

# Data sharing and intellectual property in a genomic epidemiology network: policies for large-scale research collaboration

Dave A Chokshi,<sup>a</sup> Michael Parker,<sup>b</sup> & Dominic P Kwiatkowski<sup>a</sup>

**Abstract** Genomic epidemiology is a field of research that seeks to improve the prevention and management of common diseases through an understanding of their molecular origins. It involves studying thousands of individuals, often from different populations, with exacting techniques. The scale and complexity of such research has required the formation of research consortia. Members of these consortia need to agree on policies for managing shared resources and handling genetic data. Here we consider data-sharing and intellectual property policies for an international research consortium working on the genomic epidemiology of malaria. We outline specific guidelines governing how samples and data are transferred among its members; how results are released into the public domain; when to seek protection for intellectual property; and how intellectual property should be managed. We outline some pragmatic solutions founded on the basic principles of promoting innovation and access.

Bulletin of the World Health Organization 2006;84:382-387.

Voir page 386 le résumé en français. En la página 386 figura un resumen en español.

يمكن الاطلاع على الملخص بالعربية في صفحة 387.

## Introduction

One of the most important discoveries of genome sequencing projects is the extent of genomic diversity in humans<sup>1</sup> and in human pathogens.<sup>2,3</sup> We now have many of the tools required for genomic epidemiology — the systematic investigation of how natural genomic variation affects the clinical outcome of disease. Infectious diseases are a central focus for genomic epidemiology, because pathogens are a major force for evolutionary selection of the human genome, and pathogen genomes are continually evolving to counter adaptations in the human immune system, and to survive the drugs used against them. Genomic epidemiology has important practical applications for diseases of the developing world, particularly in tackling drug resistance and guiding vaccine development.

Although genomic epidemiology operates by the same fundamental principles as other forms of genetic research, the scale of research projects is much larger. Studies are currently under way that involve testing over half a million genetic variants (known as single nucleotide polymorphisms or SNPs) in thousands of individuals with different diseases, and in healthy people, to identify

those regions of the genome that are associated with resistance or susceptibility to those diseases. It is because the whole genome is being investigated that a large number of SNPs need to be tested. The need to detect effects of modest magnitude necessitates a very large number of subjects for a given study.<sup>4</sup>

Thus large-scale epidemiological studies, often conducted in multiple populations, need to be combined with high-throughput genome technologies and advanced statistical computations. The consequent increase in scale often requires the formation of research consortia for investigations in genomic epidemiology.

This paper details issues arising during the formation of an international research consortium known as MalariaGEN (Malaria Genomic Epidemiology Network; [www.MalariaGEN.net](http://www.MalariaGEN.net)), whose aim is to use genomic epidemiology to identify molecular mechanisms of protective immunity against malaria — and thereby to guide malaria vaccine development. The consortium, which is funded through the Grand Challenges in Global Health initiative<sup>5</sup> brings together clinical researchers, epidemiologists, immunologists, genome researchers and statisticians from 20

countries in Africa, Asia, Europe and North America. The consortium is funded to analyse DNA and clinical data from tens of thousands of individuals, generating billions of genetic data points. This large undertaking raises many ethical issues which have been summarized elsewhere, particularly relating to consent, genetic database governance and the fact that the research involves communities in the developing world.<sup>6</sup> Here we focus specifically on the questions of data sharing and how intellectual property may be managed for the greatest return to society.

## Innovation and access

The issues of data-sharing and intellectual property are closely connected. For example, if a research consortium decides to release all data immediately into the public domain, this precludes the possibility of patenting discoveries. If there are good reasons for immediately releasing data into the public domain, but also good reasons for patenting discoveries, then we need policy guidelines to determine which option should take precedence in a given situation.

We propose two fundamental principles upon which to base policy

<sup>a</sup> Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, England. Correspondence to: Dave Chokshi (email: [daveash@med.upenn.edu](mailto:daveash@med.upenn.edu)).

<sup>b</sup> Ethox Centre, University of Oxford, Oxford, England.

