

Will Gene Patents Impede Whole Genome Sequencing?: Deconstructing the Myth That 20% of the Human Genome Is Patented

Christopher M. Holman
Associate Professor of Law
University of Missouri – Kansas City

During oral arguments before the Court of Appeals of the Federal Circuit in *Association for Molecular Pathology v. US Patent and Trademark Office (AMP v. PTO)*, a high-profile lawsuit challenging the validity of so-called “gene patents,” Judge Bryson asked the attorney representing Myriad Genetics (owner of the specific patents under attack) whether the company's patents would be infringed by the sequencing of an individual's genome.¹ The attorney did not seem to have previously considered the question, but ventured that in his view whole genome sequencing (WGS) would not constitute infringement because Myriad's patent claims are directed towards isolated genes, not the entire chromosome.² Later during the proceedings, the ACLU attorney representing the plaintiffs begged to differ, offering up his opinion that Myriad's gene patents, particularly claims directed to “isolated DNA that codes for the BRCA polypeptide,” would in fact be infringed by WGS.³

The question of whether Myriad's patents would be infringed by WGS is entirely hypothetical - no one has ever been sued for infringing a human gene patent based on the sequencing of an individual gene, let alone an entire genome.⁴ But Judge Bryson is understandably concerned with the policy implications of *AMP v. PTO*, and his question reflects a widely held perception that gene patents threaten to substantially impede the implementation not only of WGS, but of other genetic diagnostic technologies that simultaneously test for variations in multiple human genes.

To a large extent, the fear that gene patents might prove an obstacle to WGS arises from a widely held belief that 20% of human genes are “patented,” and as a consequence those engaged in developing and implementing WGS will be forced to either obtain licenses under a large number of gene patents, which would be prohibitively expensive and time-consuming, or risk a substantial threat of being sued for patent infringement by a host of gene patent owners. For example, Judge Bryson specifically indicated that his inquiry into the preclusive effect of gene

¹ Oral argument in *Association for Molecular Pathology v. US Patent and Trademark Office*, Federal Circuit Docket Number 10-1406, available at <http://www.ca9c.uscourts.gov/oral-argument-recordings/search/audio.html>, at 11-14 minutes into the proceedings.

² *Id.*

³ *Id.*, at 34-35 minutes into the proceedings.

⁴ Christopher M. Holman, Trend in Human Gene Patent Litigation, *SCIENCE* 322:198-99 (2008); Christopher M. Holman, The Impact of Human Gene Patents on Innovation and Access: A Survey of Human Gene Patent Litigation, 76 *UMKC L. REV.* 295 (2007).

patents on personal WGS was prompted by the concern that sequencing a person's genome might require obtaining licenses to hundreds or even 1000 gene patents.⁵

The notion that access to 20% of human genes is effectively precluded by gene patents has become something of an urban legend, and is routinely cited as fact, often without any supporting reference.⁶ However, a review of the literature reveals that the idea stems from a single study by Jensen and Murray published in 2005 (referred to hereafter as "the Study") which purports to map the "intellectual property landscape of the human genome".⁷ In this article, I deconstruct the Study and explain why it in no way supports a conclusion that 20% of human genes are "patented," or that WGS will result in the infringement of a plethora of gene patents. To the contrary, my analysis of the claims in a large sampling of the patents identified in the Study as "gene patents" indicates that the fear that gene patents will impede WGS has in all likelihood been greatly overstated.

Deconstructing the Jensen/Murray Study

Although Jensen and Murray state in their article that "20% of human genes are *explicitly claimed* as US IP," their supporting online materials make clear that what they actually did was identify issued US patents in which a human gene sequence, or the protein encoded by human gene sequence, was *explicitly mentioned* in the patent claims.⁸ The critical distinction between a gene being "claimed" versus a gene being "mentioned" in a patent claim has apparently been lost on those who assume that the Study shows that 20% of human genes are patented. The mere fact that a patent claim mentions a gene sequence (or the protein encoded by it) in no way implies that the patent completely precludes access to the gene, nor does it warrant an assumption that sequencing the gene would result in patent infringement.

In order to better understand the actual significance of the Study, and more generally the potential preclusive effect of gene patents on WGS, I analyzed the claims from a random sampling of 533 of the 4270 patents identified as "gene patents" in the Study.⁹ The 533 selected patents span the entire temporal range of the Murray and Jensen data set, from some of the earliest patents, issued in 1993, to the most recent issued in 2005 (the year they conducted the

⁵ Oral argument in *Association for Molecular Pathology v. US Patent and Trademark Office*, at 11-14 minutes into the proceedings.

⁶ See, for example, Sam Kean, *The Human Genome (Patent) Project*, 331 Science 530 (2011)(making unsupported assertion that 20% of human genes are patented).

⁷ Kyle Jensen & Fiona Murray, *Intellectual Property Landscape of the Human Genome*, 310 Science 239 (2005).

⁸ More specifically, they conducted an automated search to identify all US patents reciting the canonical term "SEQ ID NO." in the claims, and wherein the "SEQ ID NO." term is used in conjunction with a specific genetic sequence corresponding to a known human gene. Jensen & Murray's Supporting Online Materials, available at <http://www.sciencemag.org/cgi/data/310/5746/239/DC1/1> (last visited March 26, 2010).

