. For example, Yoruba community leaders used the occasion of a site visit by a staff member from the NIH (the agency that funded the community engagement) to request funds to contribute to the building of a local health center. Emerging local standards of bioethics in Nigeria, as stated in that country's proposed National Code of Health Research Ethics, direct that in certain international collaborative studies, 'research should be integrated with comprehensive capacity building, technology transfer and health care delivery strategies that address significant local health problems' [18]. Consistent with this guideline, and in recognition of the fact that Nigeria, unlike the other participating countries, would receive no benefits from participating in the genotyping (the most heavily funded part of the Project), the NIH had already provided modest funds at the beginning of the Project. These funds had been used to enhance the basic preventive and primary care services already available locally and the local collaborators received training and equipment. However, later in the course of interacting with the community, some additional funds were sought as a demonstration of reciprocity for the community's contribution to the Project.

When the request was made, the community had already committed itself to participating and sample collection had already begun. Thus, there was little potential

for undue influence; individual participants there, as in all communities, were compensated only for time and travel. Nonetheless, deciding how to respond to the request raised ethical and practical challenges.

Several factors argued in favor of providing such funds. First, as already noted, unlike in the 3 other countries, where local investigators benefited directly by participating in the genotyping, no one in Nigeria was in a position to do this, and thus that country was differently situated. Coupled with this, the HapMap itself also will provide no direct, immediate health benefits to donors; yet the samples will be made available to multiple investigators around the world, including many in biotechnology and pharmaceutical companies in countries with better-developed biomedical research infrastructures. These investigators will receive considerable financial benefit from future studies based on the HapMap, as they develop and commercialize useful therapeutic and diagnostic applications. Realistically, these applications will take much longer to reach Nigeria than countries with better-developed health care delivery systems.

On the other hand, some members of the Project's Ethical, Legal and Social Implications Group and of the Project's Steering Committee were concerned that providing the requested funds might create a troublesome precedent, leading to a climate in which future investigators – especially local investigators without support from large funding agencies – would find it hard to conduct their studies. Questions were also raised about potential inequities with other participating HapMap communities. Existing sets of international guidelines on biomedical and research ethics provided limited help. While such guidelines recognize the appropriateness of providing capacity building or other forms of community benefit, especially for population-based studies carried out in resource-poor countries [8, 9, 11, 19, 20], and while Nigeria's own proposed National Code of Health Research Ethics specifically recognizes the appropriateness of such strategies [18], the rules of most funding agencies, including the NIH, do not explicitly recognize 'capacity building' as an allowable cost item.

In the end, the NIH did offer – with approval from the relevant ethics committees and IRBs in Nigeria and the US – the provision of some additional funds to compensate for various tangible cost items that had not earlier been provided for, subject to the receipt of the documentation to support their disbursement. These funds will be released to the chair of the CAG and another named community leader, who will hold them in trust for the community. The community may then choose to use the

funds to contribute toward the desired health center, along with funds being sought from other sources.

The Project was similarly challenged to respond to some issues that arose in the course of the community engagements in China and Japan. In both countries, concerns were expressed about whether the Coriell Institute would provide the communities (through the CAGs) sufficient information on an ongoing basis to enable them to assess whether the samples really were being used in the agreed-upon ways.

In response to these concerns, the Coriell Institute modified its Statement of Research Intent, a form that all investigators who order samples must submit. Investigators who order the samples are now required to include a statement describing their proposed research in terms that lay people can understand. These statements are then provided to the CAGs by the Coriell Institute on a quarterly basis, along with a list of the names and institutions of the investigators who requested the samples. As negotiated with the communities, each quarterly report also includes a listing of all HapMap Project publications, major publications from studies that have used the HapMap and major publications from other studies that have used the community's samples. Periodic newsletters, translated into the languages of all participating communities, are also produced; these include additional information about the Project, the participating communities and how the HapMap is being used. The newsletters are made available both to the CAGs (for further dissemination within each community) and to the public through the Coriell Institute website (<http://coriell.umdj.edu/ccr/hapmap.html>). CAG members are invited to suggest to the Project management and the repository ways to improve the usefulness of the information disseminated.

In further response to concerns initially expressed in China and Japan about sending their samples to a US repository, the Director of the Cell Repositories at the Coriell Institute traveled to meet with CAG members in both countries. A similar visit to Nigeria is being discussed. These visits, which also included Project representatives from the NIH, provided an opportunity for community members to learn about the repository's commitment to serve as a responsible custodian of the samples, and what policies and procedures are in place to do this. These visits also gave the repository and the Project management a chance to listen to, and thus better understand, community members' expressions of hopes for the Project and future research with their samples, as well as lingering concerns.

