

University of Michigan Law School

## University of Michigan Law School Scholarship Repository

---

Articles

Faculty Scholarship

---

2012

### Unblocked Future: Why Gene Patents Won't Hinder Whole-Genome Sequencing and Personalized Medicine

W. Nicholson Price II

*University of Michigan Law School, wnp@umich.edu*

Available at: <https://repository.law.umich.edu/articles/2285>

Follow this and additional works at: <https://repository.law.umich.edu/articles>



Part of the [Health Law and Policy Commons](#), [Intellectual Property Law Commons](#), and the [Science and Technology Law Commons](#)

---

#### Recommended Citation

Price, W. Nicholson, II. "Unblocked Future: Why Gene Patents Won't Hinder Whole-Genome Sequencing and Personalized Medicine." *Cardozo Law Review* 33, no. 4 (2012): 1601-31.

This Article is brought to you for free and open access by the Faculty Scholarship at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Articles by an authorized administrator of University of Michigan Law School Scholarship Repository. For more information, please contact [mlaw.repository@umich.edu](mailto:mlaw.repository@umich.edu).

UNBLOCKED FUTURE:  
WHY GENE PATENTS WON'T HINDER  
WHOLE GENOME SEQUENCING AND  
PERSONALIZED MEDICINE

W. Nicholson Price II\*

ABSTRACT

*Whole-genome sequencing has been hailed as the crucial next step in personalized medicine. It has also been described as likely violating hundreds—if not thousands—of pre-existing patents on individual genes. These claims of patent infringement, however, are usually made without detailed analysis. Instead of stating that infringement definitely occurs, or in what circumstances it occurs, the discussion of whole-genome sequencing mentions that some claims may be typically infringed, but some may be invalid, and leaves the matter there. This Article seeks to provide a detailed analysis of the ways that whole-genome sequencing may infringe extant gene patents, focusing on the common basic structure of most such patents. In particular, the sequencing step itself may infringe the composition-of-matter claims of isolated DNA molecules in only a very few gene patents, with novel nanopore sequencing technology appearing to bypass infringement altogether. Gene patents often also include methods claims for comparing the personal sequence with a reference sequence for diagnostic purposes; these claims are much more likely to be infringed by any plausible whole-genome sequencing effort, but appear to fall into a two-class trap whereby comparison-only methods claims are vulnerable to section 101 patentable subject matter challenges and determination-and-comparison methods can be relatively easily avoided by having different entities perform the sequencing and analysis steps. The Article concludes with a brief analysis of policy considerations for whole-genome sequencing and suggestions for moving forward.*

---

\* Visiting Scholar, UCSF/UC Hastings Consortium on Law, Science and Health Policy, UC Hastings College of the Law. Columbia Law School, J.D. 2011. Columbia Graduate School of Arts and Sciences, Department of Biology, Ph.D. 2010. The author wishes to thank Hal Edgar, Bob Cook-Deegan, Jaime King, Henry Lebowitz, Jerry Sobel, and Ana Bračić for their very helpful comments and feedback.

## TABLE OF CONTENTS

INTRODUCTION .....	1602
I. WHAT DO GENE PATENTS LOOK LIKE?.....	1607
A. <i>DNA Claims Cover Molecules Based on DNA or Encoded Protein Sequence</i> .....	1609
1. Composition-of-Matter Claims Claim Sequences Specified as DNA, Protein-Encoding, or Short DNA Sub-Sequences ....	1609
2. In General, Composition-of-Matter DNA Claims Are for “Isolated” Nucleic Acids .....	1613
B. <i>Method Claims Usually Cover the Comparison of Sequence to References</i> .....	1616
II. DOES WHOLE-GENOME SEQUENCING INFRINGE COMPOSITION-OF-MATTER CLAIMS TO ISOLATED GENES?.....	1618
A. <i>“Traditional” Whole-Genome Sequencing Techniques: De Novo and Reference Sequence Assembly of Shotgun Sequences</i> .....	1619
B. <i>A Novel Sequencing Technique: Nanopore Sequencing</i> .....	1624
III. DRAWING MEDICAL CONCLUSIONS AND COMMUNICATING WITH THE PATIENT LIKELY INFRINGES CURRENT PATENTS, BUT THOSE CLAIMS ARE OF QUESTIONABLE VALIDITY UNDER ASSOCIATION FOR MOLECULAR PATHOLOGY.....	1626
CONCLUSION.....	1629

## INTRODUCTION

Personalized medicine has been hailed as the next great evolution in health care.<sup>1</sup> People differ, and the diseases they have differ as well. Personalized medicine aims to identify and use those variations to improve medical care.<sup>2</sup> Genetic differences can influence which diseases an

<sup>1</sup> C. Li, *Personalized Medicine—The Promised Land: Are We There Yet?*, 79 CLINICAL GENETICS 403–12 (2011).

<sup>2</sup> See generally Edward Abrahams & Mike Silver, *The Case for Personalized Medicine*, 3 J. DIABETES SCI. & TECH. 680 (2009); Suraksha Agrawal & Faisal Khan, *Human Genetic Variation and Personalized Medicine*, 51 INDIAN J. PHYSIOLOGY & PHARMACOLOGY (2007); Geoffrey S. Ginsburg & Jeanette J. McCarthy, *Personalized Medicine: Revolutionizing Drug Discovery and Patient Care*, 19 TRENDS IN BIOTECHNOLOGY 491 (2001); Geoffrey S. Ginsburg & Huntington F. Willard, *Genomic and Personalized Medicine: Foundations and Applications*, 154 TRANSLATIONAL RES. 277 (2009); Kyung-Won Hong & Bermseok Oh, *Overview of Personalized Medicine in the Disease Genomic Era*, 43 BMB REP. 643 (2010); Yuqiu Jiang & Mu Wang, *Personalized Medicine in Oncology: Tailoring the Right Drug to the Right Patient*, 4 BIOMARKERS MED. 523 (2010); Laviero Mancinelli et al., *Pharmacogenomics: The Promise of Personalized Medicine*, 2 AAPS J. 29 (2002); Robert Roberts, *Personalized Medicine: A Reality Within This Decade*, 1 J.

individual experiences, differences in the course of those diseases, and how the individual responds to treatment. The paradigmatic field of personalized medicine is pharmacogenomics, which focuses on determining how people react differently to drugs based on their genetic differences.<sup>3</sup> As the directors of the FDA and NIH jointly put it, personalized medicine is about “the best ways to develop new therapies and optimize prescribing by steering patients to the right drug at the right dose at the right time.”<sup>4</sup> Many factors play into the growth of personalized medicine: identifying the right genetic markers to study, clinically demonstrating relationships between genetic variants and drug responses; and limiting side effects of these targeted therapies.<sup>5</sup>