⁹ Kyle Jensen generously provided me with a spreadsheet listing all of the patents identified in their study, and granted me permission to post the data on the Internet. It is available through Google documents at https://spreadsheets.google.com/spreadsheet/ccc?key=0At_IJGo9WK0dG00dUh4WXINcm5UeFpKcTntMEhraVE&hl=en_US (last visited July 18, 2011).

Study).¹⁰ After reviewing the patent claims, I assigned the patents to three relevant (and sometimes overlapping) categories based on the nature of the subject matter encompassed by the claims.

To summarize, Category I comprises the substantial number of patents in the sampling that do not include a single claim that could under any circumstance be infringed by any form of genetic testing, including not only genetic testing that involves DNA sequencing, but more generally other methods, such as those that involve measuring gene expression by RNA hybridization. Category II comprises patents that include one or more product claims that purport to encompass a DNA molecule corresponding in sequence to at least some portion of a human gene. Category III comprises patents with one or more method claims that appear to have at least some potential for being infringed by some form of genetic testing, although not necessarily by DNA sequencing. A more detailed discussion of the patents falling within the three categories is provided below.

As a preliminary matter, it is important to recognize that interpreting patent claims outside the context of patent infringement litigation is a notoriously unpredictable process, and as a practical matter it is generally impossible to definitively state the precise scope of the subject matter covered by the claims in these patents. The vagaries of claim interpretation have been well documented, and are reflected in the high rate at which the Court of Appeals of the Federal Circuit reverses the claim interpretation rulings of district courts (and it is not uncommon for the Federal Circuit judges to disagree amongst themselves as to the proper interpretation of a claim).¹¹ Indeed, the ambiguity of the scope of coverage defined by patent has led some, including at times the Federal Trade Commission, to characterize patents as nothing more than "probabilistic" property rights.¹² The uncertainty is especially pronounced with respect to gene patents, which have rarely been litigated, particularly in the context of genetic testing, as a consequence of which there is little guidance to be found in the case law.¹³ This explains why Judge Bryson was unsure whether individual WGS would infringe Myriad's claims, and the two attorneys could plausibly give diametrically opposite answers - in fact, nobody knows for sure.

¹⁰ Supporting online materials for this article provide the entire list of patent identified in the Study, as well as a list of the patents I specifically analyzed. The sampling excludes patents assigned to Incyte, which were studied separately, as discussed below.

¹¹ David L. Schwartz, *Practice Makes Perfect? An Empirical Study of Claim Construction Reversal Rates in Patent Cases*, 107 Michigan Law Review 223 (2008); *Phillips v. AWH Corp.*, 415 F.3d 1303 (Fed. Cir. 2005) (en banc).

¹² Petition for Writ of Certiorari at 16-17, *FTC v. Schering-Plough Corp.*,

126 S. Ct. 2929 (2006) (No. 05-273) (citing Mark A. Lemley & Carl Shapiro, *Probabilistic Patents*, 19 J. Econ.

Perspectives 75 (2005)), available at

<http://www.ftc.gov/os/2005/08/050829scheringploughpet.pdf> (last visited July 12, 2011). See also Christopher M. Holman, *Do Reverse Payment Settlements Violate The Antitrust Laws?*, 23 SANTA CLARA COMPUTER & HIGH TECH. L.J. 489, 532-33 (2007).

¹³ Christopher M. Holman, *Trend in Human Gene Patent Litigation*, SCIENCE 322:198-99 (2008); Christopher M. Holman, *The Impact of Human Gene Patents on Innovation and Access: A Survey of Human Gene Patent Litigation*, 76 UMKC L. REV. 295 (2007).

Nonetheless, the scope of patent claims is delimited by their text, and language used to claim an invention can only be stretched so far. As explained below, it is clear from a facial reading of the claims that many of the "gene patents" identified in the Study would not be infringed by DNA sequencing under any plausible interpretation. In fact, to varying degrees there is reason to believe that few if any of the patents analyzed include claims that a court would necessarily find valid and infringed by all forms of DNA sequencing.

Category I: Patents That Would Clearly Not Be Infringed by Genetic Testing

140 of the 533 patents in the sampling fall into Category I, including 96 that should not be characterized as "gene patents" under even the most expansive interpretation of this amorphous but widely used nomenclature. These patents are directed primarily towards proteins and methods of using proteins, and were spuriously identified as gene patents by the methodology employed in the Study because the amino acid sequence of the protein was referenced in the claim by a "SEQ ID NO." that also happened to be associated with the gene encoding the protein.¹⁴

Category I also includes 44 patents that mention DNA sequences, and thus might plausibly be characterized as gene patents, but that clearly would not be infringed by any form of genetic testing. Some of these patents were flagged in the Study because they have claims directed towards cells or organisms that have been genetically modified by the introduction of a DNA sequence encoding a human gene.¹⁵ Others are directed towards a fusion gene (man-made recombinant constructs made by fusing together gene sequences in a manner that does not occur in nature), or a vaccine, or a specific complex between a DNA molecule and a protein.¹⁶ Some of the patents in this category are directed towards specific methods of using or analyzing DNA, but were classified as gene patents solely because they include a dependent claim covering use of the method in connection with a specific human gene sequence.¹⁷

While none of these 140 patents in Category I could be infringed by genetic testing or WGS, they were all tallied as gene patents in the Study, and as a result all of the genes

¹⁴ Since 1990, the PTO has required patent applicants to identify most protein and DNA sequences by means of a sequence ID number, using the canonical term "SEQ ID NO." The Study identified patents having claims that include the term "SEQ ID NO.," and wherein the SEQ ID NO. was associated with a known human gene. However, sometimes the same SEQ ID NO is used to refer not only to the gene sequence, but also the amino acid sequence of the protein encoded by the gene, which led to these spurious hits. An example of a patent misidentified as a gene patent in the Study owing to a claim identifying a protein sequence by SEQ ID NO. is US Patent number 5, 843,888 (Claim 1. A non-naturally occurring mutant human hemoglobin wherein the valine residue at position 96 of the alpha chain (SEQ ID NO: 1) is replaced by a tryptophan residue.)

