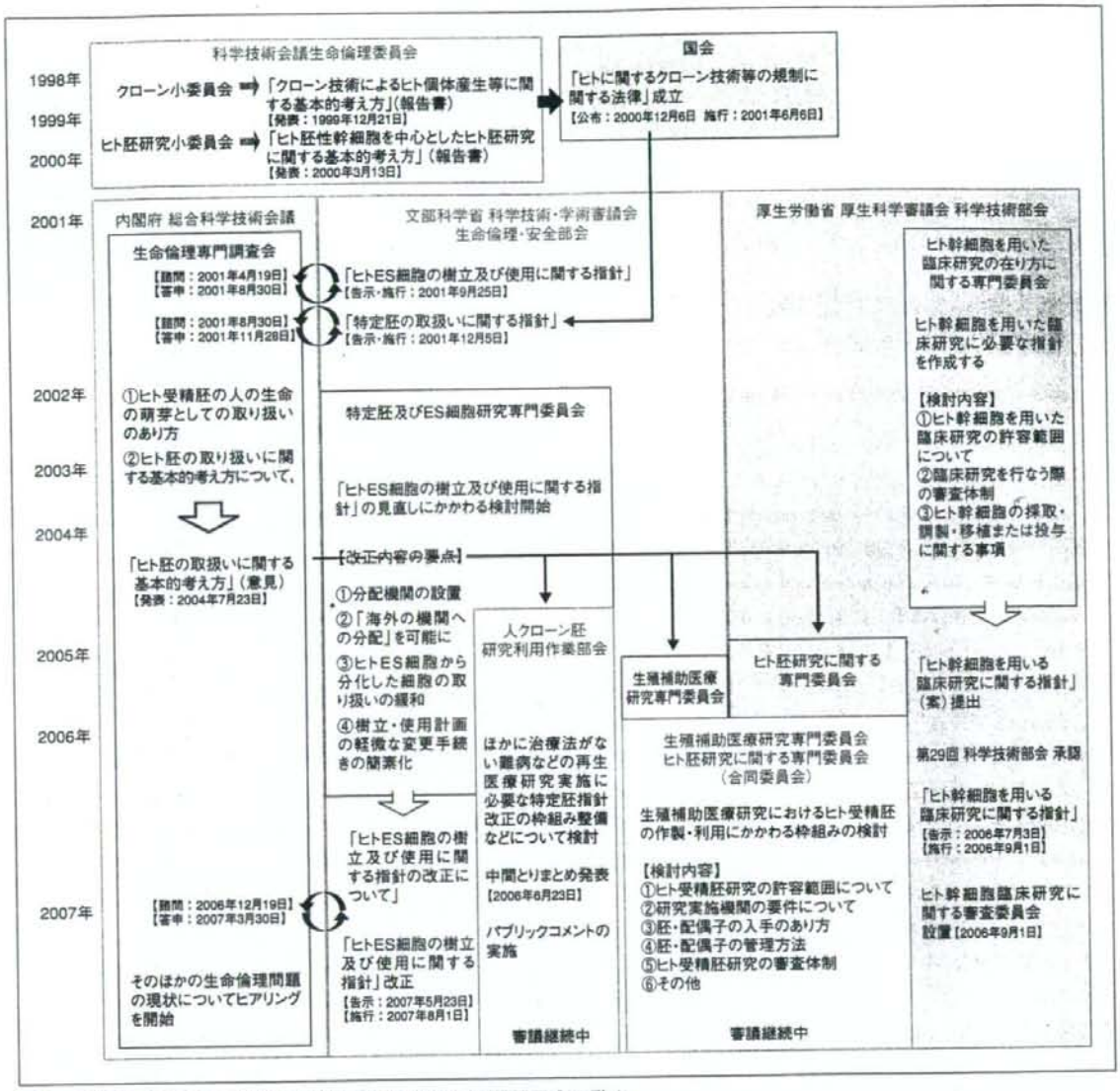

図1　幹細胞研究関連の指針や法律の整備を行なう審議委員会の動き

わが国におけるクローン技術およびヒトの幹細胞研究関連の議論は，科学技術会議生命倫理委員会におかれたクローン小委員会とヒト胚研究小委員会ではじまった（左上）．主としてクローン小委員会の報告をうけるかたちで“ヒトに関するクローン技術等の規制に関する法律”が成立した（中央右上）．2001年の中央省庁改編後は，内閣府 総合科学技術会議の生命倫理専門調査会が調査・検討を行ない，内閣府 総合科学技術会議にて国の方針を決定している（左端）．具体的な指針作成は文部科学省と厚生労働省の専門委員会や作業部会が担当し（文部科学省 ライフサイエンスの広場 生命倫理・安全に対する取組 http://www.lifescience.mext.go.jp/bioethics/index.html，厚生労働省関係審議会議事録など http://www.mhlw.go.jp/shingi/kousei.html#kagaku-hito），必要な場合は 内閣府 総合科学技術会議に諮問し確認する．ヒトES細胞研究に関する指針は2001年に施行され，2007年に改正された（中央）．ES細胞を除くヒト幹細胞の臨床研究に関する指針は厚生労働省の専門委員会で作成され，2006年に施行されている（右下）．また，2004年に総合科学技術会議で決定された“ヒト胚の取扱いに関する基本的考え方”（左端）をうけ，人クローン胚研究利用作業部会によるヒトクローン胚の利用に関する審議と（中央下），文部科学省と厚生労働省の合同委員会による生殖補助医療研究に関する審議が継続中である（中央右下）．図中の報告書や答申，および，意見の日付は，上部組織（科学技術会議生命倫理委員会，内閣府 総合科学技術会議）が承認した日を示す．

のヒトES細胞の譲渡や分化細胞の譲渡・保存を可能にすることで研究者は研究が進めやすくなったといえる．ただし，現状でも細部を指示しすぎであるという声はあり(後述)，たとえば，ヒトES細胞研究に参加する研究者の教育計画(科学面・倫理面)の国への報告など，国内のほかの分野の倫理指針では課せられていない点などは今後の検討に値するのではないだろうか．