For truly personalized medicine, widespread availability of whole-genome sequencing is a crucial step.<sup>6</sup> Whole-genome sequencing (WGS) differs from genetic testing as commonly performed today in the same way that the detailed satellite map of a parcel differs from a surveyor’s listing of identifiable landmarks. Current genetic testing is usually for a specific gene—for instance, the BRCA1 gene which relates to a predisposition for breast cancer—and determines which variant of the gene the patient has at that particular locus. Even the companies that today offer generalized genetic services—companies like Navigenics<sup>7</sup> or 23andMe<sup>8</sup>—actually only examine a relatively small number of specific loci on the genome of the patient, correlating the results at those loci with already-known genetic variations.<sup>9</sup> WGS, on the other hand, is the technique of actually determining the entire sequence—base by base—of the full genome of a patient, resulting in a string of around six billion

---

CARDIOVASCULAR TRANSLATIONAL RES. 11 (2008); Gregory E. Simon & Roy H. Perlis, *Personalized Medicine for Depression: Can We Match Patients with Treatments?*, 167 AM. J. PSYCHIATRY 1445 (2010); Michael J. Wagner, *Pharmacogenetics and Personal Genomes*, 6 PERSONALIZED MED. 643 (2009); Gary E. Marchant, *Personalized Medicine and the Law*, ARIZ. ATT’Y, Oct. 2007, at 12.

<sup>3</sup> See, e.g., Mancinelli et al., *supra* note 2; Liewei Wang et al., *Genomics and Drug Response*, 364 NEW ENG. J. MED. 1144 (2011); Teresa Kelton, *Pharmacogenomics: The Rediscovery of the Concept of Tailored Drug Therapy and Personalized Medicine*, HEALTH LAW., Jan. 2007, at 1.

<sup>4</sup> Margaret A. Hamburg & Francis S. Collins, *The Path to Personalized Medicine*, 363 NEW ENG. J. MED. 301, 301 (2010).

<sup>5</sup> *Id.*

<sup>6</sup> See, e.g., Boris Pasche & Devin Absher, *Whole-Genome Sequencing: A Step Closer to Personalized Medicine*, 305 JAMA 1596, 1596–97 (2011). *But see* Wylie Burke & Bruce M. Psaty, *Personalized Medicine in the Era of Genomics*, 298 JAMA 1682, 1682–84 (2007) (arguing that genomics will make only moderate contributions to personalized medicine).

<sup>7</sup> NAVIGENICS, <http://www.navigenics.com> (last visited Feb. 15, 2012).

<sup>8</sup> 23ANDME, <https://www.23andme.com> (last visited Feb. 15, 2012).

<sup>9</sup> Navigenics and 23andMe both offer genotyping services, where genetic markers are identified, rather than full sequencing services. See FAQs, NAVIGENICS, [http://www.navigenics.com/visitor/what\\_we\\_offer/faqs](http://www.navigenics.com/visitor/what_we_offer/faqs) (last visited Feb. 15, 2012); *Using Your Personal DNA Test*, 23ANDME, <https://www.23andme.com/howitworks> (last visited Feb. 15, 2012).

As, Cs, Gs, and Ts, representing the adenine, cytosine, guanine, and thymine bases actually present in the individual's chromosomes.<sup>10</sup>

The additional information provided by WGS facilitates personalized medicine in two ways. First, widespread WGS increases the availability of the information needed to determine how drugs work with different genetic variations.<sup>11</sup> Genomic association studies correlate drug reactions to genetic variations. For instance, individuals with a certain gene might have a higher predisposition to breast cancer,<sup>12</sup> while those with another gene might be hypersensitive to a particular drug.<sup>13</sup> WGS allows those correlations to be made at a more nuanced level, and as WGS becomes more prevalent, such studies will have larger potential pools of data on which to draw. Second, widespread availability of WGS allows individuals being treated to know and use their own detailed genetic information.<sup>14</sup> This, in combination with the correlations just discussed, can be used to select or tailor treatment according to the patient's own genetic characteristics.<sup>15</sup>

The enterprise of WGS, however, faces substantial barriers, not only technical and financial, but also based on intellectual property constraints. Many thousands of patents have been issued covering individual human genes.<sup>16</sup> Each of these could potentially be infringed by a WGS technique, which, by definition, involves looking not only at specific places in the human genome, but also at every base and every gene—and, consequently, at every gene which is covered by any human gene patent. (Hereinafter, unless otherwise specified, all gene patents are assumed to be for human genes, though some descriptions, such as the format of common claims, may apply to non-human gene patents as well.) Descriptions of WGS and personalized medicine often mention this problem, but usually as a relatively unspecific problem of “likely infringement” which must be addressed by broad policy changes.<sup>17</sup> This

---

<sup>10</sup> Pauline C. Ng & Ewen F. Kirkness, *Whole Genome Sequencing*, 628 *METHODS MOLECULAR BIOLOGY* 215, 216 (2010).

<sup>11</sup> See, e.g., Wang et al., *supra* note 3, at 1144–51.

<sup>12</sup> See, e.g., Francine Durocher et al., *Mutation Analysis of the BRCA1 Gene in 23 Families with Cases of Cancer of the Breast, Ovary, and Multiple Other Sites*, 33 *J. MED. GENETICS* 814 (1996).

<sup>13</sup> See, e.g., Mark McCormack et al., *HLA-A\*3101 and Carbamazepine-Induced Hypersensitivity Reactions in Europeans*, 364 *NEW ENG. J. MED.* 1134 (2011).

<sup>14</sup> See, e.g., Pasche & Absher, *supra* note 6.

<sup>15</sup> Abrahams & Silver, *supra* note 2; Mancinelli et al., *supra* note 2; Wagner, *supra* note 2.

<sup>16</sup> The most recent comprehensive study of human gene patents found that in 2005, patents covered over 20% of human genes, or 4382 genes out of around 23,000. (Because of continuing research, the number of sequences classified as “genes” fluctuates somewhat.) Kyle Jensen & Fiona Murray, *Intellectual Property Landscape of the Human Genome*, 310 *SCIENCE* 239, 239 (2005).