The Project's Ethical, Legal and Social Implications Group, along with the investigators who collected the samples, continue to explore ways for CAGs in different participating communities to initiate contact with each other (such as through linked websites), and the Coriell Institute has made additional funds available to help support such activities. The hope is that such efforts, over time, will lead to greater transparency and trust, and to the development of a sense of the Project as a truly global enterprise.

Conclusion

Like all genetic variation research, the HapMap Project raises complex ethical, social and cultural issues. We have tried to address some of these issues, in part, through the processes of community engagement and individual informed consent described in this paper. More time must pass before we can fully reflect on the lessons we have learned. New issues may arise that cannot yet be anticipated – especially as the HapMap and the samples begin to be widely used. The CAGs have only recently been formed and their effectiveness has not yet been tested. Community engagement is still underway in other localities with other populations and, as we have noted, the experiences at those sites may be quite different from those described here.

As with our understanding of the science of genetic variation, our understanding of how to responsibly conduct research aimed at the study of individual and group differences is evolving. We do not claim to have found the perfect model for engaging communities, and indeed, we do not believe that a perfect model exists. We also do not think that exercises in public dialogue alone can substitute for care in research design, data analysis and reporting of the findings of genetic variation studies.

Nor do we mean to suggest that a community engagement process as extensive as that used for this Project must always be undertaken when samples are being collected for genetic research with identified populations. However, the experience of this Project does suggest that, at the least, approaching such research in a spirit of openness can improve the presentation and interpretation of the science, help investigators better understand the attitudes and concerns of the communities whose samples they seek to study and, simultaneously, help communities become engaged in the science.

Appendix 1

Community Advisory Groups

The samples collected for the HapMap Project were not only used for the HapMap Project, but will also be used for future genetic variation studies. Such future studies, by building on the HapMap data, will enhance the usefulness of the HapMap itself. However, the Project recognized that members of the participating communities have a legitimate interest in remaining informed about the nature of these future studies in which their samples will be used.

Thus, in accordance with the policy of the repository at the Coriell Institute, a CAG was established at each site where new samples were collected to serve as a liaison between the community and the repository. Each CAG consists of about 6–8 community representatives (who may or may not themselves be sample donors). The CAGs meet periodically on a schedule of their choosing to discuss any matters of interest or concern, such as the status of the Project, how the HapMap and the samples are being used, or developments in genetic variation research generally. Each CAG is kept informed about general Project developments and future studies that use their community's samples through a periodic newsletter (translated into the language of each participating community) and quarterly reports. The CAGs can then disseminate this information within their broader local communities.

The Coriell Institute will work with the CAGs to resolve any concerns about future uses of the samples as they arise. In the unlikely event that a community's samples were to be used in a manner inconsistent with the community's stated wishes, as documented in the consent forms, the community could ask that all of its samples be withdrawn from further distribution and the Coriell Institute would comply with that request.

Appendix 2

Language and Comprehension

Language and comprehension issues were, as predicted, quite formidable. For example, in the Yoruba language, no word exists for the concept of 'genetics'. Although most Yoruba people speak both Yoruba and English and understand the idea of inheritance (e.g. 'diseases passed down in families from the mother and father to their children'), explaining the meaning of 'SNPs' and 'haplotypes' was quite difficult, especially because almost none of the Yoruba people engaged for participation had formal training in genetics or biology. Even in Tokyo and Beijing, where many discussions were held in university settings, and in the CEPH community, where donors are unusually conversant about genetics due to their long history of involvement in other genetic studies, explaining the Project and the principles of genetic variation research in terms that could easily be understood was challenging.

One contributing factor may have been that the Project is not directly related to the study of any particular disease. While most people seem to understand the idea of looking at blood samples to unravel the genetic component of common diseases, it is much harder to comprehend the purpose of creating a general resource that is not immediately related to the study of a specific, named disease, and for which only blood samples, without identifiers or

medical data, are being collected. This experience underscores the importance of developing robust informed consent and engagement or consultation processes when conducting non-disease-specific genetic variation research. Such activities can be informative in modifying recruitment materials and consent documents to ensure that they are accessible to lay and culturally responsive persons.

Acknowledgements

We thank many people who contributed to developing the models for community engagement/public consultation and informed consent used in this Project, and generally to helping us address the ethical, social, and cultural issues raised by the Project. In particular, we thank: J. Greenberg, R. Anderson, J. Beck, D. Coppock, A. Leach, J. Mintzer, and the staff of the Coriell Institute for support for the newsletters, quarterly reports, and

CAGs; C. Chick, E. DeHaut, M. Inaba, E. Jordan, A. Peck, S. Saylor, and J. Witonsky for general support; the people of Tokyo, Japan, the Yoruba people of Ibadan, Nigeria and the community at Beijing Normal University who were generous by participating in public consultations and community engagements; the people in these communities who donated their blood samples; and the people in the Utah CEPH community who allowed the samples they donated earlier to be used for the Project.