3. ヒトクローン胚とヒト胚の取り扱いをめぐる議論

ヒトES細胞研究の指針策定と並行して，当時，新しく登場したばかりの体細胞クローン技術を規制するため"ヒトに関するクローン技術等の規制に関する法律"が策定され，2000年12月に公布された．そのなかで，体細胞クローン技術のヒト個体産生への適用(つまり，クローン人間をつくること)に関しては罰則つきの禁止となったが，治療用のヒトES細胞の作製をめざすヒトクローン胚作製の是非については，国としてのヒト胚に対する考え方をまとめてから指針を整備するとされ，モラトリアム(当面の禁止)となった．

そこで設置されたのが，総合科学技術会議の生命倫理専門調査会である．総計32回，3年におよぶ議論を経て，2004年7月に提出された最終報告書"ヒト胚の取扱いに関する基本的考え方"では，ほかに治療法が存在しない難病に関する再生医療の研究にかぎって，ヒトクローン胚およびヒト受精卵の作製・利用を認めることが示された．それまで，わが国にはヒト胚や配偶子の取り扱いについての統一的な考え方は存在しなかったが，この報告書によってようやく国としての方向性が示されたことになる．

報告書をうけて，大きく2種類の活動がはじまった．まず，文部科学省の特定胚及びES細胞研究専門委員会に人クローン胚研究利用作業部会が設けられ，ヒトクローン胚研究に関する指針の策定をめざした議論がはじまった(**図1**)．未受精卵の提供や無償ボランティアを許容するかどうかを中心に議論が行なわれ，2006年6月に中間とりまとめが発表された．そこでは，未受精卵の入手方法としては手術により摘出された卵巣や卵巣断片からの未受精卵の採取など少数の方法に限定し，無償ボランティアからの未受精卵の提供は当面認めないこととした．ところが，この方法では質の高い未受精卵を得ることは非常にむずかしいこともあり，ヒトES細胞株の樹立に成功した中辻憲夫教授(京都大学再生医科学研究所)は，現状ではヒトクローン胚研究を行なう考えはないと言明している．作業部会では現在，パブリックコメントの結果も取り込み，再検討を行なっている．