¹⁵ See, e.g., US Patent Numbers 5,861,310 and 5,932,780.

¹⁶ See, e.g., US Patent Numbers 5,712,121; 6,537,594; and 5,670,621.

¹⁷ See, e.g., US Patent number 5,821,089 (Claim1. A method of incorporating an amino acid analog into at least one polypeptide produced by a cell selected from the group consisting of prokaryotic cell and eukaryotic cell comprising: providing a cell selected from the group consisting of a prokaryotic cell and eukaryotic cell; providing hypertonic growth media containing at least one amino acid analog selected from the group consisting of trans-4-hydroxyproline and 3-hydroxyproline; and contacting the cell with the growth media wherein the at least one amino acid analog is assimilated into the cell and incorporated into least one polypeptide.).

corresponding to the DNA and protein sequences referenced in the claims were classified as "patented," thus contributing to the myth that 20% of human genes are patented.

Category II: Patents with Product Claims Covering DNA Molecules

366 of the 533 patents fall into Category II, which comprises any patent that includes at least one product claim that covers a DNA molecule having a sequence that Jensen and Murray found to be associated with a naturally occurring human gene, or some portion thereof. These are the quintessential human gene patents that most people probably envision when they think of genes being patented, and many [presumably non-lawyers] have apparently jumped to the conclusion that such patents would necessarily be infringed by any testing or use of the "claimed" gene.¹⁸ But this is an unwarranted assumption. These patents claim molecules, not genes *per se*, and would only be infringed by a DNA sequencing methodology that necessarily entails "making or using" the DNA molecule defined by the claim.¹⁹

For example, claims to DNA molecules having the sequence of a naturally occurring human gene are almost invariably limited to forms of the DNA that are isolated, purified and/or recombinant (of these limitations, "isolated" appears to be by far the most commonly used). The inclusion of these limitations is critical - it is well-established that a DNA molecule cannot be claimed in a manner that would cover the molecule as it exists in nature, e.g., as a gene residing in its normal location on a human chromosome. DNA sequencing will only constitute infringement of a claim to an isolated DNA molecule if the sequencing methodology involves the isolation of the entire DNA molecule as defined by the claim. To my knowledge, after extensive research, no court has ever interpreted a claim to an isolated or purified DNA molecule so broadly, and such a claim has never been found infringed by DNA sequencing. In fact, there are judicial decisions that, while not directly on point, suggest that a court might very well interpret a claim to isolated DNA in a relatively narrow manner such that it would not be infringed by at least some forms of DNA sequencing, particularly certain next-generation technologies that do not require the amplification of DNA.²⁰

In normal parlance, isolation of a naturally occurring molecule arguably implies not only separating it from its native context, but doing so in a manner that results in some sort of preparation in which the isolated molecule is known to be present, and thus available for retrieval and further processing or use at a later time. A Dutch court recently took this approach in interpreting a claim broadly reciting an isolated gene, rejecting the patent owner's argument that

¹⁸ See, e.g., Statement of Dr. Mark Grodman, CEO of Bio-Reference Laboratories, Inc. before the House Judiciary Subcommittee on Courts, the Internet and Intellectual Property in Connection with its hearing on "Stifling or Stimulating-The Role of Gene Patents in Research and Genetic Testing" (October 30, 2007), available at <http://judiciary.house.gov/hearings/pdf/Grodman071030.pdf> (last visited May 3, 2011).

¹⁹ Christopher M. Holman, Gene Patents under Fire: Weighing the Costs and Benefits, book chapter in *Biotechnology and Software Patent Law: A Comparative Review on New Developments* (publication anticipated in 2011).

²⁰ See, e.g., *Synaptic Pharmaceuticals v. MDS Panlabs*, 265 F. Supp. 2d 452 (D.N.J. 2002).

the claim encompasses any DNA sequence removed from its natural environment, and concluding that “the average person skilled in the art would understand the term isolated DNA as DNA that has been retrieved from the cell (core) of an organism for further treatment in a manner as is usual in the relevant profession.”²¹ Based on this interpretation, this gene patent claim was found not to be infringed, illustrating that gene patents do not necessarily cover all recombinant uses of the claimed gene.²²

If a court were to interpret a limitation such as "isolated" in this relatively narrow manner, it seems doubtful that all forms of DNA sequencing would be found infringing. Large-scale DNA sequencing methodologies, such as those that would be applied to the sequencing of entire genomes, generally do not require the "isolation" of defined preparations of DNA. This is particularly the case with respect to some of the next-generation sequencing technologies. For example, the Pacific Biosciences gene sequencing technology relies on the observation of DNA synthesis as it occurs on an immobilized DNA polymerase, thus requiring no discernible isolation of defined DNA molecules.²³

In order for all modes of DNA sequencing to result in infringement of claims to isolated DNA molecules, the term "isolated" would have to be given a very broad interpretation, so as to basically encompass any form of DNA existing in a form distinct from its native environment. The Dutch court mentioned above rejected such a broad interpretation, but that does not mean that a US court might not reach a different outcome, at least with respect to some claims directed to isolated DNA molecules. In fact, the district court in *AMP v. PTO* adopted a very broad interpretation for "isolated" in the Myriad patent claims, apparently based on a definition of the term in the corresponding patent specifications.²⁴ But the court appears to have devoted little attention to that particular issue, and it is hard to say whether this definition would stand if challenged on appeal. Moreover, it does not necessarily mean that such a broad interpretation of "isolated" would be applied to DNA claims in general, particularly in different patents in which the specification defines the term "isolated" more narrowly, either explicitly or implicitly.

