17 Tex. Intell. Prop. L.J. 413

Texas Intellectual Property Law Journal
Spring 2009

Article

OF MICE AND MEN: WHY AN ANTICOMMONS HAS NOT EMERGED IN THE BIOTECHNOLOGY REALM

Chester J. Shiu[a1]

I.	Introduction	414
	A. Problem	414
	B. Thesis	415
	C. Outline	416
II.	Basic Science Background and Applications	417
	A. Basic Science Background	417
	B. Applications	421
	C. Anticommons	425
III.	The Structure of Biomedical Research	427
	A. Publicly Funded Research--The National Institutes of Health	427
	1. The NIH Liberally Licenses Upstream Innovations Under Its Control	428
	2. The NIH Uses Its Market Power to Force Others to Adopt Similarly Liberal Licensing Terms	430
	3. The NIH Indirectly Discourages Lawsuits Against Academic Researchers	432
	4. The NIH Mitigates Any Incentives to Adopt an Anticommons Business Model	434
	5. The Future	439
	B. Summary	442
IV.	Confusion in the Literature About the Definition of Upstream Research	442
	A. Classifying Different Types of Inventions	444
	B. Robustness of Patents Among the Four Categories	446
V.	Empirical Research--The Missing Anticommons	450

VI. Conclusion 454

*414 I. Introduction

A. Problem

Modern biology has its origins in the discovery of microscopy during the seventeenth century,[1] though the modern experimental disciplines that make up the life sciences today did not emerge until the end of the nineteenth century.[2] As a practical matter, this relative youth means that a great number of fundamental discoveries have yet to happen in the biological sciences. It also means that the ability of the patent system to fairly apportion the intellectual property rights stemming from these nascent fundamental discoveries is relatively unsettled and is still the subject of much discussion within both the legal and scientific communities.[3]

In particular, there has been a worry that excessive patenting of these new pioneering discoveries may inhibit further exploratory work in these fields. In a seminal paper, Michael Heller dubbed this type of problem a "tragedy of the anticommons."[4] A tragedy of the anticommons can occur when "multiple owners are each endowed with the right to exclude others from a scarce resource, and no one has an effective privilege of use. When too many owners hold such rights of exclusion, the resource is prone to underuse--a tragedy of the anticommons."[5] In the biomedical realm,[6] Heller and Rebecca Eisenberg identified two specific anticommons scenarios in which patents might unduly increase the costs of downstream product development. In the first scenario, patents on numerous ***415** "upstream" technologies, or research tools, act like "tollbooth[s] on the road to product development, adding to the cost[s] and slowing the pace of downstream biomedical innovation."[7] In the second scenario, reach-through license agreements on patented upstream technologies are used to obtain "rights in subsequent downstream discoveries" (e.g., royalties on sales and licenses on future discoveries).[8] However, in the ensuing decade, the predicted biomedical anticommons problem has largely not come to pass.[9]

B. Thesis

This Article argues that Heller and Eisenberg's feared biomedical research anticommons has not come to pass for two overarching reasons: first, the structure of biomedical research keeps upstream innovation in the hands of public entities committed to maximizing the dissemination of discoveries; and second, the definition of upstream research in much of the legal literature has been too imprecise to accurately frame the question. Together, these reasons suggest that the impetus toward a biomedical research anticommons has been overestimated, and thus, the associated fear has been overblown.

With regard to the first reason, most upstream research is structured and funded by the public sector, led by the National Institutes of Health (NIH).[10] The NIH has inhibited the creation of an anticommons in three ways: mandating liberal licensing of any funded discoveries; fostering a critical mass of upstream innovators that private for-profit firms dare not sue; and launching initiatives that pre-empt potential anticommons-based business models,[11] often by releasing discoveries and information into the public domain.

***416** With regard to the second reason, the various patents that touch upon upstream innovation vary widely in their usefulness, and without appropriate definition we cannot identify and characterize those patents that might actually contribute to an anticommons. Part IV identifies four categories of upstream patents--those that describe a therapeutic protein, describe a marker correlated with a disease, describe a potential drug target, or disclose a research tool. The relative scientific value of each category is different, and we will see that the majority of patents filed describe either a marker or a potential drug target--two categories that, despite their impressive sounding labels, border on practically worthless. Thus, there is no anticommons problem, in part, because far less of the biomedical research domain has been fenced in than previously supposed, a fact obscured by the conflating of all four categories.[12] Furthermore, this imprecision has masked the relative ease with which scientists in the biomedical research field can circumvent many patents, thereby further relieving any tendency toward an anticommons.[13]

Hence, an anticommons in biomedical research has not arisen for two interrelated reasons: powerful public actors have successfully kept the biomedical research domain open, and the privately held portions of the domain are smaller and less

valuable than previously feared--a fact obscured by imprecise definitions in the literature.

C. Outline

Part II of this Article provides an overview of the basic science required to understand the arguments in this Article and concludes with an illustrative example. Part III of this Article outlines the structure of biomedical research in the U.S., particularly the role played by the NIH. Part IV details the confusion in legal scholarship over the types of upstream patents in the biomedical research space and attempts to clarify the debate. Part V considers empirical evidence that an anticommons has not formed and analyzes this using the framework set out in Part IV. Part VI concludes this Article and considers areas for future research.

*417 II. Basic Science Background and Applications

A. Basic Science Background

Deoxyribonucleic acid (DNA) is the primary carrier of hereditary information for life on Earth.[14] DNA is a continuous chain, with each link in the chain made of one of four types of molecules.[15] By convention, each of these four molecules is represented by a particular letter--A, T, C, or G--and the DNA chain spells out a sequence of letters.[16] A gene, the smallest unit of heritability, is any part of the sequence that spells out the code for a protein.[17] A genome is the sequence of all the unique DNA chains in an organism.[18]

The central dogma of molecular biology states that information flows in one direction: DNA is transcribed into RNA, which is translated into protein.[19] Stretches of DNA contain genes, and each gene codes for one or more proteins, via an RNA intermediary.[20] Thus, if one alters the DNA of an organism, the protein coded for by that DNA will also be altered.[21] Likewise, one can achieve the same effect by instead directly manipulating the RNA transcribed from a portion of DNA.[22] All of this matters because proteins are the most versatile component of ***418** any living system: "they function as chemical catalysts, they transport and store other molecules such as oxygen, they provide mechanical support, they provide immune protection, they generate movement, they transmit nerve impulses, and they control growth and differentiation."[23] From this we can see that the importance of genes lies in their ability to create proteins. The process of actively transcribing a gene, the first step in creating a protein, is termed expression.[24] If an expressed gene becomes defective, the corresponding protein becomes defective, its corresponding function is altered, and the end result is often a disease.[25]

One wrinkle in this scheme is that some proteins, known as transcription factors, also regulate gene expression.[26] These transcription factors can regulate the transcription of their own genes,[27] other genes,[28] or both. These regulated genes, in turn, might also code for other transcription factors that regulate even more genes.[29] Thus, a mutation in one expressed transcription factor gene might trigger a cascade of changes in gene expression.[30] Wide-scale changes in gene expression would alter the available quantities of various proteins, wreaking havoc on the chemical reactions and other functions that these proteins mediate.[31] We generally term the end result a disease--though not all diseases are of this type.[32] ***419** The important point here is that biological systems are deeply interconnected, and slight modifications to one element can have wide ranging implications.

Of particular interest to us is the "advent of the molecular era in biology in the 1940s and 1950s, and in particular the development of . . . recombinant DNA technology in the mid-1970s."[33] The tools of molecular biology permitted scientists to, for the first time, "isolate individual genes and determine their chemical composition."[34] "The ability to map and sequence genes has . . . yielded highly detailed knowledge of the structure of evolutionary trees, increased our understanding of genetics, and led to the development of new diagnostics and therapeutics for diseases such as hypertension and cancer."[35] With the completion of the Human Genome Project (HGP),

> [R]esearch has progressed beyond creating an inventory of human genes (mapping and sequencing) to efforts aimed at elucidating gene functions, comparing the human genome with those of other species, studying the interactions between genes and the environment, analyzing the structures and functions of proteins encoded by genes, and ultimately determining the role of genes and proteins in human as well as in animal and plant biology.[36]

What distinguishes post-HGP efforts is their comprehensiveness. For example, prior to the sequencing of the human genome,

studying gene-environment interactions involved making educated guesses about which genes were worth studying.[37] Even allowing for excellent scientific intuition, this process was fraught with error. Now, instead of guessing, in a quick experiment one can directly measure the change in expression of every single gene in the genome in response to a particular environmental change, identify those genes whose expression differs from baseline, and proceed to study this subset of responsive genes in greater detail.[38] Thus, instead of blindly choosing a set of genes to study, ***420** a biomedical researcher can rapidly conduct a comprehensive search for genes of interest before zooming in to do deeper research. To reiterate, what makes the post-HGP world so exciting is the ability to conduct life science experiments with unprecedented scale and accuracy. One of the motivating factors in this continuing research is to improve our ability to diagnose and cure diseases.

The invention of recombinant DNA technology allowed scientists to insert and manipulate genes in a live organism.[39] Beyond the knowledge gained from this new experimental technique, the ability to insert a copy of a gene into an organism meant that researchers could efficiently generate large quantities of the protein that a particular gene encodes.[40] The first application of this new capability was the large-scale synthesis of pure insulin for diabetic patients.[41] Previously, insulin was obtained by grinding up pig pancreases--an expensive, inefficient, and potentially unsanitary process.[42] Now, cells could be modified to pump out endless quantities of therapeutic insulin, an event which ushered in the dawn of the biotechnology industry.[43]

In this Article, biotechnology consists of two areas: molecular and large-scale.[44] Molecular biotechnology consists of the products and processes of isolating, preparing, and replicating fragments of DNA and RNA to produce proteins, as well as the use of any molecule to manipulate the physical and ***421** chemical processes of living organisms.[45] Large-scale biotechnology, which is a subset of bioinformatics, deals with analyzing the large amounts of information created by recent advancements in computer processing power and molecular biology in an attempt to find useful correlations.[46] For example, the HGP helped spur genomics, the study of an entire organism's genome.[47] These large-scale techniques have two broad uses: improving clinical diagnoses by correlating certain genes with diseases and directing scientists, via interesting correlations, to areas of further research.[48] These two endeavors are interrelated, as the end result of a scientific study may be a clinically useful diagnostic test. At a high level of abstraction, the definition of biotechnology in this Article is simply any use of genetic information, either to diagnose a disease or to alter an organism's physical or chemical pathways through physical manipulation. This manipulation might be to treat a disease or to create some product that aids in the therapeutic process.

B. Applications

Modern molecular biology has provided a number of insights into the molecular basis of disease, insights which are leading to new strategies for diagnosis and therapy. DNA testing for specific genes provides definitive diagnoses for certain heritable diseases without the ambiguities of previous indirect phenotypic measures (though genetic tests sometimes have their own associated uncertainties for reasons that are beyond the scope of this Article).[49] DNA testing also can be used to defensively screen for diseases, permitting action to be taken before overt symptoms develop; for example, if a patient is found to possess the BRCA1 gene, which predisposes the patient for breast cancer, her doctor would urge more frequent mammograms, among other prophylactic measures.[50]

"Exploiting molecular insights . . . to craft alternative therapies has proven to be more challenging than developing new diagnostic tools."[51] In the late 1980s, ***422** gene therapy, which tried to correct the underlying genetic defect by providing a patient's cells with a properly functioning copy of the defective or missing gene, was highly touted as the future of medicine.[52] This was particularly appealing as it seemed to cleanly apply the central dogma to disease treatment: fix the defective gene to fix the disease-causing protein, and thereby fix the disease.[53] Unfortunately, practically implementing gene therapy has turned out to be exceedingly complicated, and the technique has yet to achieve much success.[54]

More encouraging, however, is the growing ability to indirectly ameliorate the problem by altering the surrounding chemical pathways.[55] As noted in this Part, biological systems are deeply interconnected.[56] In many cases, the true impact of a defective protein is its effect on the overall balance of the interconnected pathways that it participates in.[57] Thus, if a defective protein no longer performs its function, perhaps we can fashion the molecular equivalent of a crutch to either replace the defective function or balance out the deficit elsewhere to bring the overall system back into balance. For example, if the defective protein is a catalyst in the middle of a chemical reaction, then we might alter other parts of the reaction to ensure that the appropriate end product is still produced.

To gain a sense of how the approach plays out in real life, let us consider the development of the small-molecule drug

imatinib (Gleevec/Glivec), an astoundingly effective anti-cancer drug.[58] In 1960, researchers in Philadelphia noticed that one of the chromosomes of patients suffering from chronic myelogenous leukemia (CML), a form of blood cancer, was too short.[59] DNA in humans is organized into physical units called chromosomes, and an unusually short chromosome suggested that some DNA had gone missing.[60] In 1973, another ***423** researcher discovered that the missing DNA had been fused to a different chromosome.[61] As an aside, note how the 13 year gap between identifying the missing segment and locating its new home highlights the enormous advantage of contemporary whole-system surveys, which would have made short work of the problem. In 1986, researchers showed that the merged DNA, created from the fusion of part of one chromosome with another, had created a new gene.[62] This new gene, bcr-abl, of course, had a corresponding new protein, BCR-ABL, which was shown to be the cause of CML.[63]

In CML patients, BCR-ABL functions as the master control that regulates the production of white blood cells.[64] Unfortunately, BCR-ABL is jammed in the on position, leading to the uncontrolled cancerous growth we call CML.[65] The gene ***424** therapy strategy would be to replace the bcr-abl gene with a gene coding for a properly functioning regulator--a staggering, Nobel-prize-worthy feat of molecular engineering that would require selectively modifying millions of copies of bcr-abl in the relevant cells, and only the relevant cells, without killing the patient in the process. A far more practical approach is to just block the out-of-control production signal being given off by BCR-ABL.[66] That is, if one could jam its continuous growth signal, the entire growth system would be brought back into balance and the unchecked progression of CML would be stopped.[67]

This was the strategy employed by the creators of the drug imatinib, which has yielded spectacular results in patients who would otherwise have died within a few years of diagnosis.[68] In 1992, the basic compound that would become imatinib was synthesized.[69] The first clinical trial started in 1998, and the drug was approved by the FDA in 2001.[70] What happened between 1986 and 1998? Dr. Juerg Zimmerman, a medicinal chemist, developed a range of compounds that attempted to block the continuous growth signal emanating from BCR-ABL.[71] There are a number of large hurdles to overcome when designing a compound. Essential factors considered include (1) activity, the ability to jam BCR-ABL; (2) specificity, the ability to selectively jam only BCR-ABL; (3) efficacy, how well it jams BCR-ABL; (4) toxicity, harmfulness to the patient; (5) permeability, ability to penetrate into the cancerous cell so that it can effectively jam BCR-ABL; and (6) bioavailability, the ability of the body to absorb it in pill form.[72] A compound that fails any of these factors is unacceptable.[73] A team of biologists screened Zimmerman's compounds to check for effectiveness, giving Zimmerman feedback as he iteratively refined the design of his compounds.[74] Then, the most promising of these compounds were studied by a leading CML specialist, who flagged the ***425** prototype for imatinib for further development.[75] The molecular biology background provided by other scientists greatly informed this search, but it was still immensely time consuming.[76] Thus, 1992 to 1998 was spent fine tuning all 6 factors to optimize imatinib's effectiveness in anticipation of clinical trials.[77] As one can see, simply finding a drug that is active can take years of dedicated work by highly skilled chemists. Even with extensive knowledge about the gene, protein, and affected biological system, development of an effective drug took 12 years.

