CHAPTER NINE

NIH DATA AND RESOURCE SHARING, DATA RELEASE AND INTELLECTUAL PROPERTY POLICIES FOR GENOMICS COMMUNITY RESOURCE PROJECTS

Claire Driscoll[1]

INTRODUCTION

Most observers predict significant health-related gains from genomics research. Policy and legal decisions made by government institutions, the courts and legislatures have the potential to make a significant impact on both the quantity and quality of effective and innovative healthcare-related products ultimately derived from the vibrant genomics research enterprise. In particular, the careful management of the intellectual property (IP) aspects of this promising area of research will be necessary to maximise scientific progress, provide appropriate incentives for investment, and ultimately ensure optimal public benefit.

It is the mission of the US National Institutes of Health (NIH), which is comprised of 27 individual institutes and is an agency of the US Department of Health and Human Services, to facilitate the translation of basic biomedical research discoveries into useful healthcare services and products. Within the NIH, the National Human Genome Research Institute (NHGRI) is the agency's lead entity for advancing human health through genetic research.

---

[1] Director of Technology Transfer for the National Human Genome Research Institute, National Institutes of Health. This chapter was first published as: 'NIH Data and Resource Sharing, Data Release and Intellectual Property Policies for Genomics Community Resource Projects' (2005) 15(1) *Expert Opinion on Therapeutic Patents* 1–8.

The opinions expressed in this article are solely those of the author and do not necessarily represent the views of the NHGRI, NIH or DHHS.

Through its stewardship of an array of infrastructure and research projects, including several innovative public-private consortia efforts, the NHGRI seeks to contribute to the genomic tools, data and knowledge bases. In general, I believe that scientific progress in this still young field will be best served by early, open and continuing access to: i) comprehensive, high-quality data sets containing basic biological and biochemical data; and ii) critical biological materials such as animal models and genes. Data such as the complete nucleotide sequence of many different organisms' genomes, information on genetic variation within and among populations, and results on how gene expression is regulated at the cellular and molecular level are often referred to as 'pre-competitive' information, and in my view should be made rapidly available to all, without restrictions on use. Adherence by data and resource producers and users to this simple strategy should ensure that industry and academic researchers will be able to build upon this strong foundation.

At the NIH we are expected to support basic scientific discovery whilst simultaneously facilitating the appropriate commercial research and development of the results of our formidable research programs. A sizeable number of end users for these resource projects are employed with private sector companies. For this constituency the terms governing the data use, data release, the sharing and distribution of research resources and intellectual property rights of derivative inventions are of particular importance. Policies that limit companies' ability to file patent applications or licence downstream uses could end up having an unintended inhibitory effect on the development of biomedical products. Government policies need to balance the important dual goals of continuing to rapidly place huge amounts of data in the public domain and encouraging restriction-free sharing of genomic tools, whilst also ensuring that more applied inventions, notably those closer to being an actual product, can be patented. US taxpayers, and especially patients, would like the government to appropriately foster the commercialisation of promising inventions derived from use of the data and reagents generated by these efforts. Currently, the NHGRI is actively involved in the development and vetting of policy options aimed at ensuring that genomic tools, resources and databases of genomic information are used in a manner that promotes scientific research and the practice of medicine.

Relevant policies implemented by NIH-supported public private consortia efforts such the International Human Genome Sequencing Consortium (IHGSC),[2] the Trans-NIH Mouse Initiative,[3] the Mammalian Gene Collection (MGC)[4] and the International Haplotype Map Project (HapMap)[5] are specifically covered in this review.

THE NATIONAL HUMAN GENOME RESEARCH INSTITUTE: WHERE HAVE WE COME FROM AND WHERE ARE WE GOING?