In any event, a broad interpretation of a term such as "isolated" could be a double-edged sword. In general, the broader a claim is interpreted the more likely it is to be invalidated by a

²¹ *Monsanto Technology LLC v. Cefetra BV*, Court of Justice of the European Union, Case C-428/08, discussed on Holman's Biotech IP Blog <http://holmansbiotechipblog.blogspot.com/2010/07/monsanto-v-cefetra-eu-court-of-justice.html>, last visited May 3, 2011.

²² See also Christopher M. Holman, *Learning from Litigation: What Can Lawsuits Teach Us About the Role of Human Gene Patents in Research and Innovation?*, 18 KAN. J.L. & PUB. POL'Y 215 (2009)(identifying multiple instances wherein an asserted human gene patent was found not to have been infringed).

²³ SMRT™ Technology, Pacific Biosciences Website, <http://www.pacificbiosciences.com/smrt-biology/smrt-technology> (last visited July 11, 2011).

²⁴ *The Association for Molecular Biology, et al. v. United States Patent and Trademark Office, et al.*, 702 F.Supp.2d 181 (SDNY 2010).

court during litigation, for failure to comply with one or more requirements of patentability.²⁵ For example, in order to be patentable a claimed invention must be novel and nonobvious. The actual legal analysis would be quite complex, but the fundamental principle is straightforward: in general, subject matter that would infringe a patent claim after the patent issues will invalidate the patent if it occurred prior to the date of invention.

This implies that a claim reciting isolated naturally occurring DNA, if interpreted very broadly, might prove vulnerable to invalidation based on activities that occurred prior to the patent, such as electrophoretic separation of genomic DNA fragments (e.g., in Southern blotting), or the creation of genomic DNA libraries in vectors such as cosmids. Both of these techniques were widely used and reported long before any of these gene patents were filed, and inherently involve the isolation of large fragments of genomic DNA from their native environment. A full-blown legal analysis is beyond the scope of this paper, but the interested reader is directed to some recent Federal Circuit decisions clearly establishing that prior art need not explicitly describe a claimed feature of an invention in order to invalidate a claim covering subject matter inherently described in the prior art.²⁶ The 2005 *In re Crish* decision is particularly on point, because it unambiguously establishes that a prior art reference need not disclose the sequence of a DNA molecule in order to invalidate by anticipation.²⁷

A very broad interpretation of terms such as "isolated" could also raise substantial validity issues under the enablement and written description requirements.²⁸ These doctrines have been increasingly invoked in recent years to prevent inventors from claiming "more than they invented," especially with respect to inventions based on fundamental biological discoveries.²⁹ It is very possible that if a claim were interpreted so broadly as to cover essentially any and all forms of a DNA molecule removed from its native context in the genome, a court would invalidate the claim under one of these doctrines.

Furthermore, even if a mode of DNA sequencing were found to necessarily entail the "isolation" (as that term is defined by a court) of a claimed DNA molecule, there are other limitations in many gene patent claims that would appear to preclude infringement by DNA sequencing. For example, many of the patents only include claims covering an isolated cDNA molecule encoding a full-length protein. Most genomic genes include introns, which would be present in any isolated DNA arising in the course of sequencing a human genome, and thus avoid infringement of claims limited to full-length cDNA. My review of hundreds of gene patents

²⁵ F Scott Keiff, *Patents for Environmentalists*, 9 WASH. U. J. L. & POL'Y 307, 310 (2002) (explaining Judge Giles S. Rich's famous statement that "the stronger a patent the weaker it is and the weaker a patent the stronger it is").

²⁶ Cases applying the doctrine of "inherent anticipation" include *In re Cruciferous Sprout Litigation*, 301 F.3d 1343 (Fed. Cir. 2002); *Schering v. Geneva Pharmaceuticals*, 339 F.3d 1373 (Fed. Cir. 2003); *SmithKline Beecham v. Apotex*, 403 F.3d 1331 (Fed. Cir. 2005)

²⁷ *In re Crish*, 393 F.3d 1253 (Fed. Cir. 2005) (finding patent claim invalid for lack of novelty even though prior art did not disclose sequence of claimed DNA).

²⁸ 35 USC 112 is the statutory basis for the enablement in written description requirements.