<sup>17</sup> See, e.g., SEC’Y’S ADVISORY COMM. ON GENETICS, HEALTH & SOC’Y, DEP’T OF HEALTH & HUMAN SERVS., *GENE PATENTS AND LICENSING PRACTICES AND THEIR IMPACT ON PATIENT ACCESS TO GENETIC TESTS* 17 (2010) (hereinafter SACGHS REPORT), available at

understandable assumption appears to be based on three factors: First, analysis of WGS infringement is complex; thousands of gene patents exist, and any truly exhaustive analysis is daunting in scope. Second, that WGS infringes gene patents makes intuitive sense on either a superficial or slightly deeper level: on a superficial level, gene patents cover genes, and WGS requires looking at and using genes, so infringement seems likely; on a slightly deeper level, gene patents claim *isolated* DNA molecules, and WGS does involve determining the sequence of genes through creating and using isolated DNA molecules. Third and finally, older generations of DNA sequencing may in fact have infringed current gene patents. Whether or not WGS infringes individual gene patents, however, is far from clear: in *Association for Molecular Pathology v. U.S. Patent and Trademark Office (AMP I)* (district) or *AMP II* (appellate) the high-profile litigation over Myriad Genetics' BRCA1/2 breast cancer predisposition gene patents, attorneys disagreed vigorously about whether WGS would infringe Myriad's patents.<sup>18</sup> Though the majority expressed no opinion on the subject, which was not a litigated issue, Judge Bryson claimed in his dissent that "some of Myriad's challenged composition claims effectively preempt any attempt to sequence the BRCA genes, including whole-genome sequencing."<sup>19</sup>

---

[http://oba.od.nih.gov/oba/sacghs/reports/SACGHS\\_patents\\_report\\_2010.pdf](http://oba.od.nih.gov/oba/sacghs/reports/SACGHS_patents_report_2010.pdf) (finding that required licensing of many gene patents will impede whole-genome sequencing); Subhashini Chandrasekharan & Robert Cook-Deegan, *Gene Patents and Personalized Medicine—What Lies Ahead?*, 1 GENOME MED. 92, 94 (2009) ("Our preliminary analysis suggests that microarray-based methods of genetic analysis and gene variants, and a fortiori full-genome sequencing could arguably infringe on some patent claims."); Courtney C. Scala, *Making the Jump from Gene Pools to Patent Pools: How Patent Pools Can Facilitate the Development of Pharmacogenomics*, 41 CONN. L. REV. 1631, 1661 (2008) (stating that pharmacogenomics requires addressing the gene "patent thicket"); Robert Cook-Deegan, Lecture at Columbia Univ. Med. Ctr.: The Topsy-Turvy World of Gene Patents (Nov. 18, 2010) (slides on file with author); David P. Lentini & Stephen A. Bent, *Intellectual Property: Patents and Genomic Medicine*, BIO-ITWORLD (Nov. 19, 2004), [http://www.bio-itworld.com/archive/111904/si\\_gm.html](http://www.bio-itworld.com/archive/111904/si_gm.html).

<sup>18</sup> *Ass'n for Molecular Pathology v. USPTO (AMP II)*, 653 F.3d 1329 (Fed. Cir. 2011). At oral argument before the Federal Circuit, Judge Bryson asked Gregory Castanias, Myriad's counsel, about this issue: "To me, at least, it is an important question as to how preclusive your patent—and any other patent on any particular gene—would be if, in effect, you have to get 100, 200 or 1,000 licenses before you can sequence the genome of an individual." Oral Argument at 11:53–12:13, *AMP II*, 653 F.3d 1329 (Fed. Cir. 2011) (No. 2010-1406), available at <http://oralarguments.cafc.uscourts.gov/Audiomp3/2010-1406.mp3>. Mr. Castanias, while initially stating, "I'm not sure that my client has formed a view on that," went on to argue that it would depend on the technology involved, but that sequencing either very short (thirty-two nucleotide fragments) or very long (whole chromosomes) segments would not violate the patents, as he read them. *Id.* at 12:13–18, 1:08:33–47 (based on the definition of "isolation" in the patents). Christopher Hansen, on the other hand, arguing for the ACLU on behalf of the plaintiffs, argued that WGS would violate Myriad's patents, particularly its sequence-comparison methods patents. *Id.* at 34:00–08.

<sup>19</sup> *AMP II*, 653 F.3d at 119–20 (Bryson, J., concurring in part and dissenting in part). Judge Bryson went on to cite the SACGHS report, *supra* note 17, at 49–62, referencing the "thicket of [gene] patents" that would require WGS firms to license "thousands of patents from many

This Article seeks to clarify the infringement landscape with respect to WGS, and concludes that most, if not all isolated DNA gene patents are likely not infringed by WGS. In brief response<sup>20</sup> to the three factors listed immediately above: first, gene patents can be grouped into patterns, so that a reliable analysis does not require exhaustive canvassing; second, a close look at the specific mechanics of WGS shows that first- and second-level superficial analyses deceptively suggest infringement; and third, WGS technology has moved beyond older sequencing techniques in ways that matter for infringement analysis.

For these reasons, isolated DNA gene patents are likely not infringed by WGS. Infringement is especially unlikely for a nascent technology, nanopore sequencing, which appears more likely to actually fulfill the requirements of true personalized medicine. Other gene patent claims cover not the isolated genes themselves, but instead the process of comparing genetic sequences for use in a diagnosis. These claims are more likely infringed by personalized medicine and WGS, since that comparison is at the heart of personalized medicine, but are either likely invalid or relatively easily worked around by separating the steps of the method to be completed by different actors.

As an important caveat, this Article will not argue the question of whether patents claiming isolated genes should themselves be valid. This issue has been the subject of much scholarly debate,<sup>21</sup> but has long been considered settled by the U.S. Patent and Trademark Office (USPTO)<sup>22</sup> and isolated gene patents have recently been upheld as patentable subject matter by the U.S. Court of Appeals for the Federal Cir-

---

different licensors.” *AMP II*, 653 F.3d at 138 (Bryson, J., concurring in part and dissenting in part).

<sup>20</sup> Although these three factors may most reasonably explain *why* there is a casual assumption that WGS infringes many gene patents, they are not the most productive framework for actually analyzing current infringement. Therefore, the infringement analysis will not be grouped according to these three factors; however, where relevant, their applicability will be mentioned.

<sup>21</sup> See generally Lori B. Andrews & Jordan Paradise, *Gene Patents: The Need for Bioethics Scrutiny and Legal Change*, 5 YALE J. HEALTH POL’Y L. & ETHICS 403 (2005); Daniel J. Kevles & Ari Berkowitz, *The Gene Patenting Controversy: A Convergence of Law, Economic Interests, and Ethics*, 67 BROOK. L. REV. 233 (2001); Patricia A. Lacy, *Gene Patenting: Universal Heritage vs. Reward for Human Effort*, 77 OR. L. REV. 783 (1998); Barbara Looney, *Should Genes Be Patented? The Gene Patenting Controversy: Legal, Ethical, and Policy Foundations of an International Agreement*, 26 L. & POL’Y INT’L BUS. 231 (1994); John F. Merz, *Disease Gene Patents: Overcoming Unethical Constraints on Clinical Laboratory Medicine*, 45 CLINICAL CHEMISTRY 324 (1999); John F. Merz et al., *Diagnostic Testing Fails the Test*, 415 NATURE 577 (2002); David B. Resnik, *The Morality of Human Gene Patents*, 7 KENNEDY INST. ETHICS J. 43 (1997).