Where have we come from? As leader in the international Human Genome Project (HGP), NHGRI has learned a great deal about how to coordinate and manage an international, geographically dispersed and extremely complex 'community resource project' (though this term was not in our lexicon at the time). A decision on how to handle the prepublication release of HGP sequence data was made early on in the project, and a policy was put in place in 1997; a year prior to the start of the major large-scale sequencing phase. The NHGRI's data release and data deposition policy for DNA sequence information was designed 'to make sequence data available to the research community as possible for free, unfettered use'.[6] One of many other innovations of the HGP was the decision to include a research program on the ethical, legal and social issues (ELSI) arising from the study of the human genome. ELSI research projects, including studies on the patenting and licensing of gene patents and diagnostic tests, were funded alongside the technology development, mapping and sequencing projects.

In spring 2003 the International Human Genome Sequencing Consortium (IHGSC) celebrated the completion of the sequencing of the human genome; a milestone that, by a happy coincidence, occurred

---

[2] NHGRI, 'International Consortium Completes Human Genome Project' (Press Release, 14 April 2003) <http://www.genome.gov/11006929>; Human Genome Resources <http://www.ncbi.nlm.nih.gov/genome/guide/human/>.

[3] Trans-NIH Mouse Initiative <http://www.nih.gov/science/models/mouse/sharing/1.html>.

[4] Mammalian Gene Collection <http://mgc.nci.nih.gov>.

[5] International HapMap Project <www.hapmap.org>.

[6] NGHRI, *Policy on Release of Human Genomic Sequence Data* <http://www.genome.gov/10000910>.

during the 50th anniversary year of Watson and Crick's seminal discovery of the structure of DNA. In anticipation of the 'what's next' question, the NHGRI had invested in a deliberate, transparent and in-depth consultation process, lasting almost two years, with the extended genetics community about its future. With the help of patient advocates, ethicists, biotechnology company executives and healthcare practitioners, a consensus began to emerge as to which projects should be taken on by the NHGRI over the next decade.

In a paper entitled 'A Vision for the Future of Genomics Research'[7] published in April 2003, the leaders of NHGRI outlined a series of Hilbertian[8] grand challenges. Several of these challenges concern policy development. Both the creators and end users of the fruits of genomic research need to be actively involved in the development and vetting of policy options so that practical solutions ones that facilitate the extensive use of genomic information in laboratory and clinical settings can be implemented. Which policy issues are expected to have the greatest impact on whether or not citizens of the globe will realise benefits from genomics research in the future? Among the critical ones are data release, data and resource sharing, and patent and licensing policies. The remainder of this article focuses on the specifics of the intellectual property, data release, and research tool and data sharing policies in use for several NHGRI-funded genomics projects. Lastly, an overview of some patent and licensing-related issues that have emerged in the genomics field is provided, together with a brief summary of several possible legislative and policy fixes for these thorny IP-related problems.

GENOMICS COMMUNITY RESOURCE PROJECTS

In an important recent international meeting to discuss data release for such enterprises, the concept of a 'community resource project' was born. This is defined as 'a research project specifically devised and implemented to create a set of data, reagents or other materials whose primary utility will be as a resource for the broad scientific community'.[9]

---

[7] FS Collins et al, 'A Vision for the Future of Genomics Research' (2003) 422 *Nature* 1–13.

[8] D Hilbert, 'Mathematical Problems' (1902) 8 *Bulletin of the American Mathematical Society* 437–79.

[9] Wellcome Trust, *Sharing Data from Large-Scale Biological Research Projects: A System of Tripartite Responsibility* (Report of a meeting organised by the Wellcome Trust, Fort Lauderdale, 14–15 January 2003).

Box 1 provides examples of several large-scale genomics enterprises that aim to produce data or create valuable scientific reagents of broad potential utility. Many of the collaborators in these projects are small- and medium-sized private sector biotechnology and biopharmaceutical companies and large global pharmaceutical giants; however, academic centers have generally played a more major role.