もうひとつは，文部科学省と厚生労働省とが合同で行なっている活動で，生殖補助医療研究におけるヒト受精胚の作製・利用に関する枠組みの検討である(**図1**)．どちらかというとヒトクローン胚に関する議論が先行したかたちになってしまった状況に対して，社会的に大きな影響をもつ生殖補助医療研究についての議論が必要だという考えが2省合同の議論を本格的に進める契機となったと思われる．生殖補助医療研究に関しては，政府指針の必要性が長らく指摘されながら議論の詰めができていなかった．今回の総合科学技術会議の専門調査会の報告書はその議論を後押ししたことになり，法律制定も視野に入れた議論が進むことが期待される．

4. 体性幹細胞の臨床応用に関する指針

ヒトES細胞研究やヒトクローン胚研究と異なり，体性幹細胞の研究には後述する胎児細胞の取り扱い以外にはヒト胚や卵子の取り扱いのような突出した倫理問題は存在しない．結果として，基礎研究の段階をこえて患者への幹細胞移植をともなう臨床研究が国内外で数多く進んでいる．しかしながら，わが国の研究には特別なガイドラインは存在せず，機関内倫理審査委員会の審査のみでは科学的根拠が乏しい研究が承認されているという指摘があった．

そこで，厚生労働省の厚生科学審議会にヒト幹細胞を用いた臨床研究の在り方に関する専門委員会が設けられ，4年におよぶ議論を経て"ヒト幹細胞を用いる臨床研究に関する指針"が策定され，2006年9月に施行された．指針においては，①臨床研究としての有効性と安全性の確保，②倫理性の確保，③幹細胞の提供者および患者に対するインフォームドコンセント，④取り扱う幹細胞の品質の確保，⑤情報公開，⑥個人情報の保護，などが重視すべき原則として定められ，機関内倫理審査委員会と国の委員会の二重の審査を行なうこととしている．また，対象とする疾患は重篤で生命を脅かす疾患や身体機能を著し

表1　ヒトES細胞研究に対する世界各国の規制の状況

ヒトES細胞研究に対する態度	国名または地域
ヒトクローン胚研究を含め容認 *1	オーストラリア，シンガポール，中国，韓国，日本，ベルギー，スウェーデン，英国，イスラエル，南アフリカ，など
容認 *2	インド，ニュージーランド，台湾，トルコ，チェコ，デンマーク，エストニア，フィンランド，フランス，ギリシャ，ハンガリー，ラトビア，オランダ，ロシア，スロベニア，スペイン，スイス，ウクライナ，カナダ，メキシコ，ブラジル，など
禁止 *3	オーストリア，ドイツ *5，アイルランド，イタリア，リトアニア，ノルウェー，ポーランド，スロバキア，米国 *6，など
規制なし *4	アジア・オセアニア・ヨーロッパのその他の国，南北アメリカの多くの国，中央アジア・中東・アフリカのほとんどの国

国際幹細胞学会が作成した国際ガイドライン"Stem Cell Policies by Country"などを参考に作成．2007年5月末現在
*1 ヒトクローン胚の作製を含めて，ヒトES細胞の樹立と使用研究を認めている
*2 ヒトES細胞の樹立と使用研究を認めているが，ヒトクローン胚の作製は認めていない
*3 ヒトES細胞研究を基本的に認めていない
*4 ヒトES細胞研究に関する規制が整備されていない
*5 余剰胚から作製されたES細胞の輸入とその使用については厳しい規制下で認めている
*6 大統領令により新規細胞株の樹立に対する公的助成は認めていないが，すでに樹立された余剰胚由来ES細胞を用いた研究には公的助成を認めている．しかし，民間資金による研究には連邦政府の規制はかからない．州法の規制下で研究を認める州もある

く損なう疾患などに限定し，かつ，被験者にとってヒト幹細胞臨床研究の治療により得られる利益が不利益をうわまわると十分に予想されるものにかぎることが明記されている．この指針が施行されたことにより，被験者の利益とならない安易な研究の実施を防ぐ体制ができたと期待される．

今回の指針作成の倫理的問題として大きな議論になったのは，死亡胎児由来の幹細胞の使用を認めるかどうかであった．委員会ではさまざまな立場からの意見聴取や長期間にわたる議論を重ねたが，結局のところ合意には至らず，胎児由来の幹細胞は今回の指針の対象外とすることになった．結果として，胎児由来の幹細胞を用いた研究には指針が存在しない状態がつづいている．専門委員会では，胎児由来の細胞の取り扱いは重要な問題であるので，なんらかのかたちで議論を継続させ早急に結論を出すべきという意見もあったが，現在のところそのような組織はつくられていないようである．これまでの議論を無駄にしないためにも，早急に再検討を開始すべきではないだろうか．