<sup>22</sup> See, e.g., Utility Examination Guidelines, 66 Fed. Reg. 1092 (Jan. 5, 2001) (describing guidelines for evaluating “gene patents”). For an evaluation of the guidelines, see Anna E. Morrison, *The U.S. PTO’s New Utility Guidelines: Will They Be Enough to Secure Gene Patent Rights?*, 1 J. MARSHALL REV. INTELL. PROP. L. 142 (2001); Theresa M. Summers, *The Scope of Utility in the Twenty-First Century: New Guidance for Gene-Related Patents*, 91 GEO. L.J. 475 (2002).

cuit (Federal Circuit) in *Association for Molecular Pathology v. USPTO*.<sup>23</sup> Therefore, despite the possibility of an eventual change in the law, the validity of standard composition-of-matter claims for isolated genes will not be addressed. The validity of short-sequence composition claims and some diagnostic methods claims, however, will be considered briefly when relevant.

Part I describes in some detail a generalized picture of gene patents and the types they typically contain. Part II analyzes the likelihood of WGS infringing the composition-of-matter claims in gene patents—that is, the claims for the genes themselves. Part III analyzes potential infringement of the methods claims, which generally claim the process of determining gene mutations and using that information to diagnose a patient. Part IV concludes by briefly examining policy considerations and ways forward for personalized medicine based on whole-genome sequences.

## I. WHAT DO GENE PATENTS LOOK LIKE?

Much of the uncertainty surrounding gene patents comes from their variation. There is not a single set model for a gene patent, with standard language that can be parsed and analyzed to determine what any individual gene patent covers. Instead, each patent claims different variations on two types of inventions. First, gene patents typically claim isolated DNA molecules, including the entire sequence or a partial fragment of that sequence. Second, gene patents frequently claim a method of comparing the sequence of an individual with the known reference sequence (and possibly known mutations) and (sometimes) using that identification to draw medical conclusions.

A bit of substantive background helps explain the two types of invention claimed. Gene patents related to medicine fall into two main substantive classes.<sup>24</sup> These classes can be roughly described as biotechnological and diagnostic, though the distinction is not always clear-cut

---

<sup>23</sup> *AMP II*, 653 F.3d at 1329. The Federal Circuit had previously treated gene patents' validity as a given. See, e.g., *In re Deuel*, 51 F.3d 1552, 1559 (Fed. Cir. 1995) (holding that the existence of a known method of isolating DNA molecules is "essentially irrelevant" to the question of the specific-claimed DNA's obviousness without addressing section 101 patentability); *In re Bell*, 991 F.2d 781, 783 (Fed. Cir. 1993) (addressing obviousness of amino acid sequence, not section 101 patentability).

<sup>24</sup> See, e.g., Christopher M. Holman, *Trends in Human Gene Patent Litigation*, 322 SCIENCE 198 (2008) (classifying gene patents into four categories: therapeutic proteins, research tools, diagnostic testing, and forensic testing). Only therapeutic protein (such as erythropoietin or insulin) and diagnostic testing (such as BRCA1/2 for breast cancer predisposition or AspA for Canavan's disease predisposition) are highly relevant to medicine.

at a practical level.<sup>25</sup> The basic divide, however, describes both the commercial significance of gene patents and has some implication for which claims are likely to be included in the patent, argued over academically, and enforced through litigation.<sup>26</sup> Biotechnological gene patents, whose value is in the protein they encode, are by far the most currently valuable gene patents. Amgen's Patent No. 4,703,008 for erythropoietin provides one such example; the patent allows it to prevent others from making recombinant erythropoietin, the basis of its Epogen anemia drug with over \$2.5 billion in yearly sales.<sup>27</sup> Diagnostic gene patents, on the other hand, cover genes whose commercial value is not primarily in producing large amounts of their protein product, but in evaluating how that product is already naturally produced in humans. As an example, Myriad Genetics' Patent No. 5,747,282 (BRCA1 patent) covers the BRCA1 gene, which is related to a predisposition to breast cancer.<sup>28</sup> Instead of producing the BRCA1 protein, as in a biotechnology gene patent, Myriad provides its own BRCAAnalysis service, which involves determining the sequence of the BRCA1 (and BRCA2) genes in a patient and then interpreting that information to find whether the patient has a predisposition to develop breast cancer.<sup>29</sup> Myriad's patents are used to prevent others from offering this service,<sup>30</sup> which had sales of around \$353 million in fiscal year 2011.<sup>31</sup>

These two different functional classes of gene patents are not in themselves different for infringement analyses, but have to a certain extent resulted in different claim formats. The biotechnological gene patents' most important claims are those for the full gene, whether defined by gene or protein sequence, since the full gene is needed to produce the protein of interest. Diagnostic gene patents include those full-gene central claims, but also require claims for shorter sequences, hybridizing sequences, and methods of actually making diagnoses based on the information revealed by the gene sequence. These differences in patent construction will be discussed in more detail below as each type

---

<sup>25</sup> For instance, a hypothetical gene patent covering insulin could have value both for the biotechnological production of therapeutic insulin and for a diagnostic test that could identify a predisposition towards developing diabetes.

<sup>26</sup> Holman, *supra* note 24, at 198–99.

<sup>27</sup> AMGEN, AMGEN 2010 ANNUAL REPORT AND FINANCIAL SUMMARY 171 (2011), available at <http://investors.amgen.com/phoenix.zhtml?c=61656&p=irol-reportsannual> (under the "Annual Reports" heading, click on "2010 Annual Report and 10-K" to download the report).

<sup>28</sup> U.S. Patent No. 5,747,282 (filed June 7, 1995).

<sup>29</sup> Ass'n for Molecular Pathology v. USPTO (*AMP I*), 702 F. Supp. 2d 181, 203 (S.D.N.Y. 2010), *aff'd in part, rev'd in part*, 653 F.3d 1329 (Fed. Cir. 2011).

<sup>30</sup> *Id.* at 204–06.