²⁹ *Amgen v. Chugai*, 927 F.2d 1200 (Fed. Cir. 1991); *Ariad v. Eli Lilly*, 560 F.3d 1566 (Fed. Cir. 2009).

indicates that most gene patents are based on the isolation and sequencing of cDNA, rather than genomic DNA, and thus should not be interpreted so broadly as to cover genomic DNA. The U.S. government apparently shares this view - during arguments before the Federal Circuit in *AMP v. PTO* an attorney representing the U.S. government stated that a "vast majority" of the claims to isolated DNA that have been issued by the PTO are to "cDNA, recombinant DNA, process claims and the like," and hence would not cover genomic DNA (and thus, by implication, genome sequencing).³⁰

In a large number of Category II patents, the DNA product claims only cover DNA encoding a full-length protein. Proteins are generally at least 100 (and typically several hundred or more) amino acids in length, which means that the claims would only cover DNA molecules at least 300 bases in length (and more often much longer). In practice, gene sequencing typically does not involve isolating the full-length gene as a single molecule, because only fragments of the full-length gene are physically sequenced. The sequence of the full-length gene is then determined by piecing together the sequence information for the fragments. For example, as Robert Cook-Deegan and I pointed out in an amicus brief filed in *AMP v. PTO*, conventional BRCA genetic testing involves amplifying and sequencing fragments of the full-length gene (amplicons), and hence does not appear to infringe Myriad's patent claims that only cover DNA encoding the full-length BRCA protein.³¹ Large-scale DNA sequencing protocols (of the type that would be used for WGS) can involve relatively short sequencing reads, which would only necessitate the isolation of relatively short fragments of genomic DNA, irrespective of how broadly the term "isolated" is interpreted. Yet another reason to believe that at least some forms of genomic DNA sequencing would not be found infringing of many Category II patents.

On their face, the broadest Category II claims would appear to be those that explicitly claim any polynucleotide comprising a fragment of a gene sequence. These claims often purport to cover not only the isolated fragment, but any isolated DNA molecule that includes within its sequence the claim fragment, such as a larger fragment, the full-length genomic DNA, or even potentially a large stretch of genomic DNA encompassing the gene. In some cases, the fragments are relatively long, extending many hundreds of bases and thus not likely to be infringed by DNA sequencing of relatively short fragments of genomic DNA, as discussed above in connection with claims directed to full-length coding sequences. However, a relatively small but still substantial number of Category II patents include claims directed to quite small fragments -

³⁰ Oral argument in *Association for Molecular Pathology v. US Patent and Trademark Office*, Federal Circuit Docket Number 10-1406, available at <http://www.cafc.uscourts.gov/oral-argument-recordings/search/audio.html>, at 55-56 minutes into the proceedings.

³¹ Brief of *Amici Curiae* Christopher M. Holman and Robert Cook-Deegan, *Association for Molecular Pathology v. US PTO*, Federal Circuit Docket Number 2010-1406, available at https://docs.google.com/fileview?id=0B9_IJGo9WK0OWVkyjg3NjtYzRjYi00ODIyLWIyMjAtYmJkZDQxMGZmYTZi&hl=en (last visited May 3, 2011).

for example, one of the patents purports to cover any isolated DNA molecule comprising any sequence of 10 or more contiguous bases appearing in the claimed gene sequence.³²

If interpreted broadly, these fragment claims might cover any form of genomic DNA sequencing that requires the isolation of fragments of genomic DNA as short as 10 or 15 bases in length. However, the sheer breadth of such a claim might also be its downfall - a recent article suggests that fragment claims of this type could be declared invalid based on anticipation by a large number of prior art DNA sequences falling within the scope of the claim.³³ The expansive breadth of the claim would also raise substantial issues of patentability under the enablement and written description doctrines, for claiming much more than the inventor actually invented.

In short, it is far from certain that any Category II patents would be infringed by all modes of DNA sequencing, particularly next-generation technologies that arguably do not involve isolation of the claimed DNA molecules. For a large number of these patents, particularly those directed towards cDNA and full-length genes, infringement by DNA sequencing would appear unlikely, if not entirely out of the question.

Category III: Patents with Method Claims Potentially Infringed by Genetic Testing

Category III comprises the 48 patents identified in the study that include one or more process claims that could potentially be infringed by at least some form of genetic testing (defined broadly to include all forms of genetic testing, including methodologies that do not involve determining the sequence of genomic DNA). There is some overlap between Category II and Category III - 21 of the patents analyzed include claims falling into both categories. And as was the case with Category II, it is facially apparent that some of the Category III patents could not be infringed by any form of genomic DNA sequencing (even though the claim might be infringed by some form of genetic testing that does not involve genomic DNA sequencing).³⁴

Most of the claims in the Category III included specific technical limitations that would appear to preclude infringement by at least some modes of DNA sequencing. For example, many of the claims are directed toward methods of testing for the presence of a specific genetic variation in an individual, and are limited to methods that involve determining its presence by detecting hybridization to a nucleic acid or probe. It seems likely a court would interpret many (if not all) of these claims as limited to methods of genetic testing that involve direct detection of the hybridization event, and thus not encompassing DNA sequencing methodologies that only incidentally involve DNA hybridization, e.g., in the course of PCR amplification.³⁵ The term

³² US Patent Number 5,559,023.

³³ Thomas B. Kepler, Colin Crossman and Robert Cook-Deegan, *Metastasizing patent claims on BRCA1*, *Genomics* (2010), doi:10.1016/j.ygeno.2010.03.003.

³⁴ For example, some of the patents only include claims reciting methods of analyzing the amount of RNA expressed from a gene, or other relatively limited methods that would not encompass DNA sequencing.

³⁵ See, e.g., US Patent Numbers 5,580,727 and 5,756,288.