II 国際幹細胞学会によるヒトES細胞研究のガイドライン

ヒトES細胞研究およびヒトクローン胚の研究に関しては，各国で規制の状況が大きく異なっている（**表1**）．たとえば，オーストリア，イタリア，ノルウェー，ポーランドなどでは，クローン胚研究はもちろん，ヒトES細胞研究も禁止されている．一方，インド，台湾，フランス，スペイン，スイス，カナダなどは，生殖補助医療の余剰胚にかぎって研究への利用を認めており，ヒトES細胞の研究が進められている（ヒトクローン胚研究は禁止）．また，オーストラリア，中国，韓国，シンガポールや英国では，ヒトES細胞研究に加えてヒトクローン胚研究についても認められている．

ヒトの胚を研究に用いてよいかどうかという問題は，さまざまな文化的・歴史的・宗教的背景をもつ国ごとに異なる対応がなされる問題であり，世界共通の考え方をつくることはおそらく不可能である．生命科学研究が人間社会のもつ文化的価値観・社会的価値観とぶつかり，ときに規制がなされるような時代になっていることは，好むと好まざるとにかかわらず認めざるをえない現実である．

このように国ごとに対応が異なる現実のなか，少なくともヒトES細胞研究が行なわれている国々においては共通の指針をもつことが望ましいのではないか．そのような指針は国際共同研究の促進に役立つばかりでなく，今後，新たにヒトES細胞研究を行なおうとする国が自国の指針をつくるときのガイドとなるだろう．そうした考え