This work was supported in part by Genome Canada, Genome Quebec, the Chinese Ministry of Science and Technology, the Chinese Academy of Sciences, the Natural Science Foundation of China, the Hong Kong Innovation and Technology Commission, the University Grants Committee of Hong Kong, the Japanese Ministry of Education, Culture, Sports, Science and Technology, the Wellcome Trust, the SNP Consortium, the US NIH (FIC, NCI, NCRR, NEI, NHGRI, NIA, NIAAA, NIAID, NIAMS, NIBIB, NIDA, NIDCD, NIDCR, NIKKK, NIEHS, NIGMS, NIMH, NINDS, OD), the W.M. Keck Foundation and the Delores Dore Eccles Foundation.

References

- 1 International HapMap Consortium: A haplotype map of the human genome. *Nature* 2005;437:1299–1320.
- 2 International HapMap Consortium: The International HapMap Project. *Nature* 2003; 426:789–796.
- 3 International HapMap Consortium: Integrating ethics and science in the International HapMap Project. *Nat Rev Genet* 2004;5: 467–475.
- 4 Sharp RR, Foster MW: Involving study populations in the review of genetic research. *J Law Med Ethics* 2000;28:41–51, 3.
- 5 State of Utah Census 2000 Data Profile. www.business.utah.edu/bebr/bebrCensus/State/State6.pdf.
- 6 Wikipedia: 'Yoruba'. <http://en.wikipedia.org/wiki/Yoruba> (accessed 18 October 2006).
- 7 CIA World Fact Book: 'China'. <https://www.cia.gov/cia/publications/factbook/geos/ch.html> (accessed 18 October 2006).
- 8 Human Genome Organisation (HUGO): Statement on the principled conduct of genetics research. 1996. <http://www.gene.ucl.ac.uk/hugo/conduct.htm>.
- 9 Council for International Organizations of Medical Sciences (CIOMS): International ethical guidelines for biomedical research involving human subjects. 2002. http://www.cioms.ch/frame_guidelines_nov_2002.htm.
- 10 Deschenes M, Cardinal G, Knoppers BM, Laberge C; Quebec Network of Applied Genetic Medicine (RMGA): Statement of principles on the ethical conduct of human genetic research involving populations. *Res Health* 2003;30:1–4.
- 11 United Nations Educational, Scientific and Cultural Organization (UNESCO): Universal declaration on the human genome and human rights. *J Med Philos* 1998;23:334–341.
- 12 United Nations Educational, Scientific and Cultural Organization (UNESCO): International Declaration on Human Genetic Data. 2003. http://portal.unesco.org/en/ev.php@URL_ID=17720&URL_DO=DO_TOPIC&URL_SECTION=201.html.
- 13 Bankowski Z, Levine RJ (eds): Ethics and Research on Human Subjects: International Guidelines. Proceedings of the XXVth CIOMS Conference. Geneva, CIOMS, 1993, p 63.
- 14 People's Republic of China, Ministry of Science and Technology and the Ministry of Public Health: Interim measures for the administration of human genetic resources. 1998. <http://www.ebnc.org/interim.htm>.
- 15 Human Genome Organisation (HUGO): Statement on DNA sampling: control and access. 1998. <http://www.gene.ucl.ac.uk/hugo/sampling.html>.
- 16 Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT), Council for Science and Technology, Bioethics Committee: Fundamental principles of research on the human genome. 2000. http://www.Mext.go.jp/a_menu/shinkou/shisaku/principles.htm.
- 17 The Japanese Ministry of Education, Culture, Sports, Science and Technology, the Japanese Ministry of Health, Labour and Welfare, the Japanese Ministry of Economy, Trade and Industry: Ethical guidelines for analytical research on the human genome/genes. 2001. <http://www.eubios.info/EGHGR.htm>.
- 18 National Health Research Ethics Committee of Nigeria (NHGREC), Federal Ministry of Health, Department of Health Planning and Research: National Code of Health Research Ethics. 2006. <http://www.nhrec.net/nhrec/code.html>.
- 19 Secretariat of the Convention on Biological Diversity: Bonn Guidelines on Access to Genetic Resources and Fair and Equitable Sharing of the Benefits Arising out of their Utilization. Montreal: Secretariat of the Convention on Biological Diversity, 2002. <http://www.biodiv.org/doc/publications/cbd-bonn-gdls-en.pdf>.
- 20 Human Genome Organisation (HUGO): Statement on benefit sharing. *Genome Digest* 2000;6:7–9.