C. Anticommons

Imatinib provides a good example of the drug development process in the biotechnology context. Other biotechnology innovations can be understood as assisting in some aspect of the process detailed in the previous section or in some way leveraging the same scientific knowledge to improve patient diagnosis. One point to appreciate is that various points along the research and development process described in the previous section could have fallen within the scope of one or more patents.[78] Accordingly, at these points any of the patent holders would have the ability to prevent research from proceeding.[79] A given patent holder may ask for some form of compensation to allow research to continue or may simply refuse to allow research to proceed. If the patent is valid, then perhaps we should believe that the patent holder has every right to do this.[80] However, some have worried that if too many different people hold patents in an area and, as a result, it becomes difficult to secure all the rights necessary to proceed with further research, the overall progress of research may be unduly impeded. This is one variant of the ***426** tragedy of the anticommons problem that Michael Heller and Rebecca Eisenberg worried about.[81]

Patent literature speaks of upstream and downstream development, which correspond roughly to earlier and later in the process of developing a viable commercial product.[82] An anticommons in biomedical research would occur if too many patents on upstream discoveries made it prohibitively expensive to proceed on downstream product development work.[83] In the development of imatinib, the important discoveries, from upstream to downstream, were: discovering that chromosomal fusion causes CML, identifying genes affected by chromosomal fusion, isolating the relevant gene among these (bcr-abl), isolating and characterizing the protein (BCR-ABL) coded for by the relevant gene, and designing a molecule to neutralize

BCR-ABL.[84] It is clear that if no one knew that chromosomal fusion causes CML, the rest of the discoveries would not have followed. From this, another possible anticommons scenario would be if one or more entities--acting in concert or independently-- had in aggregate patented all of the genes that were potentially affected by the fusion event and demanded a royalty from any product that results from the exploration of these genes, regardless of the number of intervening steps required to create a commercially viable treatment. Such a demand, possibly from multiple patent holders at multiple junctures, could have made it prohibitively difficult for other scientists to identify bcr-abl as the root cause of CML and derailed the eventual development of imatinib.

It is also important to note that in the above example, only one of the genes actually mattered;[85] a researcher would not know this and, in addition to spending time researching the set of all potentially pertinent genes, would have to waste resources securing the rights to this set. This uncertainty about the usefulness of various items is a hallmark of biomedical research, and it implies that there are network effects at play.[86] These network effects stem from the observation that there is great value in having a complete collection of knowledge in a particular area, whereas the value of an incomplete collection is highly variable.[87] If ***427** knowledge in that area is severely fragmented among many owners, as would be the case in an anticommons, then it might be prohibitively expensive to assemble a valuable complete collection. Thus, in this context, the anticommons problem has the potential to raise the cost of research dramatically and deter a number of desirable discoveries.

III. The Structure of Biomedical Research

A. Publicly Funded Research--The National Institutes of Health

Public funding of biomedical research is coordinated by the National Institutes of Health (NIH).[88] In its own words, the NIH "is the steward of medical and behavioral research for the Nation. Its mission is science in pursuit of fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to extend healthy life and reduce the burdens of illness and disability."[89] Its 2008 fiscal year budget was $28.7 billion spread over 27 different institutes and centers.[90] The institutes cover all aspects of biomedical research, from arthritis to cancer to mental health.[91] One institute, the National Human Genome Research Institute, was responsible for funding and coordinating the Human Genome Project, among its many other responsibilities.[92]

The influence stemming from this budgetary clout is profound. Half of all federal nondefense R&D and over 60% of federal funded research in American universities is supported by the NIH.[93] In overall biomedical research, in 2003 ***428** industry supplied approximately 57% of all funding, private not-for-profit groups contributed 7%, and combined federal, state, and local government funding made up the remaining 36% (78% of which came from the NIH).[94] However, the 57% figure is misleading because it includes all research spending, not just basic R&D. If we look only at basic R&D, the grist of upstream innovation, federal spending actually exceeded private (industry) spending nearly four-fold in 2003.[95] The influence that comes with such massive spending allows the NIH to prevent the formation of an anticommons in three important ways: liberally licensing upstream innovations under its control, indirectly discouraging lawsuits against academic researchers, and mitigating any incentives to adopt an anticommons-based business strategy.

1. The NIH Liberally Licenses Upstream Innovations Under Its Control

Before considering the NIH's licensing posture, it is worth noting the actual clout it wields when it comes to patents. The U.S. government is the second largest holder of DNA-based patents, behind only the University of California system.[96] In addition, one conservative estimate projects that, via federal funding, the federal government retains an interest in at least half of the patents held by academic institutions.[97] For some perspective, as of 2006 the NIH had 1,364 active licenses and held approximately 4,000 issued or pending patents.[98] These patents cover a number of key innovations, enough for the NIH to keep a list of its top 20 ***429** commercially successful inventions.[99] While biotech and pharmaceutical companies still hold the majority of patents, the federal government controls the single largest bloc of patents,[100] and the NIH is the dominant government entity in this area.[101]

Indeed, the NIH has been very concerned about patenting and licensing practices impeding advances in biomedical research.[102] This concern was reflected in the "Best Practices" guidelines it published in the Federal Register for all genomic inventions, including "cDNAs; expressed sequence tags (ESTs); haplotypes; antisense molecules; small interfering RNAs (siRNAs); full-length genes and their expression products; as well as methods and instrumentation for the sequencing of

genomes, quantification of nucleic acid molecules, detection of single nucleotide polymorphisms (SNPs), and genetic modifications."[103] The guidelines on licensing agreements recommend broad, non-exclusive licensing, except when another approach is needed to induce post-discovery commercialization.[104] Furthermore, in recent years, the NIH has required researchers applying for more than $500,000 in funding--a quite common sum--to submit a plan for data-sharing in their grant applications.[105] In addition, beyond the NIH, it is worth noting that individual academic institutions, as a matter of practice, generally retain the right to transfer their rights to other nonprofits for further research.[106] This reinforces the idea that both the NIH and its academic partners tend to license their discoveries liberally to advance science rather than threaten to withhold discoveries to maximize economic gain. Thus, in addition to being substantial in size, the portion of the biomedical research domain under public control is handled by its owners in an open manner that minimizes the potential for an anticommons thicket.

***430** 2. The NIH Uses Its Market Power to Force Others to Adopt Similarly Liberal Licensing Terms

Beyond making inventions under its control readily available to all comers, the NIH also uses its substantial market power to ensure that other actors adopt satisfactorily liberal licensing strategies. Two famous examples of NIH interaction with private industry center on the genetic modification of mice. Mice are very useful laboratory animals for basic research, and various lines of genetically modified mice have been created for specific types of experiments.[107] One way to modify a gene is simply to stop its expression.[108] This results in what is known as a knockout mouse,[109] so-called because one or more of its genes has been, effectively, knocked out. The NIH Knockout Mouse Project (KOMP) seeks to produce a series of knockout mice lines, one for each gene in the mouse genome, for distribution in research.[110] The KOMP can be thought of as a reaction to a potential anticommons problem.

A very popular method for creating knockout mice utilizes a technology, called Cre-lox, developed by a DuPont scientist.[111] DuPont made this technology widely available through a licensing agreement that did not require any cash payment but did provide some very aggressive terms for subsequent use, including "reach-through" provisions that seemed to imply that any invention or discovery made using these mice or Cre-lox technology would be owned by DuPont.[112] Here we see a private firm leveraging its upstream position in an attempt to profit from downstream research in ways beyond the direct use of its product, which was a large concern of Heller and Eisenberg when positing the anticommons problem.[113] ***431** Knockout mice are vital to research; although a handful of prominent universities refused DuPont's terms (e.g., University of California, Harvard, Johns Hopkins, and MIT), more than 100 did sign the license agreement.[114] Seeing that some of its most important research partners were locked out of this vital technology, the NIH intervened and persuaded DuPont to remove the reach-through provisions from its licensing agreement with all universities.[115]

How did the NIH persuade DuPont to revise the licensing terms? The NIH Office of Technology Transfer, while discussing licensing scenarios similar to the Cre-lox dispute, stated that the "NIH does not support the coupling of procurement with intellectual property rights and restrictions and expects Recipients [of NIH funding] to ensure that NIH-funded tools are not restricted as a result of such agreements."[116] In other words, the NIH threatened to forbid the purchase of Cre-lox technology by anyone who receives NIH funding--an overwhelming share of the basic R&D market, as discussed at the beginning of this Part.[117]

Another way that mice can be modified is by increasing the expression of a particular gene.[118] One example of this is the OncoMouse, developed at Harvard.[119] As opposed to knockout mice, the OncoMouse has a gene that is expressed at a higher than normal rate, in this case one that increases the likelihood of developing cancer.[120] This makes the OncoMouse particularly useful for cancer research, and DuPont has an exclusive license for the patents related to this immensely useful line of mice.[121] The original licensing terms imposed by DuPont ***432** for using OncoMice created large logistical headaches over who could or could not receive animals from supplier repositories.[122] In particular, the licensing terms reserved to DuPont the right to downstream royalties, defined on the basis of sales or profits of products developed with the mouse as a research tool.[123] Thus, research based upon the OncoMouse led to a complicated web of licensing restrictions, and distributing the mouse required being mindful of these restrictions. This restricted access to OncoMice impeded downstream research, and once again the NIH used its clout to come to an understanding with DuPont to remove the problematic downstream licensing provisions.[124]

With both Cre-lox and the OncoMouse the NIH took on a massive private entity, DuPont, and forced it to relent. Beyond the actual outcome of these disputes, such actions serve notice to other would-be aggressive licensors about the types of terms that will be tolerated by the NIH.

3. The NIH Indirectly Discourages Lawsuits Against Academic Researchers

Even if one assumes, arguendo, that the set of currently issued biotechnology patents has created an anticommons thicket, it is only on paper. An anticommons becomes problematic for downstream researchers only if the rights holders enforce their rights, or, at the very least, potential infringers believe that the rights holders would enforce their rights. Why would patent holders in the biotechnology arena hold back from enforcing their rights, or at least allow the impression of restraint to persist? There are two reasons: because they want to maintain goodwill with the people who improve their inventions and because the damages from any suit would be small.[125]

***433** Eisenberg notes that many of the pioneering discoveries in upstream biomedical research that paved the way for lesser, but more financially viable, discoveries were paid for by NIH.[126] The work leading to the perfection of DNA recombination technology discussed in Part II was done by Herbert Boyer at the University of California San Francisco and Stanley Cohen at Stanford, both with NIH funding.[127] Microarrays, which underpin much of modern genomic research, were invented by Pat Brown at Stanford with support from the NIH.[128] Celera, the private company that raced the HGP to sequence the human genome, was founded by Craig Venter with technology that he had developed at the NIH.[129] These examples highlight the pivotal role that NIH-funded technologies have played in the advancement of biomedical research.

Biotech is primarily in the business of taking pioneering innovations and further developing them toward some commercially viable product. Thus, the industry as a whole is very dependent on the upstream work produced by NIH-funded scientists and, it can be reasonably stated, is not in a hurry to anger them with lawsuits.[130] One commentator has noted that "researchers, whether public or private, are less likely to enforce their patents when it will erode the personal relationships and the information exchange integral to the scientific community."[131] For example, in the DuPont Cre-lox controversy discussed earlier in this Part, the technology at issue had been improved by a number of publically funded labs, including scientists at the University of Cologne in Germany, the NIH, and the ***434** Massachusetts Institute of Technology.[132] In a more recent example, academic researchers at Stanford developed a new type of cellular analysis tool by combining a standard type of imaging tool with statistical analysis algorithms from computer science.[133] Among other possible uses, the new tool may enable dramatically more precise cancer diagnosis, opening up a potentially lucrative new market to manufacturers of the imaging tool.[134] Furthermore, beyond funding scientifically valuable work, the NIH reinforces this social norm by actively advocating against threats to easy information exchange, as detailed in Subsection 2 of this Part.[135]

Additionally, suing would be particularly unwise because academic scientists usually have little money in comparison to large pharmaceutical companies. Suing might be worthwhile if a researcher's university was found liable for infringement,[136] but the plaintiff would still have to contend with the social problem of angering useful scientists. Furthermore, the NIH, as exemplified in its actions against DuPont in Subsection 2 of this Part, is likely to step in to deter aggressive private sector actors. Thus, the risk-benefit calculus of upstream patent holders is likely heavily tilted against taking legal action against downstream researchers. As a result, regardless of the legal state of the biomedical research domain, it is effectively being treated as open by downstream researchers.

4. The NIH Mitigates Any Incentives to Adopt an Anticommons Business Model

An anticommons business model is any business plan that focuses on conduct that, either alone or when aggregated with similar conduct by other firms, helps promote an anticommons. While the NIH cannot possibly cover every minute niche of every domain of biomedical research, it can strategically marginalize many of the private firms that might seek to profit from an anticommons.

***435** Beyond controlling the licensing of its own patents, the NIH also has the ability to take offensive action against would-be overzealous patent holders. First, it can expand the scope of the prior art by releasing copious amounts of data into the public domain. A number of industry players have reported that "[t]he publication of the human genome and the accumulation of filed patent applications have reduced the scope for discovery of novel DNA sequences."[137] While the NIH certainly contributed to the buildup of patent applications, its unique contribution was its ability and willingness to dump an entire genome into the public domain.[138] The release of the first full working draft of the human genome in 2000 marked the beginning of a decline in gene patent applications, though modifications to the USPTO utility rules also undoubtedly played a part.[139] These two events were symbiotic: beyond rendering specific gene sequences prior art, the HGP helped raise the bar for what was considered a technically valuable discovery worthy of patent protection.