Box 1. Examples of Genomics Community Resource Projects

- International Human Genome Sequencing Consortium (IHGSC)/Human Genome Project (HGP):[10] large-scale DNA mapping and sequencing of the human genome.
- Mouse Genome Sequencing Consortium (MGSC):[11] MGSC - large-scale DNA mapping and sequencing of the *Mus musculus* (mouse) genome.
- Rat Genome Sequencing Consortium (RGSC):[12] large-scale DNA mapping and sequencing of the *Rattus norvegicus* (rat) genome.
- The SNP Consortium (TSC):[13] discovery and characterisation of two million single nucleotide polymorphisms (SNPs) in the human genome; primarily a private sector initiative with some limited NIH involvement.
- Trans-NIH Mouse Initiative:[14] a group of initiatives for the creation and distribution of mouse genomic resources such as mutant mice, phenotypic and genotypic information and functional genomic data.
- Mammalian Gene Collection (MGC):[15] development and distribution of a complete collection of full-length cDNAs for all known mouse, rat and human genes.

[10] NHGRI, 'International Consortium Completes Human Genome Project' (Press Release, 14 April 2003) <http://www.genome.gov/11006929>; Human Genome Resources <http://www.ncbi.nlm.nih.gov/genome/guide/human/>.

[11] Mouse Genome Sequencing Consortium (MGSC) <http://www.genome.gov/10001859>.

[12] Rat Genome Sequencing Consortium (RGSC) <http://www.genome.gov/11511308>.

[13] The SNP Consortium <http://snp.cshl.org>.

[14] Trans-NIH Mouse Initiative <http://www.nih.gov/science/models/mouse/>.

[15] Mammalian Gene Collection <http://mgc.nci.nih.gov>.

- International Haplotype Map Project (HapMap):[16] creation of a haplotype map for the comprehensive study of human DNA variation among and between a diverse set of populations.
- Encyclopaedia of DNA Elements (ENCODE):[17] identification and characterisation of all the functional elements, such as regulatory sequences, encoded in the human DNA genome.

Attendees of the 1996 Bermuda[18] and the 2003 Fort Lauderdale[19] International Sequencing Consortium (ISC) meetings wholeheartedly agreed that the policy of rapid prepublication release of sequence data for projects, such as the HGP, should be extended beyond 'simple' sequence data to other types of more complex genomic data; for example, gene expression analysis/microarray data and protein structure information. 'The products of community resource projects have, over the past several years, become increasingly important as drivers of progress in biomedical research. The scientific community will best be served if the results of community resource projects are made immediately available for free and unrestricted use by the scientific community to engage in the full range of opportunities for creative science.'[20] The conclusion reached was that the architects of these other large-scale projects should seriously consider adopting the same data release policy. Of course, the exact details of how, in what format and under what type of schedule these more complex data will be downloaded into public databases still needs to be defined by the domain experts. For example, there are problems related to the optimal level and degree of validation needed for particular kinds of experimental data so that the data are useful and useable to other consortia scientists and non-consortia scientists.

[16] International HapMap Project <www.hapmap.org>.

[17] Encyclopaedia of DNA Elements (ENCODE) <http://www.nhgri.nih.gov/10005107>.

[18] Dr Bentley, 'Genomic Sequence Information Should be Released Immediately and Freely in the Public Domain' (1996) 274 *Science* 533–4; Summary of Principles agreed at the International Strategy Meeting on Human Genome Sequencing (Bermuda, 25–28 February 1996).

[19] Wellcome Trust <http://www.wellcome.ac.uk/en/1/awtpubrepdat.html>.

[20] Wellcome Trust <http://www.wellcome.ac.uk/en/1/awtpubrepdat.html>.

The conclusions reached above by the ISC participants are consistent with the now seven-year-old NIH research tools policy officially called 'Sharing Biomedical Research Resources: Principles and Guidelines for Recipients of NIH Research Grants and Contracts'[21] and the policies are also in line with the newer NIH policy on the sharing of data for extramural grantees and contractors.[22] It is important to note that databases, along with research reagents, such as genes, vectors and antibodies, are considered to be research tools. Under the policy, which went into effect on October 1, 2003, grantee institutions must submit proposed data-sharing plans in all grant applications that request US $500 000 or more per year in funds.