"probe" is generally used to refer to a DNA molecule that incorporates a detectable label, such as a radioactive or fluorescent group, which would exclude much of the hybridization that occurs in the course of processing samples for DNA sequencing. Other claims in this category specifically require a PCR amplification step in order for there to be infringement, which would rule out DNA sequencing technologies that do not require PCR amplification, such as some next-generation technologies capable of sequencing a single DNA molecule.³⁶

On their face, the broadest Category III patents are those with claims that purport to cover detection of a particular genetic variation by any means, without any explicit methodological limitations. A recent article identified broad method claims of this type as the most difficult to design around in genetic diagnostic testing, apparently based on an assumption that they would necessarily be infringed by virtually any mode of diagnostic testing for the claimed genetic variation, regardless of the analytical technique used.³⁷ Notably, of the 533 patents analyzed, I only found twelve that included claims of this nature. This finding suggests that these broad method claims, which some have identified as particularly problematic for genetic testing and genomic sequencing, are actually quite rare, and probably only appear in the relatively few instances in which, at the time the claims are drafted, the gene had been shown to be directly correlated with a medically significant Mendelian phenotype.

There is a good chance that a court would interpret most (if not all) of these twelve patents in a manner that would require not only that an alleged infringer physically determine the sequence of the DNA molecule, but also that the same entity analyze the resulting sequence data for the presence of the claims genetic variation, in order for infringement to occur. For example, in all but one of the twelve patents there is a clause in the claim specifying that the genetic variation "indicates" the presence of a medically significant phenotype, such as, for example, a lower likelihood of being diagnosed with bipolar disorder.³⁸ In a recent case involving claims directed towards a method of personalized medicine, the court interpreted a very similar "indicates" clause as requiring that a doctor be "warned or notified" of the relevant phenotype in order for there to be infringement.³⁹

As a general matter, a finding of infringement requires that all steps in a claimed process are performed by a single entity.⁴⁰ In a situation where one entity simply sequences an individual's genomic DNA without recognizing the presence of the genetic variation, and a distinct and independent entity then analyzes the resulting genetic information for the existence

³⁶ See, e.g., SMRT™ Technology, Pacific Biosciences Website, <http://www.pacificbiosciences.com/smrt-biology/smrt-technology> (last visited July 11, 2011) and Clarke J, Wu HC, Jayasinghe L, Patel A, Reid A, Bayley H, *Continuous base identification for single-molecule nanopore DNA sequencing*, *Nature Nanotechnology* 4 (4): 265–270 (2009). doi:10.1038/nnano.2009.12

³⁷ Isabelle Huys et al. *Legal Uncertainty in the Area of Genetic Diagnostic Testing*, 27 *Nature Biotechnology* 903 (2009).

³⁸ US Patent Number 6,458,541.

³⁹ *Prometheus Laboratories v. Mayo*, 628 F.3d 1347 (Fed. Cir. 2010).

⁴⁰ *BMC Res., Inc. v. Paymentech, L.P.*, 498 F.3d 1373 (Fed. Cir. 2007), and *Muniauction, Inc. v. Thomson*, 532 F.3d 1318 (Fed. Cir. 2008).

of the variation, it might well be the case that no party is liable for infringement. This would occur, for example, under a scenario where one company determines an individual's genomic DNA sequence, and an independent healthcare provider identifies any clinically significant genetic variations that might be present.

Only one of the twelve patents, U.S. Patent Number 6,432,644, has a broad method claim that does not include a limitation specifying that the presence of the claimed genetic variation “indicates” a likelihood of some specific phenotype. Claim 1 of this patent purports to claim any “method for diagnosing the presence of a polymorphism in human KCNE1 . . . which causes long QT syndrome wherein said method is performed by means which identify the presence of said polymorphism.” Claim 3 of the patent depends from Claim 1 and specifically recites that presence of the polymorphism is determined by sequencing the KCNE1 gene. It is impossible to state with confidence exactly how a court would interpret this claim, but it seems likely that an entity that sequences the KCNE1 gene in the course of sequencing an individual's genome, but does not specifically look for the presence of the polymorphism (which will in any event not exist in the majority of individuals not afflicted by the disease), and thus does not engage in any diagnosis, would not be found infringing.

Furthermore, the broader the Category III method claims are interpreted, the more vulnerable they will be to invalidation, for reasons analogous to those discussed above in connection with broad Category II claims. The Federal Circuit recently provided an example of this in *Billups-Rothenberg v. Associated Regional and University Pathologists*,⁴¹ which to my knowledge is the first human gene patent case involving an allegation of infringement by genetic testing that has actually resulted in a substantive judicial decision.⁴² In *Billups-Rothenberg*, the court found all of the asserted claims to be invalid, either because of anticipation by prior art, or for violation of the written description requirement. This is a good example of a point Robert Cook-Deegan and I attempted to make in our amicus brief filed in connection with AMP v. PTO - many gene patent claims relevant to genetic testing would face substantial validity issues if their owners ever attempt to enforce them in court.⁴³ In order to cover all modes of DNA sequencing, these gene patent claims would have to be interpreted very broadly, but the broader the claim that more vulnerable it is to invalidation. This might be part of the reason why, to date, there has been so little gene patent litigation in the context of genetic testing.

⁴¹ 2011 WL 1601996 (Fed. Cir. 2011), decided April 29, 2011, Federal Circuit Docket No. 2010-1401.