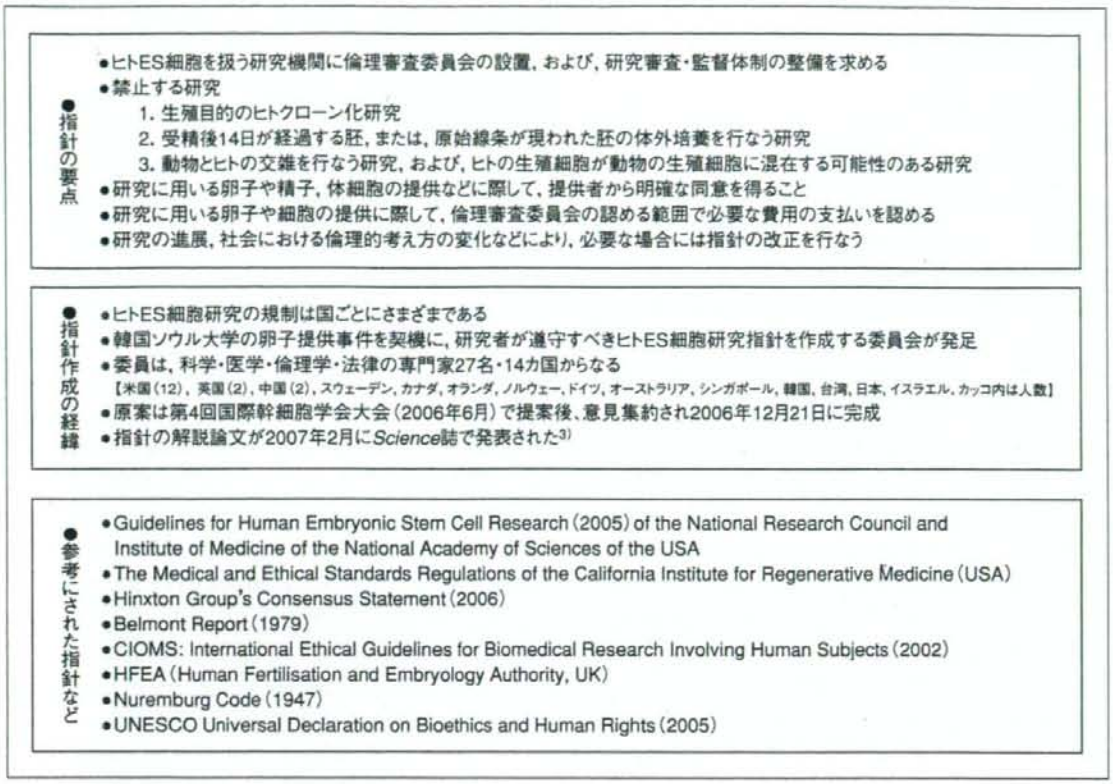

図2　国際幹細胞学会のヒトES細胞研究国際指針

のもとに国際ガイドラインを作成したのが，幹細胞研究の情報交換促進を目的に2002年に設立された国際幹細胞学会(International Society for Stem Cell Research：ISSCR，**http://www.isscr.org/**)である．

国際幹細胞学会は，2005年11月に世界14カ国の幹細胞研究者や生命倫理学者27名からなる委員会を発足させ，指針策定を開始した．委員会での議論をもとに作成された指針案は，2006年6月の学会大会で学会会員に提示され意見の募集が行なわれた．これらの意見を取り入れて完成した指針は2007年2月に公表され(**http://www.isscr.org/guidelines/index.htm**)，骨子が*Science*誌に発表されている[3]．

指針でポイントとなった点を，**図2**にまとめた．まず，ヒトES細胞研究を行なう際の研究審査体制を整備することの重要性が示されている．審査制度そのものを各国が整備することと，それにもとづき研究機関のなかに設置された倫理審査委員会が研究の審査を行なうことを求めている．また，研究を行なうグループだけでなく，専門雑誌の編集者や資金援助団体，細胞バンクや分配にたずさわる機関にも指針の遵守を求めている．

研究そのものについては，禁止すべき研究という形で，①生殖目的のヒトのクローン化の研究，②受精後14日を過ぎた胚，または，原始線条(primitive streak)が出現した胚の体外培養をともなう研究，③動物とヒトの交雑を行なう研究およびヒトの生殖細胞が動物の生殖細胞に混ざる可能性のある研究，が示されている．さらに，研究に使用する卵子や精子など配偶子の提供者や，核移植をともなう研究におけるドナーの体細胞提供者などにはインフォームドコンセントを得る必要があること，提供者に対して倫理審査委員会が認める適切な範囲で必要な費

用の支払いを認めること，などが示されている．

国際幹細胞学会の指針は，幹細胞研究を行なう研究者たちがとりまとめた学会レベルの指針であり，各国の法律あるいは国際条約のような拘束力はない．しかし，もし国際幹細胞学会がよびかけている専門雑誌や資金援助団体，細胞の分配にかかわる機関などが，関係の研究者に対して指針の遵守を条件に論文の投稿や細胞の分配を認める決定を下せば，実質的な影響は非常に大きくなる．国境をこえた科学研究の適切な実施を保障するしくみのひとつは国連による国際条約の策定であるが，もうひとつ，世界各国の研究者の集まりである国際学会の自主的な指針が実は大きな波及効果をもつ可能性があることは，もっと意識されてもいいだろう．その意味で，学会に参加する研究者には，自国の指針や議論を理解するだけでなく，日ごろから国際的な議論の状況を知り，自らの考えを深めておくことが期待される．

また，国際幹細胞学会による指針策定が行なわれたことにより，すでに指針をもつ国でも自国の指針の再検討が行われる可能性がある．たとえば，国際幹細胞学会の指針策定にわが国から委員として参加した中辻教授は，国際的基準に比べわが国の指針が厳格すぎるため研究が遅れる原因になっているとさまざまな場面で主張している[4]．少なくとも，今回の国際幹細胞学会の指針を含む諸外国の指針を検討し，わが国の指針のどの部分が厳密すぎてどの部分が世界標準に合っているかを再検討することは必要だと，筆者らも考える．

Ⅲ 社会全体で考えるために

ここ数年の幹細胞研究の変化は驚くほど激しい．本稿を準備しているあいだにも，胎仔の線維芽細胞に遺伝子導入することでES細胞とほとんど同程度の全能性をもつ細胞をつくることにマウスを使って成功したという論文が，山中伸弥教授（京都大学再生医科学研究所）のグループを含む複数の研究グループから発表された[5-7]．同様のことがヒト細胞でも可能になるのであれば（ただし，胎児由来の細胞を使用することの倫理的問題や移植した細胞の腫瘍化の問題など，さまざまな課題をクリアすることは必要である），ES細胞でもない，体性幹細胞でもない，新しいカテゴリーの幹細胞を用いた再生医療が一挙に現実味を帯びる可能性がある．その際に，既存の指針をどのように改定するのが適切なのかについて，いまから検討をはじめる必要がある．