Ref. No. 06-029843

(Submitted: 10 January 2006 – Final revised version received: 18 March 2006 – Accepted: 20 March 2006)

Dave A Chokshi et al.

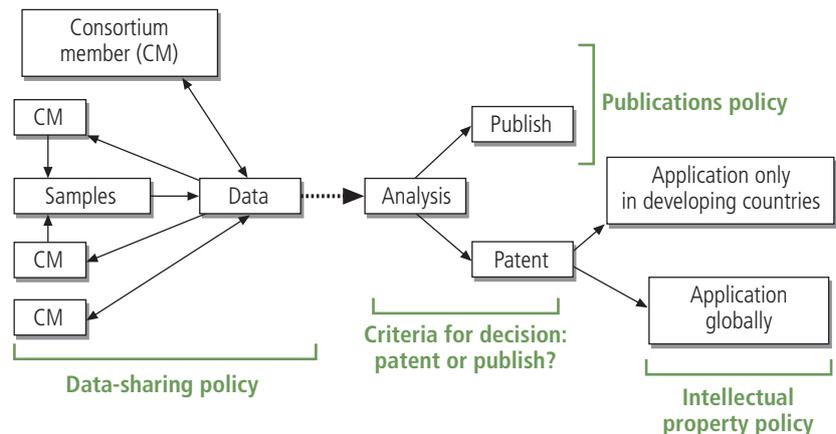
decisions about data sharing and intellectual property: (1) impediments to innovation in research processes should be minimized, and (2) the fruits of research — eventual products that result from scientific discoveries — should be made as widely accessible as possible, particularly to the people who need them the most.

In the context of genomic epidemiology, fostering innovation involves two broad goals. The first goal is to accelerate basic scientific research by making data accessible to the researchers best able to build upon promising findings. For example, primary data can be disseminated through web databases and peer-reviewed publications in such a way as to facilitate the aggregation of data sets where this is scientifically important. Of course, the release of data depends on the consent of the communities being studied and the approval of relevant ethical bodies, as there are ethical issues surrounding clinical and genetic data that must be taken into consideration. The second goal in fostering innovation is to translate the most promising scientific findings into therapeutic end products. Intellectual property management may be an important tool that a research consortium can use to maximize the chances of translating research findings into diagnostics, pharmaceuticals or vaccines.

Ensuring access, which means that therapeutic end-products are readily available to those who need them most, is an equally important responsibility for large research partnerships. It is also an area where proactive management of intellectual property may be a useful tool. For example, it has been shown that possessing a patent that is “upstream” in the drug development process gives the patent holder some influence when problems of access come to the fore.<sup>7</sup> The advent of non-profit pharmaceutical operations and public-private partnerships (e.g., the Medicines for Malaria Venture and the Institute for OneWorld Health) present opportunities to ensure that downstream development of basic research is focused on making end products available to those with greatest need.

The information flow of a research consortium, indicating the stages where policy guidelines are needed, is shown in Fig. 1. First, consortium members must establish data-sharing relationships by defining how samples and primary data

Fig. 1. Information flow and policy decisions in a research collaboration



WHO 06.43

are transferred among them. After data have been analysed, the consortium must have a procedure to determine how the analyses will be released into the public domain. This depends on guidelines about when to seek intellectual property protection, and if this is done, consensus policies must exist for how intellectual property is managed. Policy governing publications resulting from large-scale research collaboration is also important but is not discussed here because it has been explored elsewhere.<sup>8,9</sup>

### Data-sharing policy

#### Distinguishing between collective and individual resources

A data-sharing policy must begin by demarcating resources — samples, data, and infrastructure — that will be shared across the consortium and those that remain in the domain of individual researchers. For example, in the MalariaGEN consortium, most research groups contribute: (1) DNA from subjects with malaria and population controls, and (2) a limited set of standardized phenotypic information (clinical data) associated with the samples. However this does not include the full clinical database, which remains under the control of the research groups who collected the data. Thus a clear distinction may be made between “consortium experiments”, which analyze genetic data and a limited set of standardized phenotypic data across the whole consortium, and “investigator-initiated analyses”, in which individual research groups utilize all the clinical and epidemiological information that they have collected

together with any genetic data that have been generated on those samples in the consortium experiments.