SPECIAL CHALLENGES IN ACHIEVING OPEN DATA ACCESS: THE SNP CONSORTIUM AND THE HAPMAP PROJECTS

The SNP Consortium (TSC) is a non-profit foundation that focused on discovering single point mutations in the human genome called single nucleotide polymorphisms (SNPs). It was formed in 1999 with a mix of public and private funds, and the final data release occurred last September. Its mission was to 'develop up to 300 000 SNPs distributed evenly throughout the human genome and to make the information related to these SNPs available to the public without intellectual property restrictions'.[23] The main idea behind the unique structuring of the consortium was that it made sense to have all interested parties, which included large pharmaceutical companies such as GlaxoSmithKline and charities such as the Wellcome Trust, share the risk and expense of developing a high-quality publicly available human SNP data set. Many companies participated even though they were not given any special use of, or early access to, the data. Nonetheless, clearly the private sector felt that the effort was worth the investment.

The TSC members agreed to adopt a policy of waiving the right to receive patent protection on the raw SNP data and agreed to publish the

---

[21] *Report of the NIH Working Group on Research Tools*, presented to the Advisory Committee to the Director (4 June 1998) <http://www.nih.gov/news/researchtools/>.

[22] NIH, *Data Sharing Information* <http://grants2.nih.gov/grants/policy/data_sharing/>.

[23] The SNP Consortium <http://snp.cshl.org>.

mapped SNPs as quickly as was feasible. A legal instrument called a Statutory Invention Registration (SIR) was used in a defensive tactic to guard against those who might try to file patent applications on identical SNP data in advance of its public deposition. SIRs were filed on the data as it was generated and then those SIRs were abandoned once the validated SNP information was downloaded to the TSC website every three months.[24]

The goal of the HapMap consortium-based effort is to determine the common patterns of sequence variation in the human genome.[25] Because *de novo* mutations occur relatively rarely in human populations any new allele that does arise will 'travel' with other nearby alleles within a continuous block of genomic DNA on the particular chromosome. These common sets of alleles are called haplotypes. At present the genotyping of hundreds of samples and the construction of detailed haplotype maps is being carried out on set of DNA specimens from populations with ancestry from various regions in Asia, Africa and Europe.[26]

'The [HapMap] project is committed to rapid and complete data release, and to ensuring that project data remain freely available in the public domain at no cost to users.'[27] A legal review of potential intellectual property problems with the project raised one major concern: it would be theoretically possible for an unscrupulous company or entity to add on a trivial amount of information to the published HapMap data and then attempt to secure 'parasitic' patent claims such that all others would be prohibited from using the original public data. This scenario along with other related concerns and a desire to be in harmony with earlier community resource projects, led the consortium members to agree upon a new data release and data use policy. Under the policy, users of HapMap data can file patent applications on associations they uncover or verify between particular SNPs and disease and/or disease susceptibility. The only caveat is that the owner or assignee of the

[24] The SNP Consortium, *Frequently Asked Questions* <http://snp.cshl.org/about/faq.shtml>.

[25] International HapMap Project <www.hapmap.org>.

[26] The International HapMap Consortium, 'International HapMap Consortium Paper' (2003) 426 *Nature* 789–96.

[27] The International HapMap Consortium, 'International HapMap Consortium Paper' (2003) 426 *Nature* 789–96.

patents cannot try to limit or prevent other users from enjoying full and unrestricted access to the HapMap data. A click-on licence agreement describing the details of the policy is posted on the HapMap project website.[28] Users must agree to the conditions of the online policy before they are permitted to access or download HapMap data.

The HapMap participants, similarly to the members of TSC, agreed that SNPs, genotypes and haplotype data of unknown specific utility are not inventions and therefore are not appropriate subject matter for patenting.[29] Conversely, an SNP or haplotype that is strongly associated with a disease or medically important phenotype, such as susceptibility to diabetes or a poor response to a particular chemotherapy drug, would clearly have a specific utility and be patentable. However, neither TSC nor the HapMap projects include phenotype or disease association studies.