これほどに変化の激しい幹細胞研究のあり方を社会全体で検討するためには，現在のわが国にはない新しい活動（あるいは，既存の活動の活性化）が必要ではないだろうか．つぎに，これについて筆者らの考えを述べてみたい．

ひとつには，幹細胞研究そのものの現状やガイドラインなどの整備の現状を，広く世界を見渡してまとめる活動の重要性である．第Ⅰ部で述べた文部科学省や厚生労働省などにおける審議会では，その都度，必要に応じて情報収集がなされるが，指針の策定などが終わるとその収集作業は止まってしまう．やはり，大学や研究機関などに継続的に情報を収集するための組織が必要である．そうした活動は海外では進んでおり，たとえば，カナダのモントリオール大学にあるHumGenという研究センターでは，ヒトゲノム研究やヒト幹細胞研究の倫理的・法的・社会的課題に関する文献を収集してWeb上で公開している（http://www.humgen.umontreal.ca/int/）．科学研究についての情報収集・蓄積についても同様の活動が望まれる．

つぎに重要だと考えるのは，マスメディアの活動の活性化とレベルアップである．ヒトES細胞研究に関しては，ヒトES細胞の樹立計画が国の委員会で審査されていたころにはしばしば新聞の一面にも大きな記事が出ていたが，最近ではめだった記事はほとんどなくなった．だが，臨床医療への応用がいろいろな方向からみえてきたいまこそ，幹細胞研究の進め方を社会全体で真剣に議論する必要がある．話題性や事件性のある内容を派手にとりあげるだけではなく，研究の発展の状況を冷静に分析したり，自国と諸外国の状況を比較して，わが国になにが必要なのかを深く掘り下げたりするような新聞記事やテレビ番組がもっと増えることが望まれる．

そして，最後に指摘したいのは，研究者自身による情報発信の重要性である．近年，幹細胞研究は再生医療とのかかわりばかりで語られる傾向にあるが，基礎生物学としても重要な分野である．細胞核の全能性を示す実験を行なった発生学者シュペーマンによる研究をはじめ[8]，

百年以上の歴史をもつ細胞分化の研究は現代の幹細胞研究につながっている．現在のわが国でも，細胞分化の研究はプラナリア，線虫，ゼブラフィッシュ，メダカなどの生物を使って活発に行なわれている．多様な生物を使った研究発表や歴史の紹介なども織り交ぜて，幹細胞研究の現状や意義を研究者のコミュニティーから社会にダイレクトに発信することはできないだろうか．

筆者らが属する文部科学省 科学研究費特定領域研究ゲノム4領域では，100人規模の研究者が街に出て実物つき研究展示を用いながら一般市民を含む非専門家と交流する"ゲノムひろば"という催しを2002年度から継続的に実施している[9]．そこでは，研究者が非専門家や市民に研究を伝えるだけでなく，基礎生物学・医学から情報科学分野までのゲノム研究者と，政策・教育担当者などを含むさまざまな分野の専門家どうしが交流し，自分たちの研究の社会的意義をとらえ直す場としても機能している．さらには，どのような形で研究の情報を社会に発信すればよいかを研究者が学ぶ場にもなっている．同様の催しを"幹細胞ひろば"あるいは"再生医学研究ひろば"というような名称で実施することで，研究情報の発信と専門家どうしの交流の両方が実現できるだろう．このような双方向の対話，異分野の交流などを含めた活動を"科学コミュニケーション"とよぶが，社会的関心の高い幹細胞研究を対象にすれば，さまざまな形態の活動が実施できるのではないだろうか．

おわりに

幹細胞研究の社会的側面を，政府や国際学会によるガイドラインの策定，社会とのコミュニケーションなど，複数の視点からみてきた．そこであらためて気づくのは，この分野が基礎研究，医療，行政，倫理，安全といった社会における実に広い分野とのかかわりをもつことである．それぞれの専門家が分野をこえて協力すると同時に，社会との対話に取り組み，本当の意味で社会に役立つ幹細胞研究とそれを用いた医療のシステムをつくりあげていくことが期待されている．