#### Access to samples

We propose the principle that an investigator who contributes samples — here called a contributing investigator — retains full access to those samples and is responsible for ensuring that the access conditions of the consortium are consistent with those laid down by their institutional review board. Other investigators are not granted access except for the specific purpose of performing agreed consortium experiments. Samples can only be transferred to other repositories by the contributing investigator. It is likely that there will be exceptions, such as when investigators agree to share samples for projects other than consortium experiments, but these must be approved by the relevant contributing investigator and their institutional review board. A contributing investigator may withdraw samples and is obliged to do so if a research subject asks to leave the study.

#### Access to information within the consortium

First, primary genotyping data must be transmitted back to the contributing investigator who provided the samples. In MalariaGEN, it is proposed that contributing investigators should be able to access genotyping data on their samples as soon as it is generated via a web-based system. If two contributing investigators wish to collaborate on a specific project, they may use this web-based system to

share genotyping data, and as they each retain control of their own clinical databases, they are in a position to decide which clinical data to share.

Second, a decision has to be made about when and how to release genotyping data across the consortium. Such decisions are simplified if there is prior agreement about the level of data sharing for consortium experiments. There are now two separate questions: (1) at what point can this group of investigators see one another's data, and (2) at what point can the whole consortium see the whole dataset? The answers should generally be (1) immediately and (2) soon. A consortium may decide that there are reasons for delaying (2), hopefully by no more than a few months, to allow researchers who have invested time and effort in collecting samples and clinical data to have the first opportunity to analyse resulting data.

### Access to information outside the consortium

It is beyond the scope of this paper to deal with the general issue of protection of anonymity for research subjects, except to say that this is of paramount importance. A specific issue for genomic epidemiology is that genetic data may, in certain circumstances, indirectly identify individuals within a well-defined study population. Thus researchers and ethical committees need to weigh up the risks and benefits of different levels of personal genetic identification. For example, there is a difference in risk between releasing large amounts of genetic data for each individual within a small village that is identified and releasing the same data for subjects sampled randomly from a large population, even if both groups are fully anonymised. One way of reducing any potential risk to individuals is to publicly release only pooled data.

Once ethical issues have been addressed, there are two major additional considerations. Data release may have to be delayed if there are compelling reasons for intellectual property protection. Furthermore, it may be argued that putting consortium data into the public domain could undermine the proper assignment of research credit to the major contributors.

A consortium may deal with such considerations by regulating the nature and amount of data release together with the timing of data release. For example, if it is not possible to release all the data

immediately on an open web site, then it may at least be possible to release some of the data immediately on a controlled web site. If genetic data are released without any phenotypic data, this may be of immediate scientific value (e.g. in determining the population frequency of known disease susceptibility alleles) and a strong case has been made against patenting data of this sort.<sup>10</sup> This does not preclude the investigators from seeking patent protection for an innovative idea based on a novel disease association, which requires knowledge of both the genetic and the phenotypic data.

A consortium may also regulate the method of data access. For example, the Wellcome Trust Case Control Consortium (<http://www.wtccc.org.uk>) proposes to make certain data freely available to researchers, but only after their "bona fide" status has been approved by a Consortium Data Access Committee (CDAC), and they have entered into a data access agreement. This may stipulate, for example, that the data are to be used only for the advancement of medical research, according to the consent obtained from the research subjects; that the confidentiality and privacy of subjects must not be in any way compromised; and that appropriate acknowledgement or credit must be given to the study populations and investigators who provided the samples and data.<sup>11</sup>

### Investigator-initiated analyses

An example of an investigator-initiated analysis is when an investigator performs a detailed analysis of locally generated data, such as clinical and immunological measurements, combined with genetic data from those subjects as part of a consortium experiment. It is important that clinical and epidemiological investigators are empowered and encouraged to carry out such analyses, which are a crucial component of translating basic research into practical applications. A consortium may lay down guidelines specifying, for example, that investigator-initiated publications should acknowledge the use of consortium data and should not pre-empt ongoing consortium experiments. To foster collaboration, the network could help investigators to post reports online of research in progress as a way to encourage multiple research groups to pool their clinical data for a particular project.