<sup>31</sup> Turna Ray, *Posting 11 Percent Revenue Growth in FY2011, Myriad Focuses on Growing MDx Pipeline*, GENOMEWEB PHARMACOGENOMICS REP. (Aug. 10, 2011), <http://www.genomeweb.com/dxpgx/posting-11-percent-revenue-growth-fy2011-myriad-focuses-growing-mdx-pipeline>.

of claim (composition-of-matter for the gene itself or method-of-sequence-comparison) is addressed.

A. *DNA Claims Cover Molecules Based on DNA or Encoded Protein Sequence*

Most gene patents rely on composition-of-matter claims to isolated nucleic acids. This type of claim also lies at the heart of many ethical and policy debates over gene patents: when people argue that gene patents cover genes within the human body, or that gene patents restrict the common heritage of humankind, these are the claims to which they refer.<sup>32</sup> These claims come in several forms;<sup>33</sup> however, one of the broadest types of claim, for short DNA sub-sequences, are very likely invalid in most cases.<sup>34</sup> Common to these composition claims, based on USPTO guidelines,<sup>35</sup> is language claiming an “isolated” nucleic acid, a term with important implications discussed below.<sup>36</sup>

1. Composition-of-Matter Claims Claim Sequences Specified as DNA, Protein-Encoding, or Short DNA Sub-Sequences

Central to almost all gene patents are claims for isolated deoxyribonucleic acids (DNA) with a specific sequence.<sup>37</sup> To avoid falling outside the statutory subject matter of patents as defined in section 101 of the Patent Act, these claims claim isolated DNA, relying on the isolated products doctrine initially expounded in *Parke-Davis & Co. v. H.K. Mulford Co.*<sup>38</sup> The USPTO relies on this doctrine; it does not define a

---

<sup>32</sup> See, e.g., NUFFIELD COUNCIL ON BIOETHICS, *THE ETHICS OF PATENTING DNA* 21–23 (2002).

<sup>33</sup> See *infra* Part I.A.1.

<sup>34</sup> See *infra* notes 49–58 and accompanying text.

<sup>35</sup> Utility Examination Guidelines, *supra* note 22, at 1093.

<sup>36</sup> See *infra* Part I.A.2.

<sup>37</sup> Many patents claim nucleic acids in general, which includes ribonucleic acids (RNA). Since no genomic sequencing technology uses RNA (RNA is far more fragile than DNA, among other limitations), the distinction will be ignored here, and DNA will be assumed to be the relevant nucleic acid.

<sup>38</sup> 189 F. 95 (C.C.S.D.N.Y. 1911). The Federal Circuit recently held in *AMP II* that “an isolated DNA molecule is not a purified form of a natural material, but a distinct chemical entity.” *AMP II*, 653 F.3d 1329, 1352 (Fed. Cir. 2011). This implies that the initial rationale behind the “isolated” language may no longer apply; however, since the Federal Circuit held that the isolated DNA molecules in question were “markedly different—have a distinctive chemical identity and nature—from molecules that exist in nature,” based on the process of their isolation, *id.* at 1351, the practical impact of isolation making DNA patentable does not change. Furthermore, the court explicitly cited with approval the USPTO’s guidelines as longstanding practice with which Federal Circuit law should comport. *Id.* at 1354–55.

“gene patent” in its examination guidelines, but states that gene discovery may form the basis for a patent if the gene is “isolated from its natural state and . . . separate[d] . . . from other molecules naturally associated with it.”<sup>39</sup> These claims for isolated DNAs come in three primary forms: the entire gene defined by DNA sequence, the entire gene defined by encoded amino acid sequence, and a fragment of the gene defined by DNA sequence. These three forms of claims each have a specific function within the realm of gene patents.

The most obvious claim in a gene patent is for an isolated DNA identified by its DNA sequence.<sup>40</sup> This can be claimed as a 100% sequence match; claim 2 of the BRCA1 patent above claims “[t]he isolated DNA of claim 1, wherein said DNA has the nucleotide sequence set forth in SEQ ID NO:1.”<sup>41</sup> Frequently, some flexibility in the match is provided; in the erythropoietin (EPO) patent, claim 1(b) claims a “purified and isolated DNA sequence encoding erythropoietin, said DNA sequence selected from the group consisting of [specified sequences and] DNA sequences which hybridize under stringent conditions to [those sequences].”<sup>42</sup> Hybridization under stringent conditions is possible without 100% sequence identity (and must be, for that part of the claim to be independently valuable).

The next type of claim in many gene patents is for an isolated DNA encoding a protein, with the DNA identified by the sequence of the encoded polypeptide. Myriad’s BRCA1 patent has as claim 1 “[a]n isolated DNA coding for a BRCA1 polypeptide, said polypeptide having the amino acid sequence set forth in SEQ ID NO:2.”<sup>43</sup> Miami Children’s Hospital Research Institute’s Patent No. 5,679,635 for ASPA (ASPA patent), a gene related to Canavan’s disease, similarly claims as claim 1 “[a]n isolated nucleic acid molecule comprising[] (a) a nucleic acid sequence encoding a human aspartoacylase polypeptide.”<sup>44</sup> These claims could potentially be invalid inasmuch as they claim DNA molecules not disclosed pursuant to the Patent Act’s section 112 written disclosure

---

<sup>39</sup> Utility Examination Guidelines, *supra* note 22, at 1093.

<sup>40</sup> Note that patents typically claim both the DNA strand and its (typically non-coding) counterpart; that is, both sides of the double helix are claimed. In the ASPA patent, for instance, Claim 1(b) claims “a nucleic acid sequence fully complementary to nucleic acid sequence (a).” U.S. Patent No. 5,679,635 (filed Sept. 9, 1994). This typically does not change the analyses described in this Article, except inasmuch as were complementary DNA sequences *not* claimed, such absences would raise the remote possibility of a technical workaround unless such a workaround was held to be infringing. In any case, the issue generally does not arise since complementary strands are usually included.

<sup>41</sup> U.S. Patent No. 5,747,282 (filed June 7, 1995).

<sup>42</sup> U.S. Patent No. 4,703,008 (filed Nov. 30, 1984).

<sup>43</sup> U.S. Patent No. 5,747,282 (filed June 7, 1995).