Second, beyond releasing large amounts of data, the NIH has consciously engaged in projects that claim most of an area of

upstream research, thereby rendering attempts to exploit that area commercially unviable.[140] The NIH, by funding various initiatives, develops a cluster of related upstream discoveries in an area and rapidly pushes those upstream discoveries into the mainstream.[141] This push into the mainstream impacts commercial enterprises in three ways. First, by cultivating related upstream technologies, the NIH covers large portions of the intellectual domain for a given area; this limits the ability of private firms to fragment a particular domain of research and helps preclude an anticommons.[142] ***436** Second, by implementing upstream discoveries on a broad scale and making the results readily accessible at a low cost, it renders the commercial market unappealing to private firms. Third, the speed with which the NIH pushes a discovery into the mainstream greatly narrows the window of opportunity for private firms to engage in behavior that might trigger an anticommons. In essence, for a given field the NIH snaps up large portions of the domain and reduces the financial incentive for private firms to attempt to carve out an anticommons. The NIH has this first mover ability because of the dominating role it plays in funding upstream research as noted earlier.[143] Without commercial incentive, there is no reason for private industry to move in and create an anticommons or exploit one that can be crushed by the NIH. The NIH is particularly adamant about the public availability of what it terms "pre-competitive" information,[144] which roughly tracks what Heller identified as upstream innovation, so we can expect this trend to continue.

To reiterate, the relationship between public and private research is not merely symbiotic--the public side of the equation actually has the ability to deny the private side certain objectives. By the time a private firm sees an opportunity, the upstream research is often well on its way towards becoming routine,[145] so the firm must decide whether it will engage the NIH in a race. The best example of this is the race between Celera and the HGP to sequence the human genome. There was no private effort to systematically sequence the genome until the founding of Celera.[146] Indeed, the putative enormousness of the task was what justified the founding of the international Human Genome Project.[147] While Celera certainly ***437** spurred the HGP to work faster, in the end Celera was not able to profit by patenting genes or licensing databases of gene sequences.[148] In particular, Celera originally aimed to sell genomic information to drug development companies, a plan that was largely undercut by the free public availability of equivalent information, courtesy of the HGP.[149]

With respect to the DuPont Cre-lox controversy, once the licensing terms were modified, the NIH moved to create its own line of knockout mice, utilizing nineteen of its institutes and issuing grants to numerous universities.[150] The NIH also obtained licenses from two major private industry partners for their existing lines of knockout mice, as well as funding to create additional lines.[151] These actions collectively make up the KOMP.[152] Upon completion, the KOMP will render moot any anticommons problem with knockout mice because the NIH will be able to supply any researcher with the needed lines. Likewise, the NIH's headaches with DuPont's OncoMouse licenses spurred the NIH to create its own patent-free (and hence license-free) lines of mice.[153]

With both knockout mice and the OncoMouse, the NIH is taking the same measures: use its clout to negotiate licensing terms that are less problematic from an anticommons perspective and then swiftly move to undermine the potential anticommons. The key is that the NIH is not simply passively reacting to private moves towards an anticommons; it is proactively launching initiatives to foreclose ***438** the possibility. The reason firms cooperate is they recognize that moving downstream is more financially lucrative in the long term.[154] The NIH is foreclosing actors who would rather remain in this upstream space and profit by engaging in behavior that might foster an anticommons, or at least heavily tax downstream research. As one commentator noted, "[w]here patent law may have gone too far, NIH has decided to take matters into its own hands."[155]

Leaving mice behind, we can find other examples of the NIH's willingness to pursue public projects in fields that are rife with anticommons potential. For example, Small Nucleotide Polymorphisms (SNPs) are small fragments of DNA sequences that are highly correlated with the onset of a particular disease, a trait that makes them useful for disease diagnosis.[156] What makes them particularly useful is that the sequences are relatively short, making them easy to check for--one can think of a SNP as a landmark in a genome that lets a tester quickly know whether a particular disease-causing gene is nearby, thus making diagnosis easier.[157] One can readily imagine an anticommons arising with a large number of firms patenting various SNPs, leading to a scenario where running one diagnostic test could mean infringing on numerous patents held by different firms.[158] This is important because simply knowing that a SNP exists does not tell you anything about which diseases it might help diagnose.[159] Thus, this is a potential anticommons scenario because a diverse array of upstream SNP patent holders could inhibit necessary downstream research on actual diagnostic associations, never mind the final commercialized diagnostic test.

***439** In this critical area, the NIH, as well as many drug companies, has taken action to prevent this scenario from arising.[160] The identification of SNPs has been greatly assisted by the HGP, and the NIH has joined the International HapMap Project to catalog and release all SNPs into the public domain.[161] In October 2007, HapMap released a detailed analysis of 3.1 million known SNPs.[162] Work on identifying new SNPs in new populations is ongoing, but much as the HGP raised the bar for utility

in gene patenting by making possession of a sequence rather trivial,[163] it is difficult to imagine a private firm successfully patenting a significant number of SNPs at this point in time. Thus, as with the HGP and KOMP, the NIH has assisted in keeping a particular area of the biomedical research domain open. The net result is that the NIH is proactively making it difficult for business models with great blocking potential to emerge around upstream innovations.

5. The Future

As noted before, the NIH has first-mover advantage in many areas because it funds so much upstream research. There are several areas of high value current research that could develop into anticommons thickets in the somewhat distant future. These areas are protein structures and predicting the interactions between drugs and genes. The Protein Structure Initiative aims to determine the physical structure of all proteins and make them readily available.[164] The Pharmacogenomics Research Network is aimed at correlating an individual person's genes to his response to a drug.[165] For brevity, this Article will only discuss the Protein Structure Initiative.

As stated in Part II of this Article, proteins perform numerous crucial functions in living organisms. That Part focused on how defective genes give rise to defective proteins, per the central dogma. However, nothing was said about the precise nature of a protein's defect, which involves the three-dimensional physical structure of the protein.[166] In particular, changes to a gene's sequence yields ***440** changes in the resultant protein's physical structure and composition.[167] The structure of a protein enables whatever function it performs, much as the sharp blade of a knife enables its cutting function; hence a defect in a protein's structure often results in a defect in its function.[168] Thus, beyond sequencing and cataloging genes, the next major step for biomedical researchers is to determine the structure of their associated proteins and to begin to understand the relationship between gene defects and their corresponding protein structural defects.[169] From an anticommons perspective this does not seem so bad because determining a structure requires significant effort, which should deter researchers from trying to obtain patents on knowledge stemming from protein structure unless they intended to do some downstream research.[170] For some perspective, between July 2005 and June 2006 the NIH, via an intensive initiative analogous to the HGP, determined the structure of 425 proteins.[171] By comparison, with present technology it is possible to sequence one billion bases (letters) of DNA per day.[172] Even allowing for a gene length of one million letters--an absurdly large figure--in less than a day, a single machine could sequence as many genes as the NIH has determined protein structures in over a year.

However, one potential problem area is the use of computer algorithms to predict the structure of a protein based on its associated gene sequence. These methods are not, at present, very accurate,[173] but if patents were allowed on the structures predicted by these methods, there could be an enormous anticommons problem. That is, there could be a scenario where numerous patents are issued on structures that are not actually well-characterized, and those patents could be used to interfere with downstream research that would lead to commercial products based upon knowledge of that structure. In order to meet the utility requirement of patentability without expending much physical labor, such an anticommons-leaning patentee might resort to computational methods like homology-based function ***441** prediction, which attempts to infer the function of a new structure by comparing it with known structures.[174] While such an attempt seems farfetched now, remember that this hypothetical contemplates future, greatly improved algorithms. For some perspective on the potential for such rapid technological advancement, consider that when the HGP was officially launched in 1990 as an international, multi-decade effort to sequence a single genome, it would have been unfathomable that an entire bacterial genome could be sequenced in a single day, a feat that was first accomplished in 2000.[175]

Yet even in this nascent area, we see an NIH initiative that would undercut any possible anticommons. The Protein Structure Initiative aims to repeat for protein structure what the HGP did for genome sequencing.[176] While at present, no private firms appear to be threatening to create an anticommons in this area,[177] this situation is analogous to the HGP before the invention of shotgun sequencing. In the 1980s, no one would have predicted a rush to patent gene sequences, as sequencing a single gene was enormously laborious; even Heller and Eisenberg did not express their concerns in Science until 1998.[178] That the NIH is already deeply involved with protein structure determination fits with its mission of funding upstream research. No one would argue that the NIH is in this area right now to foreclose an anticommons, but it is comforting to note that, should history repeat ***442** itself, the NIH already has the Protein Structure Initiative in place to respond to a looming anticommons crisis.

B. Summary

Legal commentators tend to focus on adjusting the patent regime to optimize upstream and downstream incentives without

creating an anticommons.[179] This Part indicates that, particularly in an area where patent law is permissive of excessive fragmentation, the NIH is likely to actively intervene, both to resolve the immediate situation and to challenge the underlying conditions that permitted the situation to form in the first place. While the impact of the NIH on negotiating reasonable licensing terms has been discussed in the past,[180] its ability to actively mitigate the formation of businesses that aim to foster and profit from an anticommons has not received sufficient attention. Given the examples recited in this section, thus far NIH intervention has been quite effective, and consequently, NIH intervention in the biomedical arena represents a very large confounding factor that seems to have been either ignored or overlooked by many commentators--an omission that ought not be repeated in the future.

IV. Confusion in the Literature About the Definition of Upstream Research

Beyond the active intervention of the NIH, there is another reason that the anticommons feared by Heller and Eisenberg has thus far not come to pass. For various reasons, the patents filed thus far in the biomedical field have not been effective at fragmenting the commons. That is, the patents in this area either cover discoveries that are not very valuable, or their enforceable scope is very limited. Taken together, these patents do not encompass enough of the biomedical research domain to portend the creation of an anticommons.

The intellectual framework for this argument is grounded in the observation that the definition of upstream research in legal literature has been imprecise. This is important because it has led legal commentators to gloss over deep differences in the value of different types of upstream research. Eisenberg provides a good example of this:

> Issuing broad patent rights to upstream research performers, or otherwise permitting them to use their intellectual property to reach through to the profits from downstream product development, ensures that they recover the full social value of their inventions, ***443** including the value that they contribute to subsequent inventions that might be more directly profitable. . . . Although the improvements may have more stand-alone commercial value than the primitive versions of the invention developed by the pioneer, the pioneer has taken greater risk and shown more ingenuity by opening up a new field. Giving pioneers broad patents allows them to force subsequent improvers to negotiate for licenses, thereby capturing for pioneers some of the follow-on value created by those who merely tweak inventions to make them more marketable.[181]

The problem is that pioneering inventions, as properly understood by Eisenberg, are exceedingly rare. Part III.A.3 noted three seminal biotech inventions, and they reasonably delineate the most important, and enabling, biotech breakthroughs of the past half century.[182] By framing the discourse in terms of pioneers, when there are not many, ignores the actual dynamics between upstream and downstream players. However, there are non-pioneer patents that can block downstream research, and it is important to understand the nature of that blockage and its relation to innovation. There is a gradient of patentable upstream discoveries bounded by pioneering discoveries on one end and discoveries that have no downstream applications on the other. The problem is that commentators focus on the pioneering end of the gradient to highlight the problem, but they then say little more about the rest of the gradient; this seems to imply a cognitive bias towards the pioneering side of the gradient.[183] This is key, because patents that fall towards the other end of the gradient have less of an impact on downstream research, thus reducing their contribution towards a possible anticommons. The bias towards the pioneering side of the gradient is not well justified, and this error can largely explain why the resulting fears of an anticommons have not been borne out--patents have not encapsulated as many rights as supposed because they do not claim what scientists in the field value.

A corollary to this observation is that firms have not been able to realize value from these less-than-ideal patents. As a result, one might expect the number of patent filings to fall over time, which has indeed been happening since 2001[184] and ***444** is discussed at length in Part V. For now, the reader should keep this observation in the back of his or her mind while considering the categories of patents advanced below.

Commentators as a whole are not completely oblivious to the distinction between the pioneering and the mundane;[185] rather, they have relied too much on exceptional cases to buttress their particular theories, and this over-reliance has distorted the debate. Indeed, the author of this Article readily concedes that there is an abundance of patents that might inhibit or has inhibited downstream research unduly[186]--the question is whether these examples, numerous though they may be, can be used to support the proposition that there is a systemic anticommons problem in the biomedical research domain. As current

empirical research has failed to find an anticommons,[187] some introspection by commentators is called for, and the focus on exceptional cases might provide a partial explanation for the disconnect between the legal discussion and empirical reality.

A. Classifying Different Types of Inventions

An empirical study by Lori Pressman, Richard Burgess, and others has shed new light on what motivates licensing and patenting behavior at universities, which comprise one of the largest groups of upstream innovators in the U.S.[188] Upstream innovations in DNA-related technology can be divided into four broad categories: therapeutic proteins, markers for disease, DNA sequences for drug targets, and research tools.[189] While these categories were synthesized with regard to universities, the observations attached to each category are relevant to both public and private patent seekers.

1.) Therapeutic proteins make up the most valuable category of discoveries.[190] These are structures, or their associated sequences, that ***445** directly act upon some target to ameliorate disease. For example, discovering insulin, or the human insulin gene that codes for insulin (and knowing that it codes for insulin), would fall within this category. Because discoveries in this category produce something tangible (e.g., insulin) with obvious utility, they have clear commercial value and patentability. According to the university licensing offices who responded to the Pressman survey, they would likely patent these discoveries and license them exclusively.[191]

2.) Marker sequences are fragments of DNA shown to correlate to a disease in some way.[192] The reason for this correlation may not be known. Unless the marker strongly correlates with a disease that is commercially worth developing a diagnostic test for, universities will likely not bother patenting, or they will engage in non-exclusive licensing if they do patent.[193]

3.) DNA sequences encoding drug targets are genes involved in disease pathways. In essence, these are the targets for therapeutic proteins. In the Gleevec example given earlier, this is the bcr-abl gene. The layperson may suspect that these are quite valuable, but in practice they are of marginal commercial interest.[194] Universities accordingly only bother spending the money to patent this class of discoveries when the apparent commercial value of the potential treatment is high.[195]

4.) Research tools is a catch-all category for everything else. Research tools can best be categorized as tools for discovering other information that does not directly result in a commercial product.[196] One example is automated gene sequencing machines. Money can certainly be made by selling some ***446** tools, but the tools are only means to other research. The NIH requires that broad, nonexclusive licenses be granted for these tools.[197]

These categories seem quite reasonable, as we are ultimately concerned with the impact of patents on downstream research. This categorization scheme highlights the ability of certain patents to influence downstream research, which is appropriate in light of this concern.