THE SHARING OF VALUABLE RESEARCH REAGENTS: MOUSE RESOURCES AND THE MAMMALIAN GENE COLLECTION

The NIH encourages and actively supports the sharing of resources such as inbred mouse strains, genetically modified mice as well as the DNA vectors and murine embryonic stem cells that must be used to generate useful model knockout lines. To facilitate the timely and efficient development and distribution of a wide array of mouse resources the NIH decided to place a number of related projects under the administrative management of a new umbrella program called the Trans-NIH Mouse Initiative.[30] Beginning in 1998, workshops were convened to bring together diverse members of mouse and larger scientific community, as well as program staff from many NIH institutes, in order to research agreement on the key priorities for the initiative. Coordination and oversight of the many program such mutagenesis and phenotyping studies, genomic sequencing and the creation of mouse repositories such as the Mutant Mouse Regional Resource Centers is

[28] International HapMap Project <www.hapmap.org>.

[29] International HapMap Project <www.hapmap.org>.

[30] Trans-NIH Mouse Initiative <http://www.nih.gov/science/models/mouse/>.

carried out by a Trans-NIH Mouse Genomics and Genetics Resources Coordinating Group.

In 2003, the coordinating group also wrote a policy document entitled 'NIH Statement on Sharing and Distributing Mouse Resources'.[31] As a result of this policy, not-for-profit entities may obtain materials for use in non-commercial research using a standard Material Transfer Agreement (MTA). Importantly, the MTA cannot contain reach-through licensing terms, nor can it contain provisions that are more restrictive than those included in the widely used Simple Letter Agreement (SLA)/Uniform Biological Materials Transfer Agreements (UBMTA).[32] In most cases the decision to seek patent protection on a particular genetically modified mouse is entirely up to the grantee or contractor, as is their right under the US 1980 *Bayh Dole Act*.

The terms of some NIH funding arrangements for projects under the Trans-NIH Mouse Initiative do, however, contain an exception, known as a Determination of Exceptional Circumstances (DEC). US government grantees and contractors do not obtain title to inventions when the government makes a 'determination of exceptional circumstances'.[33] A DEC removes the standard right of ownership usually retained by the funded institution for inventions made by their employees when the research was funded, partially or in whole, with government monies. This administrative and legal tool is only used for programs in which the main goals are to create data and/or resources that are to be made widely available with minimal restrictions on their use. NIH's intention to utilise a DEC for a new program is always announced in advance of the award of a grant or contract. Regardless of whether or not a specific mouse mutant or mouse resource is patented those interested in using the animals or materials for commercial purposes must negotiate a licence with the owner.

Another important genomics research tool initiative called the Mammalian Gene Collection (MGC) program, which began in 1999, represents yet another unique undertaking by NIH. Its remit is to create

---

[31] NIH, *Statement of Sharing and Distributing Mouse Resources* <http://www.nih.gov/science/models/mouse/sharing/>.

[32] Uniform Biological Materials Transfer Agreement (UBMTA) <http://ott.od.nih.gov/NewPages/UBMTA.pdf>.

[33] As set forth at: 35 USC § 202(a) (1994).

a public collection of affordable, sequence-verified full-length complementary DNAs (cDNAs) for every known mouse and human gene, and a subset of rat genes. Feedback solicited from the NIH intramural and extramural communities at the time revealed that the lack of reasonably priced, freely available and sequence-verified 'correct' cDNA clones was indeed slowing down many research projects. After a careful study and review of the various options, the NIH decided to fund and handle the coordination for the MGC initiative. This was fortunate as perhaps no other entity had the experience, financial resources or technical wherewithal to competently take on and successfully complete such a project.

Like some programs managed under the Trans-NIH Mouse Initiative, the MGC request for funding proposals contained a DEC. The MGC also works closely with related projects to create collections of *Xenopus* and zebrafish full-length cDNAs. Not only does the MGC make available high- quality and modestly priced cDNAs but buyers are free to use the clones for research purposes and there are no onerous IP or reach-through licensing terms whatsoever.