<sup>44</sup> U.S. Patent No. 5,679,635 (filed Sept. 9, 1994).

requirement, but the additional sequences they claim are not relevant to WGS.<sup>45</sup>

The third and certainly broadest type of claim in gene patents claims smaller sequences that are fragments of the gene or its non-coding counterpart. The BRCA1 patent's claim 5 claims "[a]n isolated DNA having at least fifteen nucleotides of the DNA of claim 1 [claiming any nucleic acid *encoding the specified polypeptide*]"<sup>46</sup>; similarly, the ASPA patent claims in claim 1(c) "a nucleic acid sequence at least sixteen nucleotides in length capable of hybridizing specifically with one of said nucleic acid molecules (a) or (b)."<sup>47</sup> In his dissent in *AMP II*, Judge Bryson described the BRCA1 claim 5 as "breathtakingly broad," and argued that it should be held invalid under section 101.<sup>48</sup>

Although a fact-agnostic legal analysis of short-sequence claims' invalidity is hard to construct—indeed, thousands of such claims have been permitted by patent examiners—recent theoretical and empirical analyses suggest that all or nearly all such patents are invalid as anticipated under section 102 of the Patent Act.<sup>49</sup> A recent analysis of the BRCA1 claim 5 addressed both the expected coverage of such claims and their actual extent.<sup>50</sup> The authors calculated the expected number of the claimed 15-base sub-sequences of BRCA1 (BRCA1 15-mers) in the human genome, assuming a random genomic ordering of bases, and found that an average of one in 600 possible 15-mers would be claimed; therefore, an average-length 10,000 base human gene would be expected to contain about fifteen claimed 15-mers.<sup>51</sup> In an actual search for claimed sequences, they found over 340,000 claimed 15-nucleotide se-

---

<sup>45</sup> This type of claim relies on the redundancy of the genetic code (since there are sixty-four three-base codons and only twenty amino acids, most amino acids are encoded by multiple codons) to claim all sequences which could encode a particular protein. The number of sequences claimed is extremely large, given that the "usage-weighted geometric mean codon degeneracy per amino acid is 3.107." Thomas B. Kepler et al., *Metastasizing Patent Claims on BRCA1*, 95 *GENOMICS* 312, 312 (2010). For instance, a twenty-amino acid protein (far shorter than any human gene) could potentially be encoded by 3.107<sup>20</sup> DNA sequences—more than seven billion.

<sup>46</sup> U.S. Patent No. 5,747,282 (filed June 7, 1995).

<sup>47</sup> U.S. Patent No. 5,679,635 (filed Sept. 9, 1994). Presumably, a fifteen-nucleotide threshold was picked as a threshold of something approaching uniqueness in at least the human genome, since the odds of choosing any given fifteen-nucleotide sequence at random are  $(\frac{1}{4})^{15}$ , or a little less than one in a billion. However, the human genome is not randomly arranged, and in fact such sequence matches occur far more frequently. Kepler, *supra* note 45, at 312–13.

<sup>48</sup> *AMP II*, 655 F.3d 1329, 1379 (Fed. Cir. 2011) (Bryson, J., concurring in part and dissenting in part).

<sup>49</sup> Kepler et al., *supra* note 45, at 313.

<sup>50</sup> *Id.* at 312–14.

<sup>51</sup> *Id.* at 312–13. Judge Bryson cited a similar analysis in his *AMP II* dissent, stating that the BRCA1 '282 patent's Claim 6, covering fifteen-mers from the cDNA sequence (just the coding sequence, not the full-length gene), "covers portions of the cDNA of more than 4% of human genes. It also covers portions of the DNA of nearly all human genes." *AMP II*, 653 F.3d at 1379 (Bryson, J., concurring in part and dissenting in part).

quences in human chromosome 1 alone.<sup>52</sup> Breadth alone is not a reason for invalidity.<sup>53</sup> However, this extraordinary breadth makes anticipation under section 102's novelty requirement very likely for most, if not all of these claims. Before gene patenting became widespread (and therefore before the vast majority of gene patents were filed),<sup>54</sup> many thousands of sequences were deposited in publicly available databases, such as Genbank,<sup>55</sup> which is likely considered a "printed publication" for section 102 novelty purposes under courts' liberal readings of that provision.<sup>56</sup> Since one in any 600 random 15-mers, on average, is claimed as a BRCA1 15-mer,<sup>57</sup> it is very likely that the claim is anticipated by previously disclosed and publically available sequences. Indeed, fifteen bases of the *natural* BRCA1 DNA sequence (ignoring the degeneracy of the genetic code) appeared in a GenBank submission more than a year before the BRCA1 patent was filed, thus anticipating claim 5 under section 102(b).<sup>58</sup> A comprehensive similar analysis for all gene patents is far beyond the scope of this paper, but would likely reveal anticipating sequences for most, if not all, short-sequence claims in gene patents. Overall sub-sequence claims are almost certainly anticipated by sequences publicly deposited before patent filing or invention.

---

<sup>52</sup> Kepler et al., *supra* note 45, at 313.

<sup>53</sup> Potentially, such extraordinary and logically unbounded breadth might render a claim vulnerable to challenge under the section 112 written-description requirement. *See, e.g., AMP II*, 653 F.3d at 1379 (Bryson, J., concurring in part and dissenting in part) ("Of course, in light of its breadth, claim 5 of the '282 patent is likely to be invalid on other grounds, and thus a ruling as to patent-eligibility with respect to that claim may be superfluous."). This possibility will not be explored in depth in this Article since section 102 anticipation arguments, *infra*, make invalidity almost certain in any case.

<sup>54</sup> *See, e.g., DPD: About the DNA Patent Database*, DNA PATENT DATABASE, <http://dnapatents.georgetown.edu/aboutdpd.htm> (last updated Feb. 1, 2012).

<sup>55</sup> *GenBank Overview*, NAT'L CTR. FOR BIOTECHNOLOGY INFO., <http://www.ncbi.nlm.nih.gov/genbank> (last updated May 3, 2011).

<sup>56</sup> For prior art to be a printed publication, it "must have been sufficiently accessible to the public interested in the art; dissemination and public accessibility are the keys to the legal determination whether a prior art reference was 'published.'" *Constant v. Advanced Micro-Devices, Inc.*, 848 F.2d 1560, 1568 (Fed. Cir. 1988). Genbank and other online databases are the dominant, and frequently required, means of making sequence information available to the scientific community. *GenBank Overview*, *supra* note 55. The Federal Circuit has previously held that inclusion in a publicly accessible database (in the particular case, Westlaw) was sufficient to constitute a printed publication. *In re Lister*, 583 F.3d 1307, 1316 (Fed. Cir. 2009).

<sup>57</sup> Kepler et al., *supra* note 45, at 312.

<sup>58</sup> The fifteen-base sequence AAGGCAAAAACAGAA is found at nucleotide 2942 of the sequence for the human T-cell receptor beta variable region, deposited November 4, 1993, well over a year before the BRCA1 patent was filed on June 6, 1995. *Homo Sapiens T-cell Receptor Beta Variable Region (TCRBV) Gene Locus, Genomic Sequence*, NAT'L CTR. FOR BIOTECHNOLOGY INFO., <http://www.ncbi.nlm.nih.gov/nucore/467918> (last visited Feb. 15, 2012).