### Criteria for deciding whether to patent

A patent is essentially an agreement whereby the inventor of a technology discloses knowledge for the advancement of society in exchange for a limited period of exclusivity over that technology. Therefore, research consortia must arrive at a policy on patenting prior to the stage where data are released into the public domain. Once research findings have been published, whether this is on the web, in peer-reviewed journals or at scientific meetings, they cannot be patented in most countries, although the United States is an exception. Once a patent has been filed, a research consortium may publish research results in the normal manner.

Here we propose three criteria that a research consortium may use in deciding whether to file for a patent:

- the innovation must be directly relevant to a health application (i.e. a diagnostic test, drug, or vaccine);
- it must be highly likely that the innovation will immediately be licensed for further development;
- it must be clear that intellectual property protection is either (1) *required* as a stimulus for further development or (2) a useful tool for negotiating global access.

The first two conditions require the innovation to have real potential to be translated into an end product. This is necessary to ensure that exclusive rights are exercised with caution, because excessive patenting acts to delay the release of information and may therefore impede scientific innovation.<sup>12</sup> The third condition sets out *positive* criteria that an innovation must meet to warrant intellectual property protection: essentially it is a means to an end, where the ultimate goal is to make a product that will improve health.

To illustrate how these criteria would work, consider some textbook examples of genetic discoveries in malaria. Thirty years ago it was discovered that individuals who do not express Duffy antigen receptor on erythrocytes are completely protected from infection by *Plasmodium vivax*.<sup>13</sup> Although this was simply a genetic observation, it was clearly of great relevance to understanding the molecular mechanism of infection, and it eventually led to discovery of the *P. vivax* Duffy Binding Protein (PvDBP) which binds to Duffy antigen

Dave A Chokshi et al.

and is essential for erythrocyte invasion by this species of parasite.<sup>14</sup> The interaction between PvDBP and Duffy antigen provides a potentially important target for vaccine and drug developers; a candidate vaccine against PvDBP is currently under experimental investigation.<sup>15,16</sup> What should be the appropriate course of action for a research consortium if a similar genetic discovery were made for *P. falciparum*, which kills over a million African children each year? Should it be patented? Evaluated against the first two criteria, it is clearly relevant to a health application and, depending on the exact nature of the discovery, it might point clearly to a process by which a vaccine could be developed. The decision to patent should then be based on the third criterion: whether patent protection is necessary to stimulate further development or to negotiate global access.

It is unlikely that more than a small fraction of the discoveries arising from genomic epidemiology will meet these criteria for patenting, because it is only exceptionally that a single genetic finding will immediately suggest a set of operations leading to a specific end-product. Genetic findings are of practical value, but this is mostly because they identify molecular pathways that are involved in disease or protection, thus helping to focus research efforts on the most important areas for drug and vaccine development. Consider the two classical polymorphisms in the gene encoding  $\beta$ -globin, an important component of haemoglobin. One of these variants, sickle haemoglobin (HbS), has undergone selection in many African populations because it confers approximately 10-fold resistance against life-threatening forms of malaria in heterozygotes, although homozygotes suffer from sickle cell disease which is often fatal.<sup>17</sup> Another variant, haemoglobin C (HbC), does not cause major problems in homozygotes, and confers resistance against malaria in both heterozygotes and homozygotes.<sup>18</sup> It is potentially of considerable relevance for vaccine development to know exactly how HbS and HbC protect against malaria and how this is related to acquired immunity.<sup>19,20</sup> This is a situation where genetic findings have led to an area of clinical and scientific investigation with important practical implications, but it would be inappropriate to develop patents based on discoveries of this sort unless they pointed to a specific product.