⁴² Christopher M. Holman, *Trend in Human Gene Patent Litigation*, SCIENCE 322:198-99 (2008); Christopher M. Holman, *The Impact of Human Gene Patents on Innovation and Access: A Survey of Human Gene Patent Litigation*, 76 UMKC L. REV. 295 (2007).

⁴³ Brief of *Amici Curiae* Christopher M. Holman and Robert Cook-Deegan, Association for Molecular Pathology v. US PTO, Federal Circuit docket number 2010-1406, available at https://docs.google.com/fileview?id=0B9_IJGo9WK0OWVvkYjg3NjItYzRjYi00ODIyLWIyMjAtYmJkZDQxMGZmYTZi&hl=en (last visited May 3, 2011)

As explained above, method claims that require both analysis of DNA molecules and analysis of DNA sequence would very likely not be infringed by an entity whose only action is limited to analysis of DNA, such as by performing WGS, if that entity does not also engage in analysis of the sequence. By the same token, an entity such as a doctor or genetic counselor who only analyzes the sequence data, without physically sequencing the DNA, is unlikely to be found liable for patent infringement. Significantly, in a recent decision the Federal Circuit explicitly stated that a claim purporting to cover the mere act of thinking about a physiological correlation would be *per se* patent ineligible.⁴⁴ This implies that once an individual's genome has been sequenced, an entity (such as a doctor or genetic counselor) could not be found liable for patent infringement for simply analyzing the data. In any event, as a practical matter it is extremely rare, if not unheard of, for a healthcare provider to be directly sued simply for practicing medicine.

As DNA sequencing technology advances, it seems inevitable that more and more individuals will obtain their own genome sequences. My research has led me to conclude that human gene patents will in all likelihood not prove a significant impediment to the widespread commercialization and utilization of WGS in the US. However, if I am wrong, there are mechanisms available for circumventing U.S. patents that could be implemented, such as having DNA sequencing performed outside the U.S., or perhaps even on an Indian reservation.⁴⁵ Once an individual obtains their own personal genome sequence, it does not appear that patents will pose any obstacle to them having the sequence information analyzed. For example, a recent publication reported that with genomic sequence data in hand, one would be able to analyze for BRCA mutations without infringing any patent.⁴⁶ Significantly, the Federal Circuit has unambiguously held that importation of data generated by a patented method does not constitute an infringement, so sequence data generated outside the US could be imported for use and analysis in the US.⁴⁷

The Jensen & Murray Study Provides No Quantitative Measure of the Extent to Which Genetic Testing Would Infringe Gene Patents

At this point it should be quite apparent that the Study provides absolutely no basis to infer that 20%, or for that matter any defined percent of human genes are covered by patents that would be infringed by sequencing the gene, or for that matter studying or using the gene. To further illustrate this point, note that some of the gene patents identified in the study are directed towards DNA microarrays, which simultaneously employ DNA fragments representing many hundreds or even thousands of distinct human genes.⁴⁸ For example, Claim 1 of US patent

⁴⁴ *Prometheus Laboratories v. Mayo*, 628 F.3d 1347 (Fed. Cir. 2010).

⁴⁵ *Specialty House of Creation, Incorporated v. Quapaw Tribe of Oklahoma*, 4-10-cv-00371 (OKND January 27, 2011, Order) (Frizzell, J.) (Indian tribe enjoys sovereign immunity from patent infringement suits).

⁴⁶ Steven L. Salzberg and Mihaela Pertrea, *Do-It-Yourself Genetic Testing*, *Genome Biology* 11:404 (2010).

⁴⁷ *Bayer AG v. Housey Pharmaceuticals*, 340 F.3d 1367 (Fed. Cir. 2003).

⁴⁸ This was explicitly pointed out by Jensen and Murray in their Supporting Online Materials.

number 6,500,938 claims “A combination comprising a plurality of polynucleotide probes, wherein said plurality of probes are SEQ ID NOs:1-1490.”⁴⁹ This patent claim does not recite any full-length gene sequence, only specific probes derived from portions of cDNA fragments representing distinct human genes.

Significantly, the claim would only be infringed by a product comprising a specific combination of all 1490 unique probes, e.g., as components of a microarray. If even one of these probes is missing, the product would not infringe the claim. In fact, different probes capable of identifying the exact same genes could be used as substitutes, as long as the specifically claimed probes are not used in the specific combination defined by the patent claim. It is absurd to suggest that this patent confers ownership on these 1490 genes. In particular, this patent has absolutely no implications for WGS, or for the vast majority of research or diagnostics uses of these genes.

Bear in mind that the Study’s conclusion that 20% of human genes are mentioned in US patents is based entirely on the fact that the authors found 4382 human genes that were mentioned in patent claims (which represented about 20% of known human genes at that time). Thus, this single microarray claim, by virtue of the fact that it mentions probes representing up to 1490 human genes, could account for up to one third of the human genes identified by Jensen and Murray as "mentioned" in patent claims.

Another factor undercutting an assumption that 20% of genes are patented is the fact that a large percentage of the patents identified in the Study are no longer in force. The patents identified in the study are already beginning to reach the end of the statutory term - as of April 2011 the 15 oldest patents had already expired, and many more will join them over the next several years. On top of that, as of April 2011 30% of the remaining patents in the data set have been terminated because the patent owners have failed to pay the necessary fees to maintain them in force.⁵⁰

The Incyte Patents

The Study identified Incyte Pharmaceuticals/Incyte Genomics (“Incyte”), a company founded around a platform technology for identifying and sequencing cDNA molecules, as the top gene patent assignee. According to Jensen and Murray, Incyte’s “IP rights cover 2000 human genes,” which would represent close to half of the genes identified in the Study as "explicitly

⁴⁹ For another example, see US Patent Number 6,607,879 (Claim 1: A composition comprising a plurality of cDNAs for use in detecting the altered expression of genes in an immunological response, wherein said plurality of cDNAs comprises SEQ ID NOs:1-1508 or the complete complements thereof.).