## 2. In General, Composition-of-Matter DNA Claims Are for “Isolated” Nucleic Acids

A crucial aspect of all three forms of composition-of-matter claims is that they almost always claim an “isolated” nucleic acid.<sup>59</sup> There is no standard definition of “isolated” used in all patents, nor would such a definition be determinative, since each patentee may be his own lexicographer.<sup>60</sup> Unfortunately, the meaning of “isolated” could end up being dispositive in a significant number of cases, since it critically qualifies what is claimed. It is impossible to completely address the meaning of “isolated,” since meanings differ and would only be fully construed at the *Markman* phase<sup>61</sup> of a gene patent trial—and those trials have been quite rare, particularly with respect to diagnostic testing.<sup>62</sup> However, a reliable working sense of what “isolated” could plausibly mean can be derived from an examination of representative patents and relevant precedent, most importantly the Federal Circuit’s decision in *AMP II*.

Examples from sample DNA patents show some measure of variation with respect to the meaning of the term “isolated,” but cover essentially the same features. The BRCA1 patent includes the following in its definition section:

An “isolated” or “substantially pure” nucleic acid (e.g., an RNA, DNA or a mixed polymer) is one which is *substantially separated from other cellular components which naturally accompany a native human sequence or protein, e.g., ribosomes, polymerases, many other human genome sequences and proteins*. The term embraces a nucleic acid sequence or protein which has been removed from its naturally occurring environment, and includes recombinant or cloned DNA isolates and chemically synthesized analogs or analogs biologically synthesized by heterologous systems.<sup>63</sup>

The district court in *AMP I* accepted the plain meaning of this definition, construing the term “isolated DNA” to mean “a segment of DNA nucleotides existing separate from other cellular components normally associated with native DNA, including proteins and other DNA sequences comprising the remainder of the genome, and includes both DNA originating from a cell as well as DNA synthesized through chemical or heterologous biological means.”<sup>64</sup>

---

<sup>59</sup> Alex Osterlind, *Staking a Claim on the Building Blocks of Life: Human Genetic Material Within the United States Patent System*, 75 MO. L. REV. 617, 624–25 (2010).

<sup>60</sup> Vitronics Corp. v. Conceptoronic, Inc., 90 F.3d 1576, 1582 (Fed. Cir. 1996).

<sup>61</sup> *Markman v. Westview Instruments, Inc.*, 517 U.S. 370 (1996).

<sup>62</sup> Holman, *supra* note 24, at 198–99.

<sup>63</sup> U.S. Patent No. 5,747,282 (filed June 7, 1995) (emphasis added).

<sup>64</sup> *AMP I*, 702 F. Supp. 2d 181, 217 (S.D.N.Y. 2010).

No definition of “isolated” is provided in the ASPA patent, but reference is made to Patent No. 5,227,292 for common definitions of terms of art, including “isolated.” That patent states that “isolated nucleic acid . . . is defined as nucleic acid isolated from its natural environment (e.g. cDNA or a fragment of genomic DNA) . . . .”<sup>65</sup>

The Federal Circuit provided only a partial description of “isolation” in *AMP II*.<sup>66</sup> More broadly, the court could not define “isolated DNA” once and for all, since each patent may define the term in its own terms. The court did describe isolated DNA in terms supporting the conception of isolation as generally a physical and chemical process rather than, for instance, an informational one.<sup>67</sup> After describing native DNA, which “exists in the body as one of forty-six large, contiguous DNA molecules[, each] an integral part of a larger structural complex, a chromosome,”<sup>68</sup> the court characterized isolated DNA as “a free-standing portion of a native DNA molecule, frequently a single gene[, that] has been cleaved (i.e., had covalent bonds in its backbone chemically severed) or synthesized to consist of just a fraction of a naturally occurring DNA molecule.”<sup>69</sup> This language confirms that “isolated DNA” is chemically isolated from its genetic neighbors; it also “must be removed from its native cellular and chromosomal environment.”<sup>70</sup> This

---

<sup>65</sup> U.S. Patent No. 5,227,292 (filed July 12, 1990).

<sup>66</sup> *AMP II*, 653 F.3d 1329, 1351–53 (Fed. Cir. 2011).

<sup>67</sup> An argument could be made that “isolation” should refer to informational isolation rather than physical separation. E-mail from Robert Cook-Deegan, Dir., Duke Inst. for Genome Scis. & Pol’y Ctr. for Genome Ethics, Law & Pol’y, to author (July 17, 2011) (on file with author). Under this conception, isolation is not an actual physical separation from other cellular components by purification and from other DNA sequences by breaking covalent bonds. *Id.* Instead, a DNA sequence is “isolated” by the very process of detecting, sequencing, or selectively hybridizing with a probe. *Id.* If we imagine the soup of DNA molecules in a biochemical preparation as a crowd of individuals, physical isolation would be pulling a person out of the crowd and having that person stand off to the side; this informational concept of isolation would instead be carried out by pointing out an individual (detection), getting her name and description (sequencing), or having her friend meet up with her while wearing bright clothing that could be identified from a distance (hybridization). The word “isolated” is certainly used in a non-physical, informational sense, such as “to isolate a problem”; in the context of composition-of-matter claims claiming a physical DNA molecule, however, it seems less plausible that “isolation” would refer to a non-physical identification or tagging process without actual physical separation. This is borne out by the language of the sample claims above; the BRCA patent defines “isolated” DNA as “removed from its naturally occurring environment,” U.S. Patent No. 5,747,282 (filed June 7, 1995) (emphasis added), not merely detected, sequenced, or hybridized.

Judge Bryson, in his dissent in *AMP II*, made a parallel argument that “[I]f we are to apply the conventional nomenclature of any field to determine whether Myriad’s isolated DNA claims are ‘new,’ it would seem to make more sense to look to genetics, which provides the language of the claims, than to chemistry.” *AMP II*, 653 F.3d at 1376 (Bryson, J., concurring in part and dissenting in part). The genetics nomenclature is an informational one: “From a genetic perspective, that [composition] claim covers one ‘composition of matter’—the BRCA1 gene.” *Id.* However, as described in detail *infra*, this reasoning was not adopted by the majority.

<sup>68</sup> *AMP II*, 653 F.3d at 1352.

<sup>69</sup> *Id.* at 1351.

<sup>70</sup> *Id.* at 1352.

description, however, does not explicitly set the boundaries of “isolation”—that is, how isolated does a DNA molecule have to be to be “isolated DNA?”