Today the MGC physical clones are currently available to both non-commercial and commercial scientists via the IMAGE consortium distributors for a modest fee.[34] In addition, as a result of NIH partner Invitrogen's open architecture licensing policy for its Gateway® and Superscript technologies, as is also outlined in an open-access agreement with NIH, academic and government purchasers of the system are permitted to share MGC clones made using the company's technologies and/or genes cloned into Gateway® vectors with others for research purposes. The HGP, the HapMap, the Trans-NIH Mouse Initiative and the MGC are key components of the burgeoning publicly available collections of scientific resources (mainly data and research reagents) that are supported by NHGRI and other NIH Institutes. The creation of additional valuable research reagents and data sets, available to all, for the global genomics-based research toolbox is a goal that cross-cuts all of NHGRI's community resource projects.

[34] Mammalian Gene Collection, *Where to Buy* <http://mgc.nci.nih.gov/Info/Buy>.

NIH PATENT POLICY AND GENOMICS-SPECIFIC INTELLECTUAL PROPERTY ISSUES

As a result of the 1986 *Federal Technology Transfer Act* (*FTTA*) American government research laboratories such as NIH were given a statutory mandate to encourage and facilitate the efficient and expeditious transfer of promising new technologies invented in NIH intramural labs to companies for further development and commercialisation. Over the next few years government agencies created technology licensing offices and began to develop and implement patent and licensing policies. The current version of the NIH Patent Policy was adopted in 1995 and includes the following recommendations:[35]

- Seek patent protection on biomedical technologies when having IP rights will facilitate the availability of the technology for research or commercial use.
- Seek patent protection for inventions when IP rights are necessary to attract commercial partners and further R&D is required.
- Do not seek patent protection for technologies if no further R&D is required and the invention has no obvious preventative, diagnostic/prognostic and/or therapeutic use (for example, has no commercial or public health value).
- Do not unduly delay or avoid the public disclosure of research results because of anticipated future patent filings (reasonable delays of one to several months are permitted).

The above policy does not distinguish between different types of inventions; it is purposefully general in nature so that it can be applied to all NIH-developed inventions.

Earlier in 2004, the NIH opted to publish a guidance document specifically aimed at outlining what are considered to be our own internal best practices for the handling of genomic inventions. This new draft publication was entitled 'Best Practices for the Licensing of

[35] *United States Public Health Service Technology Transfer Manual*, Chapter 200: PHS Patent Policy.

Genomics Inventions',[36] and it summarised the guiding principles used by licensing specialists at the NIH when they negotiate deals. One of the most provocative, and in my view the most sensible, suggestions had to do with the licensing of diagnostic applications of genomic inventions: entities are strongly encouraged to non-exclusively licence diagnostic inventions or, at a minimum, to grant only narrow exclusive licences for these kinds of inventions (limiting the licence to particular fields of use such as one type of technology platform, one or a few disease indication(s) and/or certain geographic regions). These suggestions are similar to those put forward by the Nuffield Council on Bioethics in their 2002 discussion paper.[37]

It is important to encourage the commercialisation of as many types of clinical diagnostic tools and tests as possible, as competition in the marketplace should translate to an increase in the number of innovative and affordable products available to patients and their physicians. By publishing practical details on how we at NIH handle the licensing of various types of genomic inventions, we hope to give academics ideas for how they might handle the licensing of similar inventions within their own universities and hospitals.

To get a patent in the US one must show that the invention is: useful, non-obvious or inventive, and novel. The Nuffield Council on Bioethics report[38] on the ethics of patenting DNA discusses several important issues with regard to the legal concepts of utility and inventiveness. In general, the European Patent Office (EPO) holds patent applications to a higher standard of inventiveness than does the US Patent and Trademark Office (USPTO). The EPO states 'the isolation of DNA sequences that have a structure closely related to existing sequences in which the function is known is not inventive'.[39] Using computational tools to identify homologues and assign hypothetical functions, so-called *in silico* analysis, to a DNA sequence

---

[36] *Best Practices for the Licensing of Genomic Inventions*, 69 Federal Regulations 67747–8 (2004).