B. Robustness of Patents Among the Four Categories

It is difficult to quantify the distribution of published patents among the four categories; however, we can infer some reasonable conclusions. To make these inferences, we must first know what is patentable. A patent application must fully disclose the invention, provide a written description of the invention, and do so in sufficient detail as to enable a practitioner of ordinary skill in the art to make and use the invention without undue experimentation.[198] The invention must have utility (do something useful), have novelty (not disclosed before in publications, etc.), and be nonobvious over the prior art.[199] In general, novelty requires that a patent claim describe an invention that has not been described completely by a previous reference,[200] and nonobviousness requires that a claim not describe an invention that could be described by a relatively straightforward combination of references.[201] With these criteria in mind, let us consider the patentability of the four categories and how this might affect the number of patents filed within a given category.

Therapeutic proteins are, in essence, drugs. Assuming that the proteins are novel and nonobvious, properly written patents covering therapeutic proteins would easily meet the utility, written description, and enablement requirements that will vex some of the other categories, as described below. A therapeutic protein ameliorates a disease state, thus it clearly does something useful, and hence it has utility. In addition, it is fairly easy to fully describe a therapeutic protein and do so in a way that would enable others to use it, satisfying the written description and ***447** enablement requirements. Accordingly, it can safely be concluded that therapeutic protein patent filings would continue in a more stringent patent environment, save an

explicit statutory bar.

Marker sequences may run into written description problems or, more fundamentally, novelty issues in light of extensive public sequencing efforts. As stated previously, markers are useful because they correlate with some known disease. However, in that same vein, knowing a marker sequence in itself does not tell a practitioner of ordinary skill in the art how to productively utilize it. Or, even if it does tell the practitioner enough, it might not be novel, because a public effort has already released the sequence claimed in the patent application into the public domain. Only a marker that can serve as a diagnostic test might be worth patenting.[202]

Drug targets are at the greatest risk in light of University of Rochester v. G.D. Searle.[203] The fundamental problem is that drug targets, by themselves, rarely meet the written description requirement; they provide the target, not the ammunition to take out the target.[204] In University of Rochester, the court upheld the invalidity of a patent claiming a method of treatment involving selective inhibition of a particular enzyme.[205] The touchstone of the case was that the patent's written description of the claimed selective inhibitor did not "describe the claimed subject matter in terms that establish that [the inventor] was in possession of the . . . claimed invention."[206] In essence, the applicants specified a wish list of desirable properties in the ideal therapeutic agent, but could not describe the therapeutic agent itself.[207] This problem will always exist--if one already has the therapeutic agent needed to treat the drug target, one would just file a patent on the therapeutic agent.

Furthermore, beyond the written description problem, drug targets also suffer from a scientific uncertainty problem. Biological systems are exceedingly ***448** complicated; it is extremely difficult to target a particular gene and, even if one does successfully target it, there is no guarantee that the disease will be treated.[208] One way of looking at this is that even though one might have a promising target, it is very hard to know if it is the right target, in that it can be used in some way to successfully treat a disease. Accordingly, even if one were to somehow get around the University of Rochester written description problem, a patent still very well might be worthless. The bcr-abl gene was anomalous in this regard, because it happened to code for a pivotal protein in a straightforward pathway, so it was plainly obvious that it was the right target to go for.[209] On balance, though, the combination of the written description and the scientific uncertainty problems make the economics of drug target patents very dicey. David Adelman has argued that this scientific uncertainty problem implies that the biomedical research commons is continually expanding and, therefore, highly resistant to being fragmented into an anticommons.[210] This line of reasoning is not at odds with the arguments that underline this Article. Indeed, they complement each other: not only is the biomedical research commons, per Adelman, expanding, it is being patrolled by the NIH, and its exact composition has not been properly characterized by legal commentators.

As a category, research tools have received by far the greatest number of patents, accounting for almost 50% of the biotechnology patents granted from 1990 through 2004.[211] The primary problem facing the patentability of research tools is establishing utility.[212] To have utility, the tool must do something substantial; the Federal Circuit has held that simply revealing the presence of a fragment of a gene is not enough, the claimed invention must "provid[e] an immediate, well-defined, real world benefit."[213] At the opposite end of the research tool spectrum, ***449** microarrays, a tool designed to reveal the expression of thousands of genes simultaneously, are found in a number of issued patents.[214] Utility is a somewhat awkward vehicle for evaluating research tool patents, but after In re Fisher, it is essential to carefully consider the utility of a patented research tool when trying to determine its enforceability against downstream research.[215] The primary difference between the two extremes seems to be that the patentee in In re Fisher was trying to extend the scope of his discovery into downstream research,[216] whereas the microarray patents only deal with gene fragments insofar as they relate to microarrays.[217] Thus, insofar as research tools can do something useful (e.g., quantify the expression level of many genes) without requiring follow-up work, we should expect to see patent filings continue.

Thus, all things being equal, therapeutic proteins should be the most robust in the face of declining patent filings. This robustness indicates something important--namely that therapeutic protein patents cover something very valuable and, by extension, hold a more valuable portion of the biomedical research domain than the other three categories. Thus, if large swaths of the biomedical domain were cordoned off with therapeutic protein patents, we might expect an anticommons. This is not to say that an anticommons cannot arise with one or a combination of the other categories, just that a large number of therapeutic protein patents could potentially cause an anticommons.

***450** Along the same lines, valid patents in the other categories tend to be rather narrow in how they can affect downstream research due to their limited scope. This is important because, ultimately, anticommons impose a societal cost when they impede valuable downstream research. Thus, because these patents are limited in how they can impact downstream research, they weigh less on our fears of a potential anticommons. Of course, there could be a handful of important patents in these

categories that enables its owners to create an anticommons replete with onerous reach-through licensing agreements. But on balance, such a scenario seems less likely than with therapeutic proteins because of the narrow nature of the discoveries under patent.

Markers are essentially not patentable--or at least no one would bother trying--unless they can yield a diagnostic test, which is a downstream commercial product.[218] Drug targets are not patentable under University of Rochester unless the inventor already possesses a therapeutic agent,[219] in which case the downstream commercial product is also close to realization.

Research tools are more problematic. They are, by definition, not attached to any downstream product, so there is a potential to create an anticommons problem, at least in light of this analysis. For example, use of a research tool might be contingent upon agreeing to a license with reach-through provisions on downstream discoveries. Also, tools have the potential to be tied to many different downstream discoveries, increasing their reach relative to the other categories. Fortunately, attempting to leverage research tools to reach downstream developments is fairly easy to spot, and the NIH has acted decisively when confronted with this scenario.[220]

If many of the patents issued over the past decade fall within the marker or drug target class, or perhaps even the research tool class, then it is possible that the most important parts of the biomedical domain have not been fenced in yet. This idea is the focus of the next Part.

V. Empirical Research--The Missing Anticommons

The results of a report commissioned by the National Research Council contained a surprising finding: practicing scientists are largely oblivious to ***451** patents.[221] This seems to indicate that an anticommons problem has not come to pass. In fact, the study concluded, "Our results offer little empirical basis for claims that restricted access to IP is currently impeding biomedical research . . . ."[222] This is particularly perplexing because, as a matter of law, scientists are liable for any patents that they infringe, even for academic research.[223] This Article argues that this seeming indifference has two root causes. First, in spite of the holding of Madey v. Duke,[224] investigators continue to benefit from a de facto research exemption. Second, most of the patents in question are not as valuable as they appear; consequently, investigators either ignore or readily circumvent them.

Part III.A.3 of this Article argued that the private sector indulges infringement by public researchers because of their pivotal role in spinning out new technologies. In particular, for-profit firms recognize that academic use may improve their inventions, that maintaining goodwill ensures access to future innovations, and that the potential damages are commonly relatively small.[225] Beyond the indulgence of the private sector, it is possible that the patents held by would-be plaintiffs are simply not strong enough to stand up to litigation. It is useful to note that the confusion in legal literature about the nature of upstream innovation tracks an interesting observation: not all upstream innovation is valuable. As noted previously, the number of DNA-based patent applications filed around the world has been falling since 2001.[226] The discussion in Part IV.B concluded that therapeutic protein patents were the most valuable portion of the biomedical research domain, in that they seem to cover something more substantial than the other three categories of patents. It is unsurprising, then, to discover that therapeutic-protein related patents (protein structures) are the only category to not suffer a decline in filings since 2001.[227] The small number of such patents filed (no more than eleven in a year) is also reassuring from an anticommons perspective--the rate of fencing in is quite low.[228]

The decline in filings in the other three categories of patents coincides with the implementation of the U.S. Patent and Trademark Office's elevated utility ***452** requirements, which private firms report having difficulty meeting.[229] Some in private industry have admitted that some of the patents filed previously were not grounded in solid biological data.[230] One way to interpret this is that what was known when the patents were filed was far less than what was actually claimed in the patents. Accordingly, such patents may be difficult to enforce. In addition, the heightened utility requirement may have resulted in new patent applications with far narrower scope than past applications. Thus, even if there was a will to enforce them, the patents, both pre- and post-heightening, may not effectively encircle very much of the biomedical research domain.

This means that researchers should find it relatively easy to work around a patent if the threat of litigation actually arises. For example, as noted previously, over 3.1 million SNPs, a form of disease marker, are now publicly available.[231] A marker patent holder would find it exceedingly difficult to exclude a scientist from working on a particular disease, since the scientist could most likely find a substitute marker in the public domain.[232] The one exception to this would be work on developing a specific

clinical diagnostic test using a marker, where scientists regularly abandon work on a test due to potential infringement.[233] This can best be thought of as research where substitutes from the public domain may not suffice, so it makes sense that specific patents have a deterrent effect.

Drug target patents seem more problematic in that if one wants to investigate a patented target gene, there is no real substitute. However, once again, patents in the area are less valuable than they initially appear. Approximately 20% of all known human genes (4,382 of 23,688 genes) have been explicitly claimed in U.S. patents, and approximately 2,690 of these genes are claimed by private firms.[234] However, 2,000 of the 2,690 privately claimed genes are claimed at least in part by Incyte, and its claims are primarily concerned with the use of the genes as probes on microarrays.[235] In other words, Incyte's claims do not reach into the actual function of the genes, so researchers are still essentially free to research the genes. ***453** Yet there are two confounding factors when considering the 2,000 gene figure. First, other entities almost certainly claim some of the same genes as Incyte and do so in broader terms, so it is overly simplistic to claim that 2,000 privately claimed genes are completely open to downstream researchers. Second, however, Incyte is hardly the only private entity concerned with gene probing,[236] so 2,000 is only a lower bound on the genes that might be claimed in a similarly narrow fashion. In addition, consider the following figures: as of 2005, there are 6,145 issued patents on genes and gene regulation compared to 7,428 issued patents for gene probing; at the same time, 7,105 applications related to genes and gene regulation were pending compared to 16,983 for gene probing.[237] Of course, these figures do not explicitly relate to how many genes are being claimed and the scope of specific claims, but they help to develop an intuition about the relative proportion of broader gene patents versus narrower gene probing patents. Setting Incyte aside, patents on the remainder of the 2,690 privately claimed genes may indeed potentially block downstream research. However, as the Incyte patents demonstrate, there is little reason to believe, prima facie, that the claims have sufficient scope to substantially encompass the activities of downstream researchers.

In addition, given the decrease in patents filed since the U.S. Patent and Trademark Office increased the stringency of its evaluations, we might reasonably assume at least some of the patents issued in the past are defective with regard to utility or written description and therefore would not survive litigation in the post- In re Fisher and University of Rochester legal regime. This seems particularly likely given the qualitative admission by public and private-sector officials interviewed by Hopkins that past gene patents may have been speculative or lacked sufficient biological data to back up the claims.[238] Thus, the remaining genes covered by privately held patents represent a loose upper bound on the area carved into an anticommons, if one has been carved at all. From a downstream research perspective, private patenting of potential drug targets has not encompassed a particularly large part of the biomedical research domain because either the claims do not reach relevant downstream research, the patents themselves are suspect, or there simply is no patent coverage in the area.

Patents related to research tools have also fallen since 2001.[239] Given the declining trend, it is likely that the same critique about the enforceability of previously issued patents and the scope of newly issued patents applies. In addition, a study conducted by the National Academy of Sciences and the National ***454** Science Foundation in 2003 concluded that "drug discovery has not been substantially impeded by [research tool patents]. We also find little evidence that university research has been impeded by concerns about patents on research tools."[240] Thus, considering that research tool patents did not particularly hinder research during the peak of patent issuance, it seems unlikely that an anticommons thicket will arise as the rate of issued patents falls. Fortunately for this analysis, the relationship between rate of issuance and anticommons formation is somewhat irrelevant, as the NIH has repeatedly demonstrated a willingness to use its leverage to break anticommons-related activities that affect downstream research.[241]

Thus, from a downstream research perspective, private patenting of both markers and potential drug targets has not encompassed a particularly large part of the biomedical research domain. Either the claims do not reach relevant downstream research, or the patents themselves are suspect. This is not to discount the potential downstream blocking power of particular privately held patents; this Article simply argues that the potential for such blocking is relatively low for the reasons stated above, and these patents in aggregate likely do not carve out enough to create a true anticommons, which is consistent with empirical observations. Furthermore, patents on therapeutic proteins are quite rare and therefore unlikely to produce the sort of patent thicket that characterizes an anticommons. At the opposite end of the spectrum, there have been many research tool patents, but as an empirical matter they have yet give rise to an anticommons. Taken together, empirical research in the four categories describes a relatively open research domain.

VI. Conclusion

There is still a potential for a biotech anticommons, but much of the dissonance between past literature and reality can be traced to undervaluing the primacy of public efforts in forestalling an anticommons and to imprecise definitions. Once we take into account better definitions, we see that many of the patents issued over the past decade or so have not carved out large blocking upstream positions. Accordingly, in the context of these patents, an anticommons has not emerged because not enough of the biomedical research domain has been ***455** fenced in. Insofar as the remaining patents do matter, the NIH has taken proactive steps to mitigate their effects on downstream research either by negotiating appropriate licenses or by placing the necessary upstream information into the public domain via its own research initiatives. In addition, as the main driver of upstream biomedical research in the U.S., the NIH's ability to beat private-sector firms to various key upstream positions is exceptionally potent. This ensures, at least so far, that the NIH sets the agenda in the upstream research world.

What might go wrong? Technological breakthroughs could boost the number of therapeutic protein patents, an area of the biomedical research domain that has not been extensively tested. However, the existence of the NIH Protein Structure Initiative suggests that the NIH is well equipped to handle any problems in that area. On the other hand, the NIH could be blindsided by a private firm that develops a pivotal upstream technology. In that case, since the NIH still controls so much of the funding and, thus, the potential market for the nascent technology, it will still have some bargaining power to ensure fair licensing terms, much as it has in the past.