Neither the patent-based definitions nor *AMP II*'s description specifies the necessary degree of isolation, or the boundaries of that definition. This is not wholly surprising; in an individual diagnostic testing case, the targeted sequence usually is precisely what is covered by the patent claim. If a researcher amplifies a patient's BRCA2 gene to look for mutations, that amplified DNA sequence is precisely what is described in the claims of Myriad's BRCA2 patents. In the case of WGS, to the contrary, sequencing is not targeted, so the boundary of what counts as “isolated” becomes significantly more important, as will be described below.<sup>71</sup>

While the definitions available do not specify boundaries, they do suggest a set of potential breadth-based definitions for isolation, which are summarized in Table 1. The narrowest definition conceivable would seem to be perfect and complete isolation: only the sequence described, with no flanking sequences (i.e., nucleotides attached to either end), no bound proteins or other nucleic acids (e.g., ribosomes, polymerases, or histones), and no other cellular components or structures (e.g., lipid membranes). This seems to be the definition hinted at, though not stated, by the Federal Circuit.<sup>72</sup> A slight broadening would include only substantial separation from those other elements, rather than total separation. Finally, a still broader definition would include substantial separation from other cellular elements, but would allow longer flanking DNA sequences.<sup>73</sup> How long can flanking sequences be while retaining “isolation”? In an intuitive, plain-meaning sense, “isolated” most likely captures isolation from adjacent genes, but could potentially include

---

<sup>71</sup> See *infra* Part II.

<sup>72</sup> Judge Moore stated that, “Instead of being connected to many thousands of additional nucleotides at the 3' and 5' ends of the sequence in question, as is the case in the chromosome, the isolated DNA molecules terminate in, for example, a hydroxyl and a phosphate group, respectively.” *AMP II*, 653 F.3d at 1363 (Moore, J., concurring in part). This language implies a very stringent definition of “isolated DNA” as a chemically defined single molecule, with precisely delineated ends—that is, not as a mixture of a population of very similar DNA fragments with some impurities coming from other cellular components. The majority opinion echoes this conception, less obviously, in its repeated description of isolated DNAs as “distinctive chemical molecules,” *id.* at 1351 (majority opinion), or having a “distinctive chemical identity,” *id.* at 1351–52, 1354.

<sup>73</sup> This would involve some linguistic twisting for the nucleic-acid-specified claims, since it would construe “isolated” to refer to the nucleic acid as opposed to non-nucleic-acid cellular components (i.e., including nontrivial flanking sequences), but would then presumably refer only to the coding region of that nucleic acid for purposes of specifying the sequence. This dichotomy does not occur in the composition claims based on the encoded protein or on the acid containing the short sequence overlaps, since in each of these types of claim it is clear that the nucleic acid merely contains the identifying sequence, and does not consist *entirely* of that sequence.

them under a very broad reading of the claims. “Isolated,” however, certainly includes separation from the chromosome taken as a whole, as the court laid out in *AMP II*.<sup>74</sup> Therefore, for the remainder of this analysis, the broadest feasible construction of “isolated nucleic acid” will be used: a nucleic acid sequence which is substantially isolated from other cellular components but which may contain non-trivial flanking sequences, including other genes but not a large fraction of the chromosome.

**Table 1. Potential Definitions of “Isolated Nucleic Acid”**

Breadth	Name	Cellular Components	Flanking Sequences
Narrowest	Completely Isolated	None	Nothing
Narrow	Substantially Isolated	Substantially none	Minimal sequences
Broad	Non-trivial Flanks	Substantially none	Non-trivial sequences, but no adjacent genes
Broadest	Adjacent-gene Flanks	Substantially none	Adjacent genes, but not large chromosome fractions

**B. *Method Claims Usually Cover the Comparison of Sequence to References***

The second important class of claims frequently found in gene patents contains method claims relating to using the genetic information for diagnostic purposes. Obviously, this type of claim typically only appears (or at least is only of commercial importance) in diagnostic gene patents. These method claims generally comprise two main steps, which are sometimes conflated. The first step (when included) involves determining (or observing or detecting) the sequence of the gene as possessed by the patient, and the second step involves comparing that sequence to

<sup>74</sup> *AMP II*, 653 F.3d at 1351–53 (contrasting isolated DNA with native DNA as found in a large DNA molecule as part of a chromosome). Aside from the extraordinary lexical broadness that would be required to construe an isolated DNA as including an entire chromosome, that broad a definition would be practically useless. If the meaning of an “isolated DNA” including a specified sequence could encompass the entire chromosome containing that sequence, if separated from other cellular components, the vast majority of DNA patents drafted in this form would be invalid, since they would be anticipated by the first patent claim for any sequence within that chromosome.

the sequences identified in the gene patent. A third step may sometimes be specified, sometimes only as a “wherein” clause, involving using the differences between the sequences to diagnose, for example, a predisposition to the disease connected to the gene. These method claims are less uniformly written, but a few examples demonstrate their general formulation. Duke’s Patent No. 5,508,167 (covering ApoE, a gene associated with Alzheimer’s disease) has as claim 1:

A method of detecting if a subject is at increased risk of developing late onset Alzheimer’s disease (AD) comprising directly or indirectly: *detecting* [step 1] the presence or absence of an apolipoprotein E type 4 isoform (ApoE4) in the subject; and *observing whether or not the subject is at increased risk* [step 3] of developing late onset AD by *observing if the presence of ApoE4 is or is not detected* [step 2], wherein the presence of ApoE4 indicates said subject is at increased risk of developing late onset AD [step 3].<sup>75</sup>

Myriad’s Patent No. 5,753,441 (again, covering BRCA1) includes steps 2 and 3 in its claim 1:

A method for screening germline of a human subject for an alteration of a BRCA1 gene which comprises comparing [step 2] sequence of a BRCA1 gene or BRCA1 RNA from a tissue sample from said subject or a sequence of BRCA1 cDNA made from mRNA from said sample with germline sequences of wild-type BRCA1 gene, wild-type BRCA1 RNA or wild-type BRCA1 cDNA, wherein a difference in the sequence of the BRCA1 gene, BRCA1 RNA or BRCA1 cDNA of the subject from wild-type indicates an alteration in the BRCA1 gene in said subject [step 3].<sup>76</sup>

While method claims are less uniformly written than the composition-of-matter claims described above, they tend to claim some variation of the three steps above; they can be broadly classed based on whether they explicitly include a sequencing step, which the Federal Circuit has indicated may be crucial for their validity.<sup>77</sup> The three-step process is certainly a prototypical method of diagnosing a disease based on gene sequence, and will be used in the analysis of WGS’s potential infringement of individual gene