If patenting is a possibility, then principles need to be established governing royalties. We propose that these should flow back to the communities participating in research. One way of achieving this, proposed by the Human Genome Organization Ethics Committee, is that pharmaceutical industries should set aside a certain proportion of their income for health-care development or as broad humanitarian assistance for developing countries.<sup>21</sup> This funding mechanism avoids the difficulties, both logistical and ethical, of tracing research participants after a number of years have elapsed. Another way of formulating this principle, which highlights the need to prevent exploitation, is that patent royalties that arise directly from genomic epidemiology studies of communities should not enrich the researchers or their institutions.

### Intellectual property policy Guidelines for intellectual property management

The Grand Challenges in Global Health initiative ([www.grandchallengesgh.org](http://www.grandchallengesgh.org)), which funds MalariaGEN and a number of other large research consortia, has developed a “Global Access Strategy.” This requires grantees to prepare both a strategy for commercialization of research discoveries and an intellectual-property management policy that ensures access to affordable health solutions for the benefit of people most in need in the developing world.<sup>22</sup> Key provisions of the Global Access Strategy include:

- a requirement that the principles of the Global Access Strategy apply to licences and other contracts using the consortia’s intellectual property;
- a specific condition that prevents downstream licensees of the consortia’s intellectual property from applying for secondary patents in the developing world that would impede access to affordable health solutions;
- a stipulation that prohibits exclusive licensing of consortia’s intellectual property except in cases where it is necessary as a development or marketing incentive.

The challenge is to balance celerity in commercialization with equitable access. The key question is whether a patented innovation would lead to a product that: (1) has a global market in which there are incentives for private companies to develop pharmaceuticals, or (2) a

predominantly developing-country market with few incentives for private companies. Discoveries with an anticipated global market may require monitoring to ensure access for people most in need of the drug, while discoveries that are applicable primarily to developing-country markets may require more active work to spur commercialization and development. At the extreme of this spectrum are treatments for neglected diseases such as sleeping sickness, leishmaniasis, and Chagas disease.<sup>23</sup>

### Innovations with an expected global market

A decision must be made about whether to license the innovation to a non-profit organization or to a for-profit company for further development. Non-profit organizations, often organized as public-private partnerships (e.g. Medicines for Malaria Venture, Malaria Vaccine Initiative, and the Institute for OneWorld Health) often already have access plans in line with the above Global Access Strategy.<sup>24</sup> However, there may be cases where a for-profit company would be able to dedicate more resources to development and would therefore be more likely to commercialize an end product. In this situation, steps can be taken to accelerate commercialization while maximizing global access. For example, licences can be written such that private companies are given exclusive development and marketing rights in high-income countries while preserving the possibility for generic competition to lower prices in poor countries.<sup>7</sup>

### Innovations with an expected developing-country market only

When the potential market is not a sufficient incentive for companies to invest, the options for product development are: (1) liaising with non-profit drug developers or public-private partnerships (PPPs), or (2) technology transfer to developing-world institutions.

Product-development PPPs, or PDPPPs, have grown rapidly over the past decade.<sup>25</sup> Of the 61 neglected disease drug projects in the pipeline at the end of 2004, over 80% of them were conducted by PDPPPs.<sup>26</sup> Exhaustive discussions of the PPP landscape and the models of drug discovery used in the PDPPPs can be found elsewhere.<sup>25,27</sup> PDPPPs could be effective downstream developers of an innovation whether the anticipated end product has a global

or a developing-world-only market. In both cases, nontraditional licensing arrangements — those that take advantage of the synergistic public missions of consortia like MalariaGEN and PDPPPs — could provide a useful avenue to explore for management of its intellectual property.<sup>28</sup> As PDPPPs have become more common, universities have entered into such nontraditional deals for neglected-disease drug development.<sup>29</sup> For example, the University of California-Berkeley, where scientists have developed genetically-engineered artemisinin-producing bacteria, has provided a royalty-free licence for the bacteria to the Institute of OneWorld Health (iOWH), which plans to create a new microbial factory for cheaper artemisinin production using the licensed intellectual property.<sup>30</sup>