Nevertheless, it is somewhat disconcerting to have to rely on the NIH as a bulwark against a biomedical anticommons. Once we understand what the public sector does and refine the categorical definitions used to describe the biomedical research domain, there remain areas of concern, which we can now correctly bracket. In theory, a better solution might be to statutorily bar patents on innovations that are too far upstream, thus relieving the NIH of the burden of having to catch every potential anticommons-creating innovation. However, determining the appropriate parameters for such a bar is exceedingly nontrivial and beyond the scope of this Article.[242] Going forward, it would be useful if more empirical studies were performed on licensing behavior, as this would help inform the picture about industry's attitude towards academia.[243] Given that patents take years to go through prosecution, additional empirical patent studies should also be performed to update our current knowledge of patenting trends in the four categories suggested by Pressman et al. A study on the enforcement of patents would also be helpful, though it would inevitably suffer from undercounting the null result--if a firm perceives an issued patent as worthless it will just sit on it, making it difficult to distinguish truly tolerating infringement from no infringement existing.

***456** If the quality of patents in the four categories increases, their downstream effects bear close monitoring. This Article has argued that at least two of the four categories, markers and gene targets, are intrinsically less likely to play a large downstream blocking role. It is possible that this observation has held up only thanks to poorly drafted patents; if the assumption of intrinsic inability to block is incorrect, the situation will be far more critical than previously supposed.

Likewise, if the NIH suffers considerable funding cuts or sustained under-funding, it may begin to lose its ability to effectively preempt private firms in the upstream arena. Alternately, it may find itself with insufficient resources to respond to emerging threats. Furthermore, the NIH's effectiveness as a pre-emptive agent depends on the competence of its administration, which is not a given. For instance, if the NIH leadership had not so aggressively ramped up the HGP in response to the gene patenting threat posed by Celera, Celera might have succeeded in patenting much of the human genome. In addition, while the stewardship of the NIH has been quite good so far, the director is politically appointed, so there is always the possibility that criteria other than merit might shape the agency's leadership in the future.

Ultimately, the reasons that have precluded the emergence of an anticommons thus far seem likely to hold for the foreseeable future. However, it is important to closely monitor the reasons advanced in this Article, for a shift in any of them might endanger the biomedical research domain. This Article's analysis of the biomedical research domain--particularly the role of the NIH in altering the behavior of upstream actors and the technological importance of certain classes of inventions with respect to downstream research--highlights the need for commentators to carefully consider the full context, not just the legal one, in which biotechnological innovation unfolds.

Footnotes

[a1] J.D., conferral expected May 2009, The University of Texas School of Law. Chester Shiu received his M.S. in Biomedical Informatics from Stanford University School of Medicine in 2006 and B.S. in Biomedical Computation from Stanford University in 2005. Before attending law school he performed research in molecular dynamics, bioinformatics, and oncology. The author

would like to thank Professors John Golden at The University of Texas School of Law and Russ Altman at Stanford University School of Medicine, as well as Brenda Simon, a Fellow at the Stanford Law School Center for Law and the Biosciences, for their comments and encouragement.

[1] Lois N. Magner, A History of the Life Sciences 133-44 (3rd ed. 2002).

[2] See id.

[3] For a summary of the legal discussion surrounding biotech patents, see David E. Adelman, A Fallacy of the Commons in Biotech Patent Policy, 20 Berkeley Tech. L.J. 985, 990-96 (2005). One synthesis of the scientific community's reaction and concerns is presented in a report commissioned by the National Academy of Sciences, which is also summarized by Adelman. Id. at 998-1001 (citing Nat'l Research Council of the Nat'l Acads. (NRC), Comm. on Intellectual Prop. Rights in the Knowledge-Based Econ., A Patent System for the 21st Century (Stephen A. Merrill et al. eds., 2004)). For a discussion of different patenting schemes specifically related to genes, see Laurie L. Hill, The Race to Patent the Genome: Free Riders, Hold Ups, and the Future of Medical Breakthroughs, 11 Tex. Intell. Prop. L.J. 221, 246-58 (2003).

[4] Michael A. Heller, The Tragedy of the Anticommons: Property in the Transition from Marx to Markets, 111 Harv. L. Rev. 621 (1998).

[5] Id. at 622.

[6] This Article uses the terms biotech and biomedical interchangeably. Strictly speaking, biotechnology is the application of biomedical knowledge, but the literature does not always make this distinction.

[7] Michael A. Heller & Rebecca S. Eisenberg, Can Patents Deter Innovation? The Anticommons in Biomedical Research, 280 Science 698, 699 (1998).

[8] Id.

[9] See David E. Adelman & Kathryn L. DeAngelis, Patent Metrics: The Mismeasure of Innovation in the Biotech Patent Debate, 85 Tex. L. Rev. 1677, 1729 (2007) ("We conclude that the lack of concentrated control, the rising number of patent applications, and the continuous influx of new patent owners suggest that overall biotechnology innovation is not being impaired by the growth in patents issued each year."); Timothy Caulfield et al., Evidence and Anecdotes: An Analysis of Human Gene Patenting Controversies, 24 Nature Biotechnology 1091, 1092 (2006) ("The [predicted] effects are much less prevalent than would be expected if [the] hypothesized [anticommons] mechanisms were in fact operating."); John P. Walsh et al., Effects of Research Tool Patents and Licensing on Biomedical Innovation, in Patents in the Knowledge-Based Economy 285, 285 (Wesley M. Cohen & Stephen A. Merrill eds., Nat'l Acad. Press 2003) (finding that patents on inputs to drug discovery have increased and that this increase has not substantially hindered drug discovery).

[10] See infra Part III.A.

[11] This Article defines anticommons-based business models as those that seek to profit from an anticommons, or whose behavior, if emulated by others, would collectively lead to an anticommons.

[12] The term "domain" is used in this Article to denote the entire set of discoveries, potential and already realized, in a particular field of research. This is not to be confused with the public domain, which is the entire set of all discoveries that the public may freely use. A particular domain may encompass discoveries inside and outside the public domain. This technical use of domain is intended to avoid confusion with the term "commons" (i.e., intellectual commons), which, to be consistent with the usage of anticommons, must refer to some subset of the public domain.

[13] Adelman, supra note 3, at 1022 ("[T]he redundancy and intricacy of biological processes allow for multiple lines of research that enable scientists to circumvent existing problem-specific patents.... [B]iological complexity ... affords many potential routes for engineering around existing patents."). The implications of this observation are discussed generally by Adelman. Adelman, supra note 3, at 1020-30.

[14] For an overview, see Eric S. Lander & Robert A. Weinberg, Genomics: Journey to the Center of Biology, 287 Science 1777 (2000).

[15] Jeremy M. Berg et al., Biochemistry 118-19 (5th ed. 2002) ("DNA ... [is a] linear polymer[] built up from similar units connected end to end .... The chain of sugars linked by phosphodiester bridges is referred to as the backbone of the nucleic acid .... Whereas the backbone is constant in DNA and RNA, the bases vary from one monomer to the next. Two of the bases are derivatives of purine--adenine (A) and guanine (G)--and two of pyrimidine-- cytosine (C) and thymine (T, DNA only) or uracil (U, RNA only) ....").

[16] Id.

[17] Leland H. Hartwell et al., Genetics: From Genes to Genomes 11 (1st ed. 2000) ("Today, a gene is recognized as a region of DNA that encodes a specific protein or a particular type of RNA.").

[18] Id. at 2 ("DNA molecules are assembled into chromosomes.... The entire collection of chromosomes in each cell of an organism is its genome.")

[19] William K. Purves et al., Life: The Science of Biology 220-21 (6th ed. 2002). For more detail, see also Bruce Alberts et al., Molecular Biology of the Cell, 3-13 (4th ed. 2002). There are well-known, important exceptions to the central dogma, but they are not relevant to this discussion.

[20] Purves et al., supra note 19, at 221 fig.12.3.

[21] Purves et al., supra note 19, at 234 ("[S]ome base substitution mutations may change the genetic message such that one amino acid substitutes for another in the protein."); see also Purves et al., supra note 19, at 234 illus. (describing synonymous and missense mutations). There are changes that can be made to DNA that do not affect the protein product, but that distinction is not pertinent to this discussion and will not be mentioned further.

[22] Typically one manipulates by interfering with the translation of the transcribed RNA into protein. This technique goes by several names, including RNA interference, siRNA, RNAi, antisense RNA, and gene silencing. For a general idea of how the technique works, see Purves et al., supra note 19, at 322 fig.17.12.

[23] Berg et al., supra note 15, at 41. For a technically rigorous introduction, see chapters 3-6.

[24] Hartwell et al., supra note 17, at 222 ("[G]ene expression: the process by which cells convert DNA sequence information to RNA and then decode the RNA information to the amino acid sequence of a polypeptide ....").

[25] Hartwell et al., supra note 17, at 223-24 ("[S]ince the accurate one-way flow of genetic information determines protein structure, mutations that change this information or obstruct its flow can have dramatic effects on phenotype.").

[26] This is an important break from the central dogma, but the technical ramifications are not important here. For a more technical description of transcription factors, see David S. Latchman, Transcription Factors: An Overview, 29 Int'l J. Biochemistry & Cell Biology 1305 (1997).

[27] For an example of a transcription factor that regulates itself, see Jeni Pinson et al., Positive Autoregulation of the Transcription Factor Pax6 in Response to Increased Levels of Either of Its Major Isoforms, Pax6 or Pax6(5a), in Cultured Cells, 6 BMC Dev. Biology 25 (2006), available at http:// www.biomedcentral.com/1471-213X/6/25.

[28] One example is transcription factors that modulate other genes in response to an environmental stress. Purves et al., supra note 19, at 272 fig.14.16 (depicting transcription factors (stress proteins) that alter the expression of genes A, B, and C). See Latchman, supra note 26 at 1306 tbl.1 (showing six transcription factor targets that may be found on various genes).

[29] Developmental genes are the canonical example of this type of relationship. For a graphical representation and associated explanation, see Purves et al., supra note 19, at 306 fig.16.14 (explanatory text is located above the gene cascade figure). However, note that the target genes of transcription factors do not necessarily code for other transcription factors.

[30] Purves et al., supra note 19, at 303 fig.16.9 (depicting a transcription factor cascade). Note that "inducers" are, for the purposes of this Article, transcription factors.

[31] Latchman, supra note 26, at 1310 ("Given the vital role of transcription factors in a wide variety of cellular processes, it is not surprising that alterations in these factors can result in human disease.") (citation omitted).

[32] Latchman, supra note 26, at 1310.

[33] NRC, Reaping the Benefits of Genomic and Proteomic Research: Intellectual Property Rights, Innovation, and Public Health 19, 19 (Stephen A. Merrill & Anne-Marie Mazza eds., Nat'l Acad. Press 2006), available at http:// www.nap.edu/openbook.php?isbn=0309100674.

[34] Id.

[35] Id.

[36] Id.

[37] Cf. id. at 35 ("The shift from traditional, hypothesis-driven, small-laboratory, one-gene-, one-protein-at-a-time science to this new data-driven, large-scale engineering program initially engendered resistance in the molecular genetics community."). This reference discusses gene sequencing, but the same conceptual framework applies to gene-environment studies. Cf. id. at 38 ("As these cutting-edge technologies are introduced and an increasing number of causal relationships are known, the field of diagnostics will move from its current focus on single genes to a search of all the genes responsible for a particular disease.").

[38] This type of experiment requires a microarray. Invented in the 1990s, microarrays can test for the presence of up to approximately 30,000 genes simultaneously. The results of the HGP help in two ways: figuring out which genes to put on a microarray and analyzing the results afterwards. For an overview of microarray technology, see Nat'l Human Genome Research Inst., DNA Microarray Fact Sheet, http://www.genome.gov/page.cfm?pageID=10000533 (last visited Jan. 8, 2009); see also David W. Mount, Bioinformatics: Sequence and Genome Analysis 519 (1st ed. 2001) ("From this analysis, a set of genes that responds in an identical manner may be identified.").

[39] See Stanley N. Cohen et al., Construction of Biologically Functional Bacterial Plasmids In Vitro, 70 Proc. Nat'l Acad. Sci. 3240, 3244 (1973) ("The general procedure described here is potentially useful for insertion of specific sequences from prokaryotic or eukaryotic chromosomes or extrachromosomal DNA into independently replicating bacterial plasmids."). This was the seminal paper in the early development of recombinant DNA technology.

[40] Hartwell et al., supra note 17, at 167 ("Techniques for designing and constructing recombinant DNA molecules and for harnessing

bacteria to produce large quantities of a particular gene and its protein product hold great promise for medicine, agriculture, and industry.").

[41] Sally Smith Hughes, Making Dollars Out of DNA: The First Major Patent in Biotechnology and the Commercialization of Molecular Biology, 1974-1980, 92 Isis 541 (2001); see Irving S. Johnson, Human Insulin From Recombinant DNA Technology, 219 Science 632 (1983).

[42] See Tests Begin on Insulin Synthesized from Bacteria Through Gene-Splicing, N.Y. Times, July 24, 1980, at D18 ("Insulin currently used by diabetics is derived from the pancreas of animals, usually pigs or cows. In some cases, human patients can develop an infection with the animal insulins, a situation scientists hope can be avoided with the new insulin.... [The new insulin] should be available in much larger amounts than [animal insulin] ....").

[43] Eugene Russo, Special Report: The Birth of Biotechnology, 421 Nature 456 (2003), available at http:// www.nature.com/nature/journal/v421/n6921/pdf/nj6921-456a.pdf.

[44] This Article will not concern itself with biotechnology as applied to agriculture, which has its own set of statutory regulations. See generally Plant Variety Protection Act, 7 U.S.C. §§2321-582 (2006). This Article will also not consider medical or surgical procedure patents; for a discussion of those, see Leisa Talbert Peschel, Ph.D., Revisiting the Compromise of 35 U.S.C. §287(c), 16 Tex. Intell. Prop. L.J. 299 (2008).

[45] See Bernard R. Glick & Jack J. Pasternak, Molecular Biotechnology: Principles and Applications of Recombinant DNA 5-7 (3d ed. 2003).

[46] For a comprehensive introduction to bioinformatics, see Mount, supra note 38.

[47] NRC, supra note 33, at 38-39.

[48] NRC, supra note 33, at 38-39.

[49] See Francis S. Collins, Director, Nat'l Human Genome Research Inst., Address Before the World Economic Forum: A Brief Primer on Genetic Testing (Jan. 24, 2003) (transcript available at http://www.genome.gov/page.cfm? pageID=10506784). A particular biological system's genes are often expressed in response to particular external environmental factors. Upinder S. Bhalla & Ravi Iyengar, Emergent Properties of Networks of Biological Signaling Pathways, 283 Science 381, 386 (1999).