文　献

1) Takahashi, K., Yamanaka, S.: *Cell*, **126**, 663-676 (2006)
2) Thomson, J. A. *et al.*: *Science*, **282**, 1145-1147 (1998)
3) Daley, G. Q. *et al.*: *Science*, **315**, 603-604 (2007)
4) Cyranoski, D.: *Nature*, **438**, 263 (2005)
5) Okita, K., Ichisaka, T., Yamanaka, S.: *Nature*, doi: 10.1038/nature05934 (2007)
6) Wernig, M. et al.: *Nature*, doi: 10.1038/nature05944 (2007)
7) Maherali, N. *et al.*: *Cell Stem Cell*, **1**, 55-70 (2007)
8) Hamburger, V.: The Heritage of Experimental Embryology: Hans Spemann and the Organizer, Oxford University Press, Oxford (1988)
9) 加藤和人: 蛋白質 核酸 酵素, **50**, 2302-2307 (2005)

加藤和人

略歴：1989年 京都大学大学院理学研究科 修了，理学博士，英国Cambridge大学，生命誌研究館を経て，2001年 京都大学人文科学研究所 助教授(2007年より准教授)，2004年より同大学院生命科学研究科 兼任．科学研究を理解しつつ社会との接点分野の仕事ができる専門家が育つことを願い，十数名の大学院生・ポスドクたちと毎日奮闘中．

川上雅弘

略歴：2006年 近畿大学大学院農学研究科博士後期課程 修了，農学博士，同年より京都大学人文科学研究所 研究員．

研究テーマ：胚操作技術および幹細胞研究を中心とした先端科学と社会に関する研究．

シリーズ「幹細胞技術の現状と展望」は今回をもちまして終了となります．長いあいだのご愛読ありがとうございました．
ご意見・ご感想をE-mail (pne@kyoritsu-pub.co.jp) にてお寄せいただけましたら幸いです．　(編集部)

Case Report

Community Genetics

Community Genet 2007;10:186–198
DOI: 10.1159/000101761

Community Engagement and Informed Consent in the International HapMap Project

Charles Rotimi[a] Mark Leppert[b] Ichiro Matsuda[e] Changqing Zeng[h] Houcan Zhang[i] 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

[a]National Human Genome Center, Howard University, Washington, D.C., [b]Department of Human Genetics, University of Utah, Salt Lake City, Utah, [c]National Human Genome Research Institute, US National Institutes of Health, Bethesda, Md., and [d]Department of Bioethics, Case Western Reserve University School of Medicine, Cleveland, Ohio, USA; [e]Health Sciences University of Hokkaido, Hokkaido, [f]Department of Medical Genetics, Shinshu University School of Medicine, Matsumoto, and [g]University of Tsukuba, Eubios Ethics Institute, Tsukuba Science City, Japan; [h]Beijing Genomics Institute, Chinese Academy of Sciences, and [i]Beijing Normal University, Beijing, China; [j]University of Ibadan College of Medicine, Ibadan, Oyo State, Nigeria; [k]United 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.

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

1422–2795/07/0103–0186$23.50/0

Accessible online at:
www.karger.com/cmg

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

Vol 449|18 October 2007|doi:10.1038/nature06258 nature

ARTICLES

A second generation human haplotype map of over 3.1 million SNPs

The International HapMap Consortium*

We describe the Phase II HapMap, which characterizes over 3.1 million human single nucleotide polymorphisms (SNPs) genotyped in 270 individuals from four geographically diverse populations and includes 25–35% of common SNP variation in the populations surveyed. The map is estimated to capture untyped common variation with an average maximum r^2 of between 0.9 and 0.96 depending on population. We demonstrate that the current generation of commercial genome-wide genotyping products captures common Phase II SNPs with an average maximum r^2 of up to 0.8 in African and up to 0.95 in non-African populations, and that potential gains in power in association studies can be obtained through imputation. These data also reveal novel aspects of the structure of linkage disequilibrium. We show that 10–30% of pairs of individuals within a population share at least one region of extended genetic identity arising from recent ancestry and that up to 1% of all common variants are untaggable, primarily because they lie within recombination hotspots. We show that recombination rates vary systematically around genes and between genes of different function. Finally, we demonstrate increased differentiation at non-synonymous, compared to synonymous, SNPs, resulting from systematic differences in the strength or efficacy of natural selection between populations.