A key pillar of MalariaGEN's strategic goals is to help developing-country partners with capacity building. While this manifests itself primarily in funding dedicated to scientific training and research capacity in developing-country laboratories, it might also be a factor in choosing a partner for downstream development of an innovation. Technology transfer to developing countries allows industry to develop technologies appropriate to their own regional needs and, by helping generate infrastructure, enables sustainable local and regional solutions to public health problems.<sup>5</sup>

The United States National Institutes of Health (NIH) has pioneered the utilization of types of intellectual

property licensing appropriate to developing-country institutions. The Office of Technology Transfer (OTT) has used geographic exclusivity or co-exclusivity as an incentive for a licensee to develop a product for a particular regional market. For example, the NIH is licensing technology related to the development of a human-bovine vaccine for rotavirus infection to institutions in Brazil, China, India, and the USA. Depending on the country and geographic region, the licence is non-exclusive, co-exclusive, or exclusive.<sup>31</sup> The degree of exclusivity was matched to the needs of the prospective licensees in each country — by segmenting the market and granting exclusive rights only when needed to spur commercialization, the strategy increases the likelihood that a technology will be developed for worldwide distribution.

Seeking to license with a developing-country partner may help circumvent the problem of market failure or delayed market entry that occurs when Western companies have little or no interest in bringing technologies to less profitable markets. Such an approach also has the advantages of potentially stimulating local economic growth. However, it should be noted that in many cases, straightforward licensing with a commercial entity, even in a developing country, will not result in products inexpensive enough to be purchased by patients in the least developed countries. Therefore, this strategy should be reserved for situations where there may be a complete market failure — including a demonstrated lack

of interest from PPPs — for a seemingly viable technology. Combining the developing-world-institution strategy with the PPP strategy by securing advance purchase commitments may also be a practicable route.<sup>32</sup>

## Conclusions

A research consortium that encompasses large-scale epidemiology and state-of-the-art genomic technology, with both developing- and developed-country partners, poses complex issues for data sharing and intellectual property. One of the most important considerations in creating policy addressing these issues is to involve all consortium members in both the initial formulation of guidelines and subsequent evaluation of the policies over the lifetime of the consortium. Therefore, the proposals laid out here should be viewed as an attempt to clarify discussion by making discrete recommendations rather than as an attempt to pre-empt that discussion in the first place. ■

## Acknowledgements

We thank Eliza Hilton, Catherine Moyes, Jane Kaye and our MalariaGEN colleagues for many helpful discussions. DAC was funded by the Rhodes Trust and DK is funded by the Medical Research Council. MalariaGEN is supported by the Bill and Melinda Gates Foundation and the Wellcome Trust.

**Competing interests:** none declared.

---

## Résumé

### Partage des données et propriété intellectuelle dans le cadre d'un réseau consacré à l'épidémiologie génomique : politiques en faveur d'une collaboration à grande échelle en matière de recherche

L'épidémiologie génomique est un secteur de la recherche dont l'objectif est d'améliorer la prévention et la prise en charge de maladies courantes à travers la compréhension de leurs origines moléculaires. Elle suppose l'étude de milliers d'individus, appartenant souvent à différentes populations, par des techniques précises. L'ampleur et la complexité de ces recherches ont imposé la formation de consortiums de recherche. Les membres de ces consortiums doivent convenir de politiques pour gérer les moyens mis en commun et traiter les données génétiques. Cet article porte sur le partage des données et les politiques en matière de

propriété intellectuelle pratiqués par un consortium international de recherche travaillant sur l'épidémiologie génomique du paludisme. Il présente d'une manière générale les règles régissant spécifiquement le transfert des échantillons et des données entre ses membres et la diffusion des résultats dans le domaine public, les cas où l'on s'efforcera de protéger la propriété intellectuelle et la façon de gérer cette propriété. Il expose également certaines solutions pratiques reposant sur les principes de base de promotion de l'innovation et de l'accès au traitement.