[50] See Nat'l Cancer Inst., Fact Sheet: Genetic Testing for BRCA1 and BRCA2: It's Your Choice (2002), http:// www.cancer.gov/cancertopics/factsheet/Risk/BRCA (follow "View/Print PDF" hyperlink under "Page Options").

[51] NRC, supra note 33, at 37.

[52] For background information on gene therapy, see generally Purves et al., supra note 19, at 347; see also Nat'l Cancer Inst., Fact Sheet: Gene Therapy for Cancer: Questions and Answers (2006), http:// www.cancer.gov/cancertopics/factsheet/Therapy/gene (follow "View/Print PDF" hyperlink under "Page Options").

[53] See supra Part II.A for an explanation of the central dogma.

[54] NRC, supra note 33, at 37 ("To date, success has been limited to a small number of relatively special cases.").

[55] NRC, supra note 33, at 37.

[56] Supra Part II.A.

[57] Supra Part II.A.

[58] NRC, supra note 33, at 37 ("[T]he small-molecule drug imatinib (Gleevec®/Glivec®) ... has yielded spectacular results in patients who would otherwise have died within a few years of diagnosis.").

[59] Richard D. Klausner, Nat'l Cancer Inst., The Nation's Investment in Cancer Research--A Plan and Budget Proposal for Fiscal Year 2003, at 76 (2001), available at http://plan2003.cancer.gov/pdf/bypass.pdf ("Drs. Peter Nowell and David Hungerford, two Philadelphia-based physicians, made a curious discovery. They noticed that cells from CML patients were missing a short segment on one member of the 22nd pair of chromosomes.").

[60] Id.
This shortened chromosome became known as the "Philadelphia chromosome" .... While the link between the Philadelphia chromosome and CML led scientists to suspect a causal relationship, the location of the missing DNA from chromosome 22 and how it might lead to CML was a mystery to be solved over the next three decades.
Id.; Elisabeth Buchdunger & Juerg Zimmerman, The Story of Gleevec, Innovation.org (Pharm. Research & Mfrs. of Am.), http:// www.innovation.org/index.cfm/StoriesofInnovation/InnovatorStories/The_Story_of_ Gleevec/ (last visited Jan. 10, 2009) ("[I]n 1960 ... scientists at the University of Pennsylvania noticed that one chromosome in the blood cells of many CML patients was shorter than normal--it was missing a big chunk of its DNA.").

[61] A timeline is provided in the 2003 fiscal year report by the National Cancer Institute:
In the early 1970s, new staining techniques offered a way to more precisely visualize band patterns, characteristic markings that can be used to identify individual chromosomes. With this technique, Dr. Janet Rowley determined that chromosome 9 in CML patients was lengthened by the same amount that chromosome 22 was shortened. From this observation, Rowley proposed that the genetic material from the two chromosomes was reciprocally exchanged, or 'translocated.'
Klausner, supra note 59, at 76. The results were published in Nature. See Janet D. Rowley, A New Consistent Chromosomal Abnormality in Chronic Myelogenous Leukaemia Identified by Quinacrine Fluorescence and Giemsa Staining, 243 Nature 290 (1973).

[62] Brian J. Druker, Translation of the Philadelphia Chromosome into Therapy for CML, 112 Blood 4808, 4809 fig.1 (2008), available at http:// bloodjournal.hematologylibrary.org/cgi/reprint/112/13/4808.pdf; Klausner, supra note 59, at 76 ("On the shortened end of chromosome 22, the genetic rearrangement produces the abnormal bcr-abl gene, the source of CML development."). The results were reported to the National Academy of Sciences. See generally Anne-Marie Mes-Masson et al., Overlapping cDNA Clones Define the Complete Coding Region For the P210c-abl Gene Product Associated with Chronic Myelogenous Leukemia Cells Containing the Philadelphia Chromosome, 83 Proc. Nat'l Acad. Sci. 9768 (1986).

[63] Druker, supra note 62, at 4809-10, 4809 fig.2; Klausner, supra note 59, at 76 ("Several laboratories confirmed the link between the defective gene and CML through studies showing that the bcr-abl gene was all that was needed to induce leukemia in mice."); Klausner, supra note 59, at 77 (noting that the BCR-ABL protein is the "product of the bcr-abl gene").

[64] Druker, supra note 62, at 4809 ("[BCR-ABL induced] defects include increased proliferation or decreased apoptosis of a hematopoietic stem cell or progenitor cell leading to a massive increase in myeloid cell numbers....").

[65] Druker, supra note 62, at 4810 ("[A]ll of the transforming functions of BCR-ABL are dependent on its tyrosine kinase activity.").

[66] Klausner, supra note 59, at 77.
American oncologist Dr. Brian Druker was interested in determining how the [BCR-ABL] protein, the product of the bcr-abl gene, fits into the complicated circuitry of cell signaling. His research led him to believe that the [BCR-ABL] protein could be a

powerful target for a drug that could impede the activity of the protein and be an effective treatment for CML. When he learned about Ciba-Geigy's complementary research, Druker asked scientists there for candidate protein kinase inhibitors that he could test against leukemia cells.
Klausner, supra note 59, at 77.

67 Druker, supra note 62, at 4810 fig.3.

68 The following timeline is taken from Buchdunger & Zimmerman, supra note 60.

69 Buchdunger & Zimmerman, supra note 60.

70 Druker, supra note 62, at 4811 ("A standard dose-escalation phase 1 study of imatinib began in June 1998."); Buchdunger & Zimmerman, supra note 60.

71 See Buchdunger & Zimmerman, supra note 60.

72 Buchdunger & Zimmerman, supra note 60.

73 See Buchdunger & Zimmerman, supra note 60.

74 Buchdunger & Zimmerman, supra note 60.

75 Druker, supra note 62, at 4811 ("Imatinib ultimately emerged as the lead compound for preclinical development based on its selectivity against CML cells in vitro and its druglike attributes, including pharmacokinetic and formulation properties."); Buchdunger & Zimmerman, supra note 60.

76 Druker, supra note 62, at 4810.
Thus, the period from 1960 to 1990 identified BCR-ABL as an ideal therapeutic target in CML. It is expressed in all patients with CML and it has been shown to be the cause of CML. BCR-ABL functions as a constitutively activated tyrosine kinase and mutagenic analysis has shown that this activity is essential for the transforming function of the protein.
Druker, supra note 62, at 4810. Furthermore, even with 30 years of scientific study, imatinib still took 6 years (1992 to 1998) to go from the lab to clinical trials. Buchdunger & Zimmerman, supra note 60 ("The compound that would become Gleevec was synthesized in 1992.... The first Phase I study began in June 1998.").

77 See supra notes 71-75 and accompanying text.

78 Research-produced ideas, principles, laws of nature, or the like are not patentable. See Diamond v. Diehr, 450 U.S. 175 (1981); Funk Bros. Seed Co. v. Kalo Inoculant Co., 333 U.S. 127 (1948). However, processes used in research and development are patentable subject matter. See 35 U.S.C. §101 (2006).

79 Carbice Corp. of Am. v. Am. Patents Dev. Corp., 283 U.S. 27, 31 (1931).

80 Id.

81 Heller & Eisenberg, supra note 7, at 698.

[82] Heller & Eisenberg, supra note 7, at 698.

[83] Supra notes 7, 8, and accompanying text.

[84] Supra notes 59-75.

[85] This was by no means a given. In biomedical research one often finds that many or none of the studied genes end up making the crucial difference, and this sort of uncertainty is the only certainty when working in this field. See Adelman, supra note 3, at 988.

[86] See NRC, supra note 33, at 73 (explaining that databases that aggregate ESTs were what made ESTs valuable).

[87] Marina Chicurel, Bioinformatics: Bringing It All Together, 419 Nature 751, 751 (2002), available at http:// www.nature.com/nature/journal/v419/n6908/pdf/419751a.pdf.
[A]ll of those involved want more integration--integration of data across the hundreds, if not thousands, of different databases, and visual integration of data to aid interpretation.... "To answer most interesting biological problems, you need to combine data from many data sources," agrees Russ Altman, a biomedical informatics expert at Stanford University.
Id.; Suneeta D'Souza, Gene Meets Machine: Intellectual Property Issues in Bioinformatics, 12 Health L. Rev. 34, 39 (2004), available at http:// www.law.ualberta.ca/centres/hli/pdfs/hlr/v12_2/5_Souza.pdf ("Why is information sharing so important in bioinformatics? In the case of bioinformatics databases, it is by widespread and complete disclosure of information that data sets are likely to be enriched, verified and analyzed as more and more researchers continue to sift through such information.").

[88] Nat'l Insts. of Health, About NIH, http:// www.nih.gov/about/index.html (last visited Feb. 15, 2009).

[89] Id.

[90] Am. Ass'n for the Advancement of Sci., Historical Data on Federal R&D, FY 1976-2009 tbl.2 (2008) http://www.aaas.org/spp/rd/hist09p2.pdf; see Nat'l Insts. of Health, The NIH Almanac--Organization, http:// www.nih.gov/about/almanac/organization/index.htm (last visited on Mar. 27, 2009).

[91] See Nat'l Insts. of Health, supra note 90.

[92] See Nat'l Insts. of Health, Nat'l Human Genome Research Inst., Important Events in the History of NHGRI and the Human Genome Project, http:// www.nih.gov/about/almanac/organization/NHGRI.htm (last visited Mar. 27, 2009).

[93] Deepak Hegde & David C. Mowery, Politics and Funding in the U.S. Public Biomedical R&D System, 322 Science 1797, 1797 (2008).

[94] Hamilton Moses III, E. Ray Dorsey, David H. M. Matheson & Samuel O. Thier, Financial Anatomy of Biomedical Research, 294 J. Am. Med. Ass'n 1333, 1335-36 tbl.1 (2005); Rebecca S. Eisenberg, Patents and Data-Sharing in Public Science, 15 Indus. & Corp. Change 1013, 1015 (2006).

[95] See Nat'l Sci. Found., Science and Engineering Indicators 2006 app. tbl.4-9 available at http://www.nsf.gov/statistics/seind06/append/c4/at04-09.pdf (last visited Mar. 27, 2009); see also David Blumenthal et al., Relationships Between Academic Institutions and Industry in the Life Sciences-- An Industry Survey, 334 New Eng. J. Med. 368, 369 (1996) (stating that industry funds 11.7% of academic research in the life sciences and about 7% of university research in science as a whole in the mid-1990s).

[96] NRC, supra note 33, at 104 fig.4-2. Note that patents here refer to gene related technologies. Anything broader would result in too nebulous a patent query.

[97] NRC, supra note 33, at 102 ("The government also has an 'interest' in as many as 60 percent of the patents held by the leading academic institutions....").

[98] Nat'l Insts. of Health, NIH Office of Technology Transfer Annual Report Fiscal Year 2006, at 6 (2006), available at http:// www.ott.nih.gov/about_nih/AnnualReport-FY2006.pdf (noting that the Monitoring and Enforcement Branch oversaw 1,364 active licenses in 2006); see Claire T. Driscoll, Director, Tech. Transfer Off., Nat'l Human Genome Research Inst., Evolving NIH Patent & Licensing Policies and Practices for Genomic Inventions in the Post-Human Genome Project Era, http:// dnapatents.georgetown.edu/resources/ClaireDriscollNHGRINIHPresUofPenn030303., (last visited Feb. 28, 2009). The total number of patents calculated by adding all issued and pending applications since 2003 to the 2002 figures; given time it takes to successfully prosecute a patent, this seems like a reasonable approximation.

[99] NIH Office of Technology Transfer Statistics, http:// ott.od.nih.gov/about_nih/statistics.html (follow "Top 20 Commercially Successful Inventions" hyperlinks by fiscal year) (last visited Mar. 27, 2009).

[100] See Lori Pressman et al., The Licensing of DNA Patents by US Academic Institutions: An Empirical Survey, 24 Nature Biotechnology 31, 36 tbl.3 (2006), available at http:// www.nature.com/nbt/journal/v24/n1/abs/nbt0106-31.html (subscription required).

[101] See NIH Office of Technology Transfer, Policy, http:// ott.od.nih.gov/policy/policy.html (last visited Mar. 27, 2009).

[102] Lori Pressman et al., supra note 100 at 32.

[103] Best Practices for the Licensing of Genomic Inventions: Final Notice, 70 Fed. Reg. 18,413, 18,413-15 (Apr. 11, 2005).

[104] Id. at 18,415.

[105] NIH Office of Extramural Research, NIH Data Sharing Policy and Implementation Guidance, http://grants.nih.gov/grants/policy/data_sharing/data_ sharing_guidance.htm (last visited Jan. 8, 2009).

[106] See Pressman et al., supra note 100, at 35.

[107] NRC, Sharing Laboratory Resources: Genetically Altered Mice 15 (Nat'l Acad. Press 1994), available at http://www.nap.edu/openbook.php?record_ id=9156 ("Several hundred mouse stocks containing mutations have been used for a long time as models of human disease and for the study of metabolic processes.").

[108] See id. ("Alternatively, mutations can be targeted at specific loci to produce 'knockout mice[,]' which do not express the gene, or otherwise to alter gene expression.").

[109] See id.

[110] Christopher P. Austin et al., The Knockout Mouse Project, 36 Nature Genetics 921 (2004).

[111] Eliot Marshall, A Deluge of Patents Creates Legal Hassles for Research, 288 Science 255, 256-57 (2000). The technique was developed by Brian Sauer at DuPont. See Brian Sauer, Manipulation of Transgenes by Site-Specific Recombination: Use of Cre

Recombinase, in 225 Methods in Enzymology 890, 890-900 (Paul M. Wassarman & Melvin L. DePamphilis eds., 1993).

112 Gary W. Matkin, Associate Dean, Univ. Extension, Univ. of Cal. Berkeley, University Intellectual Property Management in the 20th Century: How Did We Get Here and Where Are We Going?, Presentation for the Conference on Research and Development and Economic Growth in the 20th Century, at 11 (Mar. 26-28, 1999) (transcript available at http:// cshe.berkeley.edu/events/randdconference1999/papers/matkin.doc); Eliot Marshall, The Mouse That Prompted a Roar, 227 Science 24, 24-25 (1997).

113 Heller & Eisenberg, supra note 7.

114 Matkin, supra note 112, at 11.

115 Eliot Marshall, NIH, DuPont Declare Truce in Mouse War, 281 Science 1261, 1262 (1998) ("DuPont has said that all researchers who receive federal funding--not just those who work at NIH--will be covered by the liberal rules, effectively freeing up the nonprofit world.").

116 Principles and Guidelines for Recipients of NIH Research Grants and Contracts on Obtaining and Disseminating Biomedical Research Resources: Final Notice, 64 Fed. Reg. 72,090, 72,091 (Dec. 23, 1999).