[37] Nuffield Council on Bioethics, *The Ethics of Patenting DNA: A Discussion Paper* (July 2002) <www.nuffieldbioethics.org>.

[38] Nuffield Council on Bioethics, *The Ethics of Patenting DNA: A Discussion Paper* (July 2002) <www.nuffieldbioethics.org>.

[39] Nuffield Council on Bioethics, *The Ethics of Patenting DNA: A Discussion Paper* (July 2002) <www.nuffieldbioethics.org>.

would not be considered inventive in Europe; whereas according to the USPTO rules the use of non-laboratory bench computer-based methods would not necessarily exclude the granting of a patent for a gene on the basis of non-obviousness.[40] Fortunately, the USPTO issued revised utility guidelines in early 2001. Inventions now must show a 'specific and substantial and credible utility'.[41] This is certainly an improvement but only one specific, substantial and credible utility per DNA sequence is needed and so a composition-of-matter patent on a gene sequence can easily come to encompass new uses discovered long after the initial filing, even if those uses were not known at the time or not disclosed in the original patent.[42] In my opinion this is unfair as it penalises those who do the hard work of figuring out the biological function(s) of genes and proteins. The USPTO utility guidelines should be modified to eliminate this 'loophole'.

Box 2 contains information on three published reports and one ongoing study, all of which focus on gene patent and/or genomic patent and licensing issues. Patents are a lynchpin for successful commercial entities in the genomics field; they are a driving force behind innovation and a guarantee that the discoveries will be fully disclosed and speedily delivered for scientific and societal benefit. Nonetheless, patents with overly broad claims, as well as those with unsubstantiated data or a complete lack of data for the specific claimed uses, should not be allowed to issue. Also, there is the issue of the exponentially growing body of patents in the genomics area; this has led to a so-called patent 'thicket'[43] problem. To ensure their freedom to operate in a complex marketplace many companies often must licence a large range of overlapping and related patents; the high licensing and transactional costs in terms of both time and money could be causing some companies to abandon efforts to try to commercialise needed diagnostic and therapeutic products.

---

[40] Nuffield Council on Bioethics, *The Ethics of Patenting DNA: A Discussion Paper* (July 2002) <www.nuffieldbioethics.org>.

[41] *USPTO Utility Examination Guidelines*, 66 Federal Regulations 1092 (2001).

[42] E Marshall, 'Patent on HIV Receptor Provokes an Outcry' (2000) 287 *Science* 1375–7.

[43] MA Heller and RS Eisenberg, 'Can Patents Deter Innovation? The Anticommons in Biomedical Research' (1998) 280 *Science* 698–701.

Box 2. Recent Gene Patent and Intellectual Property Policy Reports

- Nuffield Council on Bioethics, *The Ethics of Patenting DNA: a Discussion Paper* (July 2002) <www.nuffieldbioethics.org>.
- The Royal Society, *Keeping Science Open: the Effects of Intellectual Property on the Conduct of Science* (April 2003) <www.royalsoc.ac.uk>.
- The Australian Law Reform Commission, *Gene Patenting and Human Health Discussion Paper* (March 2004) <www.austlii.edu.au/au/other/alrc/publications/dp/68/>.
- US National Academy of Sciences, *Intellectual Property Rights in Genomics and Protein-Related Research* (Report expected in March 2005)<www.nationalacademies.org> (under Current Projects tab).

The sequencing and public disclosure of the human genome and the gearing up of several other genomics community resource projects has effected a shift in the commercial landscape. Over the last few years, with sequence information, SNP information and soon haplotype data publicly available at no cost, a number of companies have exited the genomic information database subscription business. Not surprisingly, many of these same 'content' companies are in the process of transforming themselves into biopharmaceutical companies. Consequently, they are now focusing exclusively on activities such as functional genomics, target validation and the screening of small molecular libraries in their quest to identify promising lead molecules.

Finally we must ponder the ramifications of a recent US court ruling. Historically, biotechnology and pharmaceutical companies have almost never sued academic investigators for patent infringement, as long as no active selling was occurring, even if the not-for-profit scientists were using the company's patented invention. The harsh reality of the lack of a formal research exemption in US law was brought to everyone's attention with the *Madey v Duke*[44] court decision. As succinctly stated by

[44] *Madey v Duke University* 307 F 3d 1351 (2002).