Advances made possible by the Phase I haplotype map

The International HapMap Project was launched in 2002 with the aim of providing a public resource to accelerate medical genetic research. The objective was to genotype at least one common SNP every 5 kilobases (kb) across the euchromatic portion of the genome in 270 individuals from four geographically diverse populations[1,2]: 30 mother–father–adult child trios from the Yoruba in Ibadan, Nigeria (abbreviated YRI); 30 trios of northern and western European ancestry living in Utah from the Centre d'Etude du Polymorphisme Humain (CEPH) collection (CEU); 45 unrelated Han Chinese individuals in Beijing, China (CHB); and 45 unrelated Japanese individuals in Tokyo, Japan (JPT). The YRI samples and the CEU samples each form an analysis panel; the CHB and JPT samples together form an analysis panel. Approximately 1.3 million SNPs were genotyped in Phase I of the project, and a description of this resource was published in 2005 (ref. 3).

The initial HapMap Project data had a central role in the development of methods for the design and analysis of genome-wide association studies. These advances, alongside the release of commercial platforms for performing economically viable genome-wide genotyping, have led to a new phase in human medical genetics. Already, large-scale studies have identified novel loci involved in multiple complex diseases[4,5]. In addition, the HapMap data have led to novel insights into the distribution and causes of recombination hotspots[3,6], the prevalence of structural variation[7,8] and the identity of genes that have experienced recent adaptive evolution[3,9]. Because the HapMap cell lines are publicly available, many groups have been able to integrate their own experimental data with the genome-wide SNP data to gain new insight into copy-number variation[10], the relationship between classical human leukocyte antigen (HLA) types and SNP variation[11], and heritable influences on gene expression[12–14]. The ability to combine genome-wide data on such diverse aspects of genetic variation with molecular phenotypes collected in the same samples provides a powerful framework to study the connection of DNA sequence to function.

*Lists of participants and affiliations appear at the end of the paper.

In Phase II of the HapMap Project, a further 2.1 million SNPs were successfully genotyped on the same individuals. The resulting HapMap has an SNP density of approximately one per kilobase and is estimated to contain approximately 25–35% of all the 9–10 million common SNPs (minor allele frequency (MAF) ≥ 0.05) in the assembled human genome (that is, excluding gaps in the reference sequence alignment; see Supplementary Text 1), although this number shows extensive local variation. This paper describes the Phase II resource, its implications for genome-wide association studies and additional insights into the fine-scale structure of linkage disequilibrium, recombination and natural selection.

Construction of the Phase II HapMap

Most of the additional genotype data for the Phase II HapMap were obtained using the Perlegen amplicon-based platform[15]. Briefly, this platform uses custom oligonucleotide arrays to type SNPs in DNA segmentally amplified via long-range polymerase chain reaction (PCR). Genotyping was attempted at 4,373,926 distinct SNPs, which corresponds, with exceptions (see Methods), to nearly all SNPs in dbSNP release 122 for which an assay could be designed. Additional submissions were included from the Affymetrix GeneChip Mapping Array 500K set, the Illumina HumanHap100 and HumanHap300 SNP assays, a set of ~11,000 non-synonymous SNPs genotyped by Affymetrix (ParAllele) and a set of ~4,500 SNPs within the extended major histocompatibility complex (MHC)[11]. Genotype submissions were subjected to the same quality control (QC) filters as described previously (see Methods) and mapped to NCBI build 35 (University of California at Santa Cruz (UCSC) hg17) of the human genome. The re-mapping of SNPs from Phase I of the project identified 21,177 SNPs that had an ambiguous position or some other feature indicative of low reliability; these are not included in the filtered Phase II data release. All genotype data are available from the HapMap Data Coordination Center (http://www.hapmap.org) and dbSNP (http://www.ncbi.nlm.nih.gov/SNP); analyses described in this paper refer to release 21a. Three data sets are available: 'redundant unfiltered'