117 Rebecca Eisenberg, who advised the NIH on possible reforms after the Cre-lox incident, has stated as much in her own writing. Arti K. Rai & Rebecca S. Eisenberg, Bayh-Dole Reform and the Progress of Biomedicine, Law & Contemp. Probs., Winter/Spring 2003, at 289, 296 n.40. ("The controversy over the cre-lox technology was resolved only after [the] NIH threatened to boycott the technology unless DuPont allowed broader use by academic licensees.").

118 NRC, supra note 107, at 15 ("Recently, scientists have gained the ability to create mutant strains by integrating foreign DNA, with increasing specificity, into the genome of mouse cells. Such transgenes can be added at random to the genome. Offspring of the resulting mouse express the trait coded for by the foreign DNA.").

119 NRC, supra note 107, at 22 ("Leder and his collaborators developed the oncomouse by inserting the MYC oncogene attached to a specific promoter into the embryo of a normal mouse. The promoter is a mouse mammary-tumor virus-promoter that is expressed directly in breast epithelial cells, and mice with this oncogene reproducibly develop breast cancer.").

120 See Press Release, Nat'l Insts. of Health, NIH and E.I. DuPont Sign OncoMouse® Agreement (Jan. 19, 2000), available at http:// www.nih.gov/news/pr/jan2000/od-19.htm.

121 Id.

122 Marshall, supra note 111.

123 NRC, supra note 107, at 23.

124 Memorandum of Understanding Between E.I. DuPont de Nemours and Company and Public Health Service, U.S. Department of Health and Human Services (Jan. 18, 2000), available at http:// ott.od.nih.gov/pdfs/oncomouse.pdf. The agreement was signed by the NIH Office of Technology Transfer.

125 Pressman et al., supra note 100, at 35; Leon Rosenberg, Perspectives from Different Sectors: Major Pharmaceutical Company, in Intellectual Property Rights and the Dissemination of Research Tools in Molecular Biology 61, 63 (Nat'l Acad. Press 1997), available at http:// books.nap.edu/openbook.php?record_id=5758.

To some extent, intellectual property law helps to provide some rationality for the use of patented research tools. Damages generally cannot be collected from an infringer who is merely engaging in research. Typically, in fact, an inhouse research program is not sufficiently far along to know whether a lawsuit would actually protect valuable property or technology or even whether that property will ultimately prove to be of no value. Frankly, we all know that it is not good form to sue researchers in academic institutions and stifle their progress. Consequently, much potential litigation has been held in check, and we have not often had to confront the vexing issues that would arise in the litigation context. I hope that this rational forbearance will continue.
Id.

126 Rebecca S. Eisenberg, Reaching Through the Genome, in Perspectives on Properties of the Human Genome Project 209, 227 (F. Scott Kieff ed., 2003).

127 See Cohen et al., supra note 39, at 3244. The findings reported were supported by NIH grants AI08619 and GM14378, as well as grant GB-30581 from the National Science Foundation. Cohen et al., supra note 39, at 3244.

128 The breakthrough was reported in Science. See Mark Schena, Dari Shalon, Ronald W. Davis, & Patrick O. Brown, Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray, 270 Science 467 (1995). The findings reported were supported by NIH grants R21HG00450 and R37AG00198, NSF grant MCB 9106011, and the Howard Hughes Medical Institute (a not-for-profit private foundation). Id. at 470 n.12.

129 See J. Craig Venter Inst., About the J. Craig Venter Institute, http://www.tigr.org/a_history.shtml (last visited Mar. 27, 2009). There is a certain irony in the fact that the NIH was engaged in a race with a company founded with technology that it helped cultivate.

130 Walsh et al., supra note 9, at 331 (stating that companies, in particular, rarely sue universities for fear of the bad press that would ensue); Rosenberg, supra note 125, at 63 ("Frankly, we all know that it is not good form to sue researchers in academic institutions and stifle their progress. Consequently, much potential litigation has been held in check, and we have not often had to confront the vexing issues that would arise in the litigation context."). Rosenberg is a representative of Bristol Myers Squibb, which is a large pharmaceutical company. See Rosenberg, supra note 125, at 61.

131 Adelman, supra note 3, at 1000. The importance of scientific norms is also reflected in the reluctance of early university inventors to patent landmark biotech methods (for example, the Cohen-Boyer process for inserting DNA). See Arti K. Rai, Regulating Scientific Research: Intellectual Property Rights and the Norms of Science, 94 Nw. U. L. Rev. 77, 93-94 (1999) (noting the early absence of emphasis on property rights).

132 Marshall, supra note 112, at 24.

133 Karen Sachs et al., Causal Protein-Signaling Networks Derived from Multiparameter Single-Cell Data, 308 Science 523 (2005).

134 See Jonathan M. Irish et al., Mapping Normal and Cancer Cell Signalling Networks: Towards Single-Cell Proteomics, 6 Nature Rev. Cancer 146 (2006). A quick search of the USPTO database turned up U.S. Patent No. 6,190,877 (filed Dec. 27, 1999), which claims a method for using flow cytometry and selective cell sorting to analyze cancer cells. See U.S. Patent No. 6,190,877 (filed Dec. 27, 1999). Without commenting on the merits of the patent, I note that claims 25-27 of the patent are likely to read on the methods developed by Irish et al. See id. cols. 15-16.

135 A non-patent example of NIH activism is the ACS-PubChem controversy, where the American Chemical Society wanted to shut down PubChem, an NIH-funded public database that cross-references individual chemical structures with all known biomedical research on that structure. Jocelyn Kaiser, Chemists Want NIH to Curtail Database, 308 Science 774, 774 (2005), available at http://www.sciencemag.org/cgi/content/summary/308/5723/774a (subscription required).

136 See Merck KGaA v. Integra Lifesciences I, Ltd., 545 U.S. 193 (2005) (involving a patent infringement suit against the Scripps Institute as a defendant).

[137] Michael M. Hopkins, Surya Mahdi, Pari Patel & Sandy M. Thomas, DNA Patenting: The End of an Era?, 25 Nature Biotechnology 185, 187 (2007), available at http://www.nature.com/nbt/journal/v25/n2/abs/nbt0207-185.html.

[138] While the public may have been captivated by Celera's race against the HGP to sequence the human genome, it is ironic to note that Celera was able to proceed so quickly in part because it could cross-compare its own sequencing results against those released into the public domain by the HGP (it did not return the favor to the HGP). Kristen Philipkoski, Celera a Cinch in Genome Race, Wired, Jan. 11, 2000, http:// www.wired.com/science/discoveries/news/2000/01/33551 ("To achieve their latest milestone, Celera combined the 9 percent of the Human Genome Project's publicly available data with their own research. By continuing to validate their own research with the public project's data, Venter said Celera will achieve a seven to eight times redundancy of the human genome sequence.").

[139] Hopkins et al., supra note 137, at 186.

[140] Eisenberg, supra note 126, at 227.

[141] See Eisenberg, supra note 126, at 227. Generally, the NIH will not patent an invention if: no further R&D is needed (e.g., research tools), the invention covers a low public health priority, or patenting the invention will hinder technology transfer/access. Driscoll, supra note 98, at 9. Conversely, the NIH generally will patent an invention if: there is a high public health priority, patenting will facilitate access to technology, or patenting is necessary for investments in R&D. Driscoll, supra note 98, at 9.

[142] For a listing of the sheer breadth of fields covered by the NIH, see Nat'l Insts. of Health, supra note 90.

[143] See supra Parts III.A.1-3; see also Nat'l Sci. Found., supra note 95 (comparing annual research expenditures by the federal government versus private industry); Eisenberg, supra note 126, at 228 (stating that upstream research is mainly financed through public funding).

[144] Driscoll, supra note 98, at 36.

[145] Eisenberg, supra note 126, at 227 ("[S]ome private firms have tried to figure out business models for generating biomedical research information to provide platforms for downstream discovery, especially in genomics and bioinformatics. But often by the time private firms see such an opportunity, the upstream research has become relatively routine and mechanical....").

[146] Shotgun sequencing, invented by the founders of Celera, was the breakthrough that made it possible for a private entity to even contemplate sequencing an entire genome. Remember, the Human Genome Project, conceived before shotgun sequencing, was an international effort of unprecedented scale. See J. Craig Venter et al., Shotgun Sequencing of the Human Genome, 280 Science 1540, 1540 (1998); Justin Gillis & Rick Weiss, Private Firm Aims to Beat Government to Gene Map, Washington Post, May 12, 1998, §1 at 1 (LEXIS News Library) (discussing the Celera founders' plan to use shotgun sequencing to map entire genomes).

[147] Robert Mullan Cook-Deegan, The Alta Summit, December 1984, 5 Genomics 661, 661 (1989).
The principal conclusion of the meeting was, ironically, that methods were incapable of measuring mutations with sufficient sensitivity, unless an enormously large, complex, and expensive program were undertaken. Technical obstacles thus thwarted attainment of the main goal of the meeting, yet the meeting left a profusion of new ideas in its wake, some of which later washed ashore to be incorporated into various genome projects. Five years later, there is still no sensitive assay for human heritable mutations, but there are genome programs at NIH, at DOE, and in several foreign nations. (emphasis added).
Id.

[148] Andrew Pollack, Scientist Quits the Company He Led in Quest for Genome, N.Y. Times, Jan. 23, 2002, at C1, available at http:// query.nytimes.com/gst/fullpage.html?res=990CE3DF103BF930A15752C0A9649C8B63.
The brash challenge spurred the public project to step up its pace so that the fundamental blueprint of human life would not be

locked up by Celera, available only by subscription.... "The original business model of just being an information company or being the Bloomberg of biology, just wasn't enough," Mr. White [chairman and CEO of Applera, owner of Celera] said. One reason for this, analysts said, is that Celera was competing with the public project, which was giving the information away.
Id.

149 Id.

150 Nat'l Insts. of Health, What is KOMP?, http:// www.nih.gov/science/models/mouse/knockout/komp.html (last visited Mar. 1, 2009). The NIH does not readily state which knockout technologies are being employed. If Cre-lox is used, then in this context the reader can view KOMP as a means for the NIH to create these mice and absorb the Cre-lox licensing costs.

151 Nat'l Insts. of Health, Deltagen and Lexicon Knockout Mice and Phenotypic Data Resource, http:// www.nih.gov/science/models/mouse/deltagenlexicon/theresource.html (last visited Mar. 1, 2009).

152 See Nat'l Insts. of Health, supra note 150.

153 Marshall, supra note 111, at 257.

154 NRC, supra note 33, at 48-50; see also Lexicon Pharmaceuticals, Inc., About Us, http://www.lexpharma.com/about-us/index.html (last visited Mar. 1, 2009).
Using our proprietary gene knockout technology, we have systematically discovered the physiological and behavioral functions of nearly 5,000 genes. Through this initiative, known as the Genome5000™ program, we have identified and validated more than 100 drug targets and created a unique, growing clinical pipeline. Our goal is to enhance the lives of patients by discovering and developing breakthrough treatments for human disease.
Id. Note that the firm has changed its name from Lexicon Genetics to Lexicon Pharmaceuticals, Inc., reflecting its move downstream. Id.

155 John M. Golden, Biotechnology, Technology Policy, and Patentability: Natural Products and Invention in the American System, 50 Emory L.J. 101, 176 (2001).

156 See Human Genome Project Information, SNP Fact Sheet, http:// www.ornl.gov/sci/techresources/Human_Genome/faq/snps.shtml (last visited Mar. 1, 2009).

157 Leslie Roberts, SNP Mappers Confront Reality and Find It Daunting, 287 Science 1898, 1898 (2000) ("Indeed, what galvanized researchers a few years ago was the possibility of building a map with SNPs peppered along the genome as landmarks. With such a map, investigators could compare individuals with a disease to a control group to see whether their SNP patterns varied.").

158 See id. ("[NIH] officials, worried that academic researchers would be shut out of the field, started a crash program in 1998 to find 100,000 SNPs.").

159 Id. ("[Most SNPs] probably have no function but could provide valuable markers for gene hunters: If they lie close to a susceptibility gene, they are likely to be inherited along with it.").

160 See Int'l HapMap Project, About the HapMap, http:// www.hapmap.org/thehapmap.html.en (last visited Mar. 1, 2009) ("The Project is a collaboration among scientists and funding agencies from Japan, the United Kingdom, Canada, China, Nigeria, and the United States.").

161 See Int'l HapMap Project, Data Release Policy, http:// www.hapmap.org/datareleasepolicy.html.en (last visited Mar. 27, 2009) ("All data generated by the Project will be released into the public domain.").

[162] Int'l HapMap Consortium, A Second Generation Human Haplotype Map of Over 3.1 Million SNPs, 449 Nature 851, 851 (2007).

[163] See supra discussion in Part III.A.4.

[164] Nat'l Inst. of Gen. Med. Sci., Protein Structure Initiative Mission Statement, http:// www.nigms.nih.gov/Initiatives/PSI/Background/MissionStatement.htm (last visited Mar. 1, 2009).

[165] Nat'l Inst. of Gen. Med. Sci., Background Information, http:// www.nigms.nih.gov/Initiatives/PGRN/Background/ (last visited Mar. 1, 2009).

[166] NRC, supra note 33, at 41.

[167] See Purves et al., supra note 19, at 234.

[168] See Berg et al., supra note 15.

[169] NRC, supra note 33, at 41.

[170] See NRC, supra note 33, at 39 (explaining that high-throughput structure determination methods have not yet been developed meaning that determining a structure is extremely difficult and time consuming.).

[171] Nat'l Inst. of Gen. Med. Sci., Protein Structure Initiative Update (May 2007), http://www.nigms.nih.gov/News/Reports/psi2_update_052007.htm.

[172] 454 Life Sciences, Products & Solutions, http:// www.454.com/products-solutions/system-features.asp (last visited Mar. 1, 2009).

[173] Protein Structure Prediction Center, 7th Community Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction, Success in Predicting CASP7 Targets, http://predictioncenter.org/casp7/meeting_ docs/charts.html (last visited Mar. 1, 2009). CASP is the competition held among protein structure predictors to gauge progress in the field. Note the success rate on the charts, which is quite low overall; the successful predictions are for small structures of mixed importance.

[174] See Toni Gabaldón, Computational Approaches for the Prediction of Protein Function in the Mitochondrion, 291 Am. J. Physiology - Cell Physiology C1121, C1123 (2006).

[175] Press Release, Dep't of Energy's Joint Genome Inst., JGI Sequences "Supergerm" Genome in One Day (May 9, 2000), available at http:// www.jgi.doe.gov/News/news_5_9_00.htm. Of course, this glosses over the fact that a bacterial genome is much smaller than the human genome; however that milestone still illustrates the many orders of magnitude increase in sequencing capability that arose in just ten years.