---

## Resumen

### Intercambio de datos y propiedad intelectual en una red de epidemiología genómica: políticas de colaboración en investigaciones a gran escala

La epidemiología genómica es un campo de investigación que aspira a mejorar la prevención y el manejo de enfermedades comunes

profundizando en el conocimiento de sus causas moleculares. Para ello estudia poblaciones de millares de individuos, a menudo de

diferentes poblaciones, valiéndose de técnicas precisas. La magnitud y la complejidad de esas investigaciones ha obligado a crear consorcios de investigación, cuyos miembros necesitan acordar políticas de gestión de los recursos compartidos y de manejo de los datos genéticos. Examinamos aquí las políticas en materia de intercambio de datos y propiedad intelectual para un consorcio de investigación internacional que trabaja en la epidemiología

genómica de la malaria. Esbozamos algunas directrices específicas para establecer cómo transferir las muestras y los datos entre los miembros; cómo poner los resultados a disposición del público; cuándo decidir proteger la propiedad intelectual, y cómo gestionar la propiedad intelectual. Por último, describimos a rasgos generales algunas soluciones pragmáticas fundadas en los principios básicos de la promoción de la innovación y el acceso.

## ملخص

### تقاسم المعطيات وحقوق الملكية الفكرية في شبكة للوبائيات الجينومية: سياسات للتعاون الواسع النطاق في البحوث

وحقوق الملكية الفكرية التي تتبعها جمعية دولية للبحوث حول الوبائيات الجينومية للملاريا، ووضحنا الدلائل الإرشادية الخاصة بالتحكم في كيفية نقل العينات والمعطيات بين الأعضاء الأطراف، وكيفية الإعلان عن النتائج في وسائل الإعلام العامة، ومتى ينبغي طلب الحماية لحقوق الملكية الفكرية، وكيف ينبغي إدارة حقوق الملكية الفكرية، واستعرضنا بعض الحلول الواقعية التي تركز على أسس أساسية لتشجيع الابتكار وإتاحة المعطيات.

تعد الوبائيات الجينومية من ميادين البحوث التي تهدف إلى تحسين الوقاية من الأمراض الشائعة ومعالجتها بتفهم المصادر الجزيئية لها. وتشمل الوبائيات الجينومية دراسة آلاف الأفراد الذين ينتمون في غالب الأحيان إلى مجموعات سكانية مختلفة، باستخدام أساليب بالغة الدقة، ويتطلب نطاق ومدى تعقيد مثل هذه البحوث تشكيل جمعيات للبحوث، وينبغي أن يوافق أعضاء هذه الجمعيات على السياسات الخاصة بإدارة الموارد المتقاسمة، ويتداول المعطيات الوراثية. وقد تطرقنا في هذه الورقة إلى سياسات تقاسم المعطيات