⁵⁰ Patent owners are required to pay maintenance fees periodically after the patent issues, and if they fail to do so the patent expires, although there are mechanisms for reviving expired patents in some instances where failure to pay maintenance fee was unintentional or unavoidable. 35 U.S.C. 41.

claimed as US IP."⁵¹ Thus, it would be informative to look closer at these patents which form much of the basis for the Myth.

A search of the USPTO Public Pairs system conducted on April 11, 2011 showed that only 37 of the 398 Incyte patents flagged in the Study are still in force - the others have all expired owing to Incyte's failure to pay the necessary maintenance fees.⁵² The claims in the Incyte patents are directed primarily to recombinant use of the gene for the expression of the encoded protein. Clearly the main value Incyte saw in patenting genes was based on the potential that someday the protein encoded by the gene might prove commercially significant, such as would be the case if the protein were to be approved for use as a biologic drug. When this failed to materialize, they made a rational decision to forgo paying the relatively modest maintenance fees required to keep the patents in force. This implies that Incyte concluded (correctly in my view) that most of its human gene patents would not cover other commercially significant uses of the genes, such as diagnostic testing or WGS.

My review of many of the other gene patents identified by the Study suggest that this is generally the case - most were written with protein production in mind, and very few include claims drafted in a manner likely to encompass DNA sequencing. Some of the owners of these patents have probably continued to maintain these patents because it is easier to pay the relatively nominal maintenance fees than to thoroughly assess the value of their patents. But because of the sheer number of gene patents owned by Incyte, the cost of maintaining them became significant, which might explain why the company has been much more aggressive in cutting its losses than gene patent owners as a whole.

Policy Implications

A recent Science article on gene patents reported that a lawyer for a company developing multiplex genetic diagnostic tests initially feared that his company would need to spend \$35 million in legal fees investigating whether their technology would infringe gene patents, but that later after he had "fully analyzed the patents on a handful of genes that [his company] might use, he was encouraged, finding plenty of room to operate."⁵³ His experience is very consistent with my findings - although many patents have issued with claims mentioning human genes, in actuality few if any are likely to prove substantial impediments to WGS and other forms of multiplex genetic testing.

As alluded to by Judge Moore during oral arguments before the Federal Circuit in *AMP v. PTO*, a decision by the court that results in the wholesale invalidation of gene patents would raise serious policy concerns by disrupting the expectations of investors in biotechnology, who

⁵¹ Kyle Jensen & Fiona Murray, *Intellectual Property Landscape of the Human Genome*, 310 Science 239 (2005).

⁵² <http://portal.uspto.gov/external/portal/pair> (last visited July 11, 2011).

⁵³ Sam Kean, *The Human Genome (Patent) Project*, 331 Science 530 (2011).

have relied upon the PTO's long-standing practice of allowing gene patents.⁵⁴ Gene patents have played an important role in incentivizing biotechnology investment for many years, and to date there appears to be no compelling evidence that they have substantially impeded access to genetic diagnostic testing, as noted in a recent Report commissioned by the Secretary of Health and Human Services.⁵⁵ For example, with respect to the Myriad BRCA test, the Report states that "one surprising finding from the case studies was that the per-unit price of the full-sequence BRCA test, which often is cited as being priced very high, was actually quite comparable to the price of other full-sequence test done by polymerase chain reaction (PCR), at both nonprofit and for-profit testing laboratories."

The fear that gene patents will have a substantial negative effect on genetic testing, and particularly WGS, has been based in large part on a greatly exaggerated perception of the preclusive effect of these patents in the aggregate. The widespread misinterpretation of the Study is a good example, with so many commentators assuming that it shows that 20% of genes are patented in a manner that restricts the ability to determine the sequence of the genes or test for genetic variations. It would be a mistake to change patent law and policy based largely on the misinterpretation of empirical studies such as that conducted by Jensen and Murray, which sheds an interesting light on the IP landscape of the human genome, but has little to say with respect to the impact of gene patents on DNA sequencing and genetic testing.

⁵⁴ Oral argument in *Association for Molecular Pathology v. US Patent and Trademark Office*, Federal Circuit Docket Number 10-1406, available at <http://www.cafc.uscourts.gov/oral-argument-recordings/search/audio.html>.

⁵⁵ Revised Draft Report on Gene Patents and Licensing Practices and Their Impact on Patient Access to Genetic Tests (February 5, 2010) available at <http://oba.od.nih.gov/oba/SACGHS/SACGHS%20Patents%20Report%20Approved%202-5-20010.pdf> (last visited May 6, 2011); see also Brief of *Amici Curiae* Christopher M. Holman and Robert Cook-Deegan, *Association for Molecular Pathology v. US PTO*, Federal Circuit Docket Number 2010-1406, available at https://docs.google.com/fileview?id=0B9_IJGo9WK0OWVvkYjg3NjtYzRjYi00ODIyLWIyMjAtYmJkZDQxMGZmYTZi&hl=en (last visited May 3, 2011)..