University of Michigan law professor Rebecca Eisenberg, '[t]he recent rejection by the Court of Appeals for the Federal Circuit of an "experimental use defence" to a patent infringement lawsuit against Duke University. . . is an alarming wake-up call to the academic community'.[45] That court decision now makes it possible, and perhaps even likely, that cash-strapped or aggressive companies with no immediate revenue streams will assert their patent rights against university researchers.[46] Rather than taking a wait-and-see approach, an attempt should be made by lawmakers and academics to create a true research exemption in US patent law.

EXPERT OPINION

The free sharing and wide dissemination of pre-competitive genomic data and research resources has numerous significant benefits to research institutions, companies, scientists and the general public. Therefore, we do not support the idea of the US enacting *sui generis* database protection legislation similar to the Database Directive passed in the EU.[47] Share and share alike 'open source code' - like data release and use policies such as the ones mentioned in this article inevitably encourage researchers to undertake different scientific approaches as they attempt to unravel the intricacies of complex biological systems. The following are just a few of the benefits that result from the creation and maintenance of centralised databases and repositories: improved ease of access; avoidance of duplication of effort; and more efficient use of limited research funds. Projects such as the IHGSC, International Mouse Genome Sequencing Consortium (IMGSC), International Rat Genome Sequencing Consortium (IRGSC), TSC and HapMap have unequivocally demonstrated the usefulness of having open access comprehensive databases that contain vast quantities of genomic information.

Initiatives such as the MGC and Trans-NIH Mouse Initiative have drawn attention to the immense value of research tools for aiding and

---

[45] RS Eisenberg, 'Patent Swords and Shields' (2003) 299 *Science* 1018–19.

[46] C Dennis, 'Geneticists Question Fees for Use of Patented "Junk" DNA' (2003) 423 *Nature* 105.

[47] EU Database Directive 96/9/EC (1996) <http://europa.eu.int/ISPO/infosoc/legreg/docs/969ec.html>.

accelerating scientific progress in all sectors: academic, government and commercial. All of these projects bring together the collective financial and intellectual resources of many diverse partners, and result in efficiencies of operation, scale and speed not normally associated with government- or academic-backed endeavours. Participants in Genomics Community Resource Projects are cognisant that the greatest opportunity for value creation using the data and research tools developed by consortia efforts will occur downstream from the discovery of basic genetic and biological information.

CONCLUSION

In summary: i) I hope that universities and others will follow the suggestions outlined in NIH's new guidance document on the best practice for licensing genomic inventions; ii) I support the concept of enacting legislation to establish a formal research exemption for the non-commercial research use of patented inventions and technologies by not-for-profit entities; iii) I support the idea of raising the utility 'bar' even higher than was done in the 2001 revisions of the USPTO's utility guidelines for biotechnology inventions; and iv) I support the concept of compulsory licensing and/or compulsory sublicensing by patent holders of genomic/genetic diagnostic technologies and inventions.[48]

Today it is much easier for scientists everywhere to rapidly build upon genomic research carried out by their colleagues. We at NIH look forward to the establishment of additional public private partnerships that are dedicated to placing pre-competitive data in the public domain. As a leader in the fields of genetic research and genomic science, NHGRI will continue to help encourage policy debates and support research,[49] which informs policy development, in order to facilitate the widespread use of genomic information in both research and clinical settings, and in order to facilitate the development of new genomics-based products.

---

[48] Proposed Genomic Research and Diagnostic Accessibility Act of 2002 (HR3967); T Abate, 'Do Patients Wrap Research in Red Tape' *San Francisco Chronicle*, 25 March 2002.

[49] Example: request for application RFA-HG-04–004: *Intellectual Property Rights in Genetics and Genomics* <http://grants1.nih.gov/grants/guide/rfa-files/RFA-HG-04-004.html>.