## References

- Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. *Nature* 2005;437:1299-320.
- Ghedini E, Sengamalai NA, Shumway M, Zaborsky J, Feldblyum T, Subbu V, et al. Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. *Nature* 2005;437:1162-6.
- Mu J, Awadalla P, Duan J, McGee KM, Joy DA, McVean GA, et al. Recombination hotspots and population structure in *Plasmodium falciparum*. *PLoS Biol* 2005;3:e335.
- Risch NJ. Searching for genetic determinants in the new millennium. *Nature* 2000;405:847-56.
- Varmus H, Klausner R, Zerhouni E, Acharya T, Daar AS, Singer PA. Public health. Grand Challenges in Global Health. *Science* 2003;302:398-9.
- Chokshi DA, Kwiatkowski DP. Ethical Challenges of Genomic Epidemiology in Developing Countries. *Genomics, Society and Policy* 2005;1:1-15.
- Kapczynski A, Chaifetz S, Katz Z, Benkler Y. Addressing Global Health Inequities: An Open Licensing Approach for University Innovations. *Berkeley Technology Law Journal* 2005;20:1031-114.
- T1DGC Publication and Presentations Policies. Winston-Salem, Type 1 Diabetes Genetics Consortium, 2004.
- Sharing Data from Large-scale Biological Research Projects: A System of Tripartite Responsibility. London: The Wellcome Trust; 2003.
- The ethics of patenting DNA. London: Nuffield Council on Bioethics; 2002.
- Data release and access policy. London: Wellcome Trust Case Control Consortium, 2006.
- Eisenberg RS. Science and the law. Patent swords and shields. *Science* 2003; 299:1018-9.
- Miller LH, Mason SJ, Clyde DF, McGinniss MH. The resistance factor to *Plasmodium vivax* in blacks. The Duffy-blood-group genotype, FyFy. *N Engl J Med* 1976;295:302-4.
- Chitnis CE, Miller LH. Identification of the erythrocyte binding domains of *Plasmodium vivax* and *Plasmodium knowlesi* proteins involved in erythrocyte invasion. *J Exp Med* 1994;180:497-506.
- Arevalo-Herrera M, Herrera S. *Plasmodium vivax* malaria vaccine development. *Mol Immunol* 2001;38:443-55.
- Yazdani SS, Shakri AR, Mukherjee P, Baniwal SK, Chitnis CE. Evaluation of immune responses elicited in mice against a recombinant malaria vaccine based on *Plasmodium vivax* Duffy binding protein. *Vaccine* 2004;22:3727-37.
- Ashley-Koch A, Yang Q, Olney RS. Sick cell haemoglobin (HbS) allele and sickle cell disease: a HuGE review. *Am J Epidemiol* 2000;151:839-45.
- Modiano D, Luoni G, Sirima BS, Simpore J, Verra F, Konate A, et al. Haemoglobin C protects against clinical *Plasmodium falciparum* malaria. *Nature* 2001;414:305-8.
- Williams TN, Mwangi TW, Roberts DJ, Alexander ND, Weatherall DJ, Wambua S, et al. An immune basis for malaria protection by the sickle cell trait. *PLoS Med* 2005;2:e128.
- Fairhurst RM, Baruch DI, Brittain NJ, Ostera GR, Wallach JS, Hoang HL, et al. Abnormal display of PfEMP-1 on erythrocytes carrying haemoglobin C may protect against malaria. *Nature* 2005;435:1117-21.
- Statement on the principled conduct of genetics research of the Human Genome Organisation. Statement approved by HUGO Council on 21 March 1996, Heidelberg, Germany. *Law Hum Genome Rev* 2000;12:253-5.
- Global Access Strategy and IP Management Policy. Seattle: Gates Foundation; 2005.
- Trouiller P, Olliaro P, Torreele E, Orbinski J, Laing R, Ford N. Drug development for neglected diseases: a deficient market and a public-health policy failure. *Lancet* 2002;359:2188-94.
- Merz JF. Intellectual Property and Product Development Public/Private Partnerships. Geneva: World Health Organization Commission on Intellectual Property Rights, Innovation and Public Health; 2005.
- Widdus R, White K. *Combating Diseases Associated with Poverty: Financing Strategies for Product Development and the Potential Role of Public-Private Partnerships*. Geneva: Initiative on Public-Private Partnerships for Health; 2004.
- Moran M. A breakthrough in R&D for neglected diseases: new ways to get the drugs we need. *PLoS Med* 2005;2:e302.
- Nwaka S, Ridley RG. Virtual drug discovery and development for neglected diseases through public-private partnerships. *Nat Rev Drug Discov* 2003; 2:919-28.
- Gardner C, Garner C. *Technology Licensing to Non-traditional Partners: Non-Profit Health Product Development Organizations for Better Global Health*. Oxford: Centre for the Management of Intellectual Property in Health; 2004.
- Chokshi DA. Universities should foster neglected-disease work. *Nature* 2005;435:143.
- Daviss B. Malaria, science, and social responsibility. *The Scientist* 2005:42.
- Salicrup LA, Rohrbaugh ML. Partnerships in technology transfer. An innovative program to enhance biomedical research and global health. *Int Microbiol* 2005;8:1-3.
- Webber D, Kremer M. Perspectives on stimulating industrial research and development for neglected infectious diseases. *Bull World Health Organ* 2001;79: 735-41.