[176] Nat'l Inst. of Gen. Med. Sci., NIGMS Protein Structure Initiative Meeting Summary (Apr. 24, 1998), http://www.nigms.nih.gov/News/Reports/protein_ structure.htm (last visited Mar. 1, 2009).

[177] Though not pertaining to patents, one area of controversy has been open access to biomedical databases in the bioinformatics context.
First, if there are intellectual property rights held right down to the level of the data, downstream problems are bound to occur in

relation to biologists who use and combine the data in the database, as well as for owners of technologies which underlie the database. Such database protection schemes may also result in researchers in the non-commercial sector being unable to afford access to such proprietary data, or to enforce access and use restrictions on such data. Furthermore, a protection scheme for databases may have the effect of causing scientists to become reluctant to share data in the same way they have done in the past, leading to a "anti-commons" problem in which vast amounts of information will be held in too many small pieces by too many people with too many rights.
D'Souza, supra note 87, at 40 (citations omitted). While an area of concern, the author is unaware of any reports of anticommons thickets forming. Also of note is the fact that the NIH supports numerous databases that are freely accessible, such as PubChem. See Nat'l Insts. of Health, PubChem, http:// pubchem.ncbi.nlm.nih.gov/ (last visited Mar. 1, 2009) (providing public access to "information on the biological activities of small molecules").

178 See Heller & Eisenberg, supra note 7, at 699 (stating that private firms patent newly identified DNA sequences before a biological function has been discovered).

179 See supra note 3. However, this is not to say that all commentators have focused solely on patent law. For example, Rai and Eisenberg have suggested NIH actions and modifications to the Bayh-Dole Act to avoid problems of overly aggressive patenting of research tools. Arti K. Rai & Rebecca S. Eisenberg, Bayh-Dole Reform and the Progress of Biomedicine, 66 Law & Contemp. Probs. 289, 310-13 (2003).

180 E.g., Marshall, supra note 111, at 257 (mentioning the NIH's efforts to negotiate less restrictive license terms with DuPont).

181 Eisenberg, supra note 126, at 226-27.

182 Reasonable minds can argue about exactly how many seminal inventions have occurred, but there is no doubt that the number is quite small. Adelman proffers three examples of his own. Adelman, supra note 3, at 1020 (noting that there are "a relatively small number of common methods (for example, Cohen-Boyer, Kohler-Milstein, and PCR) that are critical to a broad range of biotech research"). Fiona Murray and Scott Stern suggest five. Fiona Murray & Scott Stern, Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge?: An Empirical Test of the Anti-commons Hypothesis, 63 J. Econ. Behav. & Org. 648, 654 (2007) (noting recombinant DNA, discovery of the HIV retrovirus, the OncoMouse, RNA interference, and embryonic stem cells).

183 See, e.g., Murray & Stern, supra note 182, at 655 ("Economists, law and technology scholars, and policymakers have focused on key cases to highlight their concerns over the impact of this expansion in [IP rights].... Consider the Onco[M]ouse ....") (citations omitted).

184 Hopkins et al., supra note 137, at 186 ("Filing at the USPTO, EPO and JPO patent offices declined markedly once the human genome sequence was published in 2001....").

185 Indeed, though a quote from Eisenberg is used, her general awareness of the issues in this field is respected by the scientific community. Pressman et al., supra note 100, at 32 ("[T]hen-NIH director Harold Varmus invited Rebecca Eisenberg to chair a working group to recommend policies NIH might pursue to ensure maximum social benefit from NIH-funded inventions.").

186 A number of examples come to mind. Researchers had to license seventy patents in order to engineer rice that addresses vitamin A deficiency. Walsh et al., supra note 9, at 288 n.6. "Fourteen patents controlled by several organizations cover the hepatitis-B vaccine; thus, stacking royalties totaling $1.47 per dose, or thirteen to fifteen percent of sales, add to the cost of the vaccine's production." Adelman, supra note 3, at 998 (citing Walsh et al., supra note 9). Over 100 patents on adrenergic receptors have been identified. Heller & Eisenberg, supra note 7, at 699. But see Walsh et al., supra note 9, at 294-95 (discussing research done by R.K. Seide and J.M. MacCloud in response to Heller and Eisenberg showing that at most a handful of such patents would need to be licensed in order to conduct research).

187 See supra note 9.

[188] Pressman et al., supra note 100.

[189] Pressman et al., supra note 100, at 33-35.

[190] Pressman et al., supra note 100, at 33.

[191] Pressman et al., supra note 100, at 33.

[192] Pressman et al., supra note 100, at 34.

[193] Pressman et al., supra note 100, at 34.

[194] The intricacies of this surprising lack of value are explained in detail in the next section. See infra Part V.

[195] Pressman et al., supra note 100, at 34.

[196] See Pressman et al., supra note 100, at 34-35 (defining "research tools" as discoveries without known specific utilities or use in producing commercial products); see also Marlan D. Walker, The Patent Research Tool Problem After Merck v. Integra, 14 Tex. Intell. Prop. L.J. 1, 4-5 (2005).
Patented research tools are ubiquitous in the pharmaceutical industry.... They save valuable time and money for researchers. Such inventions include patented assays and procedures; patented cell lines; patented recombinant DNA constructs and methods; enzymes; DNA microarrays for high throughput drug screening; patented research animals; bioinformatic tools such as computer programs; DNA, protein, and combinatorial chemistry libraries; reagents; drugs and drug targets; and many other patented machines and apparatuses.
Id.

[197] See Nat'l Insts. of Health, Principles for Recipients of NIH Research Grants and Contracts on Obtaining and Disseminating Biomedical Research Resources: Request for Comments, 64 Fed. Reg. 28205 (May 25, 1999), available at http://ott.od.nih.gov/policy/rt_guide.html ("Although some licensing of research tools to for-profit companies is necessary and appropriate, the majority of transfers, to both not-for-profit and for-profit entities, should be implemented under terms no more restrictive than the [Uniform Biological Material Transfer Agreement]."); Pressman et al., supra note 100, at 34 ("Respondents frequently referred to the NIH guidelines" when describing "their policies and practices regarding the licensing of research tools.").

[198] 35 U.S.C. §112 (2006).

[199] 35 U.S.C. §§101-03 (2006).

[200] 35 U.S.C. §102 (2006).

[201] 35 U.S.C. §103 (2006).

[202] Pressman et al., supra note 100, at 34.

[203] See Univ. of Rochester v. G.D. Searle & Co., 358 F.3d 916 (Fed. Cir. 2004) (holding that merely describing a drug target, without more, does not meet the written description requirement).

[204] The University of Rochester court made this point while discussing COX-1 and COX-2, two targets claimed in a patent that otherwise did not satisfactorily describe anything that exploit those targets.
Even with the three-dimensional structures of enzymes such as COX-1 and COX-2 in hand, it may even now not be within the ordinary skill in the art to predict what compounds might bind to and inhibit them, let alone have been within the purview of one of ordinary skill in the art....Rochester and its experts do not offer any persuasive evidence to the contrary.
Id. at 925.

[205] Id. at 930.

[206] Id. at 926 (citing Hyatt v. Boone, 146 F.3d 1348, 1353 (Fed. Cir. 1998)).

[207] Id. at 926-27 ("[Rochester's] patent discloses nothing more than a hoped-for function for an as-yet-to-be-discovered compound....").

[208] See generally Pedro Cuatrecasas, Drug Discovery in Jeopardy, 116 J. Clinical Investigation 2837, 2837 (2006), available at http://www.pubmedcentral.nih.gov/articlerender.fcgi?&pubmedid=17080187#B1; Simon Frantz, FDA Publishes Analysis of the Pipeline Problem, 3 Nature Rev. Drug Discovery 379 (2004) (explaining that most drug candidates result in failure and never reach market).

[209] Druker, supra note 62, at 4810.
[T]he period from 1960 to 1990 identified BCR-ABL as an ideal therapeutic target in CML. It is expressed in all patients with CML and it has been shown to be the cause of CML. BCR-ABL functions as a constitutively activated tyrosine kinase and mutagenic analysis has shown that this activity is essential for the transforming function of the protein. For these reasons, an inhibitor of the BCR-ABL kinase would be predicted to be an effective and selective therapeutic agent for CML....
Druker, supra note 62, at 4810.

[210] See Adelman, supra note 3, at 988.

[211] Adelman & DeAngelis, supra note 9, at 1692.

[212] See Pressman et al., supra note 100, at 34-35 ("'If you know [a DNA sequence's] specific utility, it is not a research tool. If you are using the method to produce a commercial product, then it is not a research tool.'" (quoting Alan Paau from the University of California, San Diego)).

[213] In re Fisher, 421 F.3d 1365, 1376 (Fed. Cir. 2005).

[214] See, e.g., U.S. Patent No. 5,891,636 (filed Sept. 3, 1997); U.S. Patent No. 5,716,785 (filed Apr. 19, 1996); U.S. Patent No. 5,807,522 (filed June 7, 1995).

[215] If one could articulate the specific utility of the tool being claimed in the patent, it would not belong to the research tool category. See supra note 212. The Federal Circuit held in In re Fisher,
The claimed ESTs themselves are not an end of Fisher's research effort, but only tools to be used along the way in the search for a practical utility. Thus, while Fisher's claimed ESTs may add a noteworthy contribution to biotechnology research, our precedent dictates that [Fisher's] application does not meet the utility requirement of §101....
421 F.3d at 1376.

[216] In re Fisher, 421 F.3d at 1370.
[The government] argues that Fisher failed to meet that standard because Fisher's alleged uses are so general as to be meaningless. What is more, the government asserts that the same generic uses could apply not only to the five claimed ESTs but also to any EST

derived from any organism. It thus argues that the seven utilities alleged by Fisher are merely starting points for further research, not the end point of any research effort.... We agree....
Id.; Kevin T. Kelly, Fragging the Patent Frags: Restricting Expressed Sequence Tag Patenting Using the Enablement-Commensurate-in-Scope-with-the-Claims Requirement, 17 Tex. Intell. Prop. L.J. 49, 52-53 (2008) ("The court held that the two stated uses for the DNA sequences in the patent application-- as research tools to identify polymorphisms or to isolate promoters--were mere hypothetical possibilities.") (citation omitted).

217 See Kyle Jensen & Fiona Murray, Intellectual Property Landscape of the Human Genome, 310 Science 239, 239 (2005) (mentioning complementary DNA probes used on microarrays).

218 Pressman et al., supra note 100, at 34 ("Joyce Brinton of the Harvard University Office for Technology and Trademark Licensing wrote, 'Where disease-linked mutations that may be useful in clinical diagnostics assays are identified, they sometimes are patented; this decision depends somewhat on the diagnostic kit and service market, which is less robust than the therapeutics market.'")

219 See Univ. of Rochester v. G.D. Searle & Co., 358 F.3d 916, 926-27 (Fed. Cir. 2004) (finding that the written description requirement was not satisfied because the patent specification did not disclose "any of the compounds required to practice the claimed methods.").

220 See supra Part III.A.

221 John P. Walsh, Charlene Cho & Wesley M. Cohen, View from the Bench: Patents and Material Transfers, 309 Science 2002, 2002 (2005).

222 Id. at 2003.

223 Id. at 2002 (discussing Madey v. Duke Univ., 307 F.3d 1351, 1363-64 (Fed. Cir. 2002) (holding that there was no blanket experimental use exception for academic researchers in a university)).

224 Madey, 307 F.3d at 1363-64.

225 See supra note 125 and accompanying text.

226 Hopkins et al., supra note 137, at 186.

227 NRC, supra note 33, at 110.

228 See NRC, supra note 33, at 108.

229 See Hopkins et al., supra note 137, at 186 ("[T]here was a perception that patentability thresholds had risen substantially at the USPTO after the introduction of the US utility guidelines in 2001....").

230 Hopkins et al., supra note 137, at 186.

231 Int'l HapMap Consortium, supra note 162, at 851.

[232] Cf. Adelman, supra note 3, at 1003-04 (discussing general protein targets, which are far more difficult to substitute around than SNPs).

[233] Gene Patents and Other Genomic Inventions: Hearing Before the Subcomm. on Courts and Intellectual Property of the House Comm. on the Judiciary, 2000 WL 1005937 (2000) (statement of Jon F. Merz, Assistant Professor, University of Pennsylvania) ("Merz Statement") (noting that a pilot survey showed that 25% of physicians abandoned a clinical test they developed and 48% forewent developing a clinical test on account of patents); NRC, supra note 33, at 123.

[234] Jensen & Murray, supra note 217, at 239 (noting that 20% of human genes are claimed in 4,270 U.S. patents, with roughly 63% assigned to private firms).

[235] Jensen & Murray, supra note 217, at 239.

[236] The same study also notes Human Genome Sciences, which is another early pioneer of private gene patenting. Jensen & Murray, supra note 217, at 239.

[237] NRC, supra note 33, at 115 tbl.4-2.

[238] See Hopkins et al., supra note 137, at 186.

[239] NRC, supra note 33, at 109.

[240] Walsh et al., supra note 9, at 285. Two years after the study was released, the Supreme Court took a very broad view of the § 271(e)(1) safe harbor exemption in the Hatch-Waxman Act. Merck KGaA v. Integra Lifesciences I, Ltd., 545 U.S. 193, 205-06 (2005). This view might grant statutory immunity to some university research activities; for commentary, see Daniel Wobbekind, Integra Lifesciences I, Ltd. v. Merck KGaA: Re-examining the Broad Scope of the §271(e)(1) Safe Harbor, 23 Berkeley Tech. L.J. 107 (2008); see also Anna McMinn, Judicial Interpretation of 35 U.S.C. §271(e)(1): An Improper Expansion Beyond the Legislative Intent, 16 Alb. L.J. Sci. & Tech. 195 (2006); see also Jian Xiao, Carving Out a Biotechnology Research Tool Exception to the Safe Harbor Provision of 35 U.S.C. §271(e)(1), 12 Tex. Intell. Prop. L.J. 23 (2003).

[241] See supra Part III.A.2.

[242] See John R. Allison et al., Valuable Patents, 92 Geo. L.J. 435, 462 (2004) ("[I]f valuable patents can be reliably identified at the time of application, or at least at the time of issue, the lottery theory runs into difficulty. At best, it becomes only a partial explanation....").

[243] See, e.g., Eric G. Campbell et al., Inside the Triple Helix: Technology Transfer and Commercialization in the Life Sciences, 23 Health Aff. 64 (2004). For a consideration of the interplay between university licensing and research, see Gregory K. Sobolski et al., Technology Licensing: Lessons from the U.S. Experience, 294 J. Am. Med. Ass'n 3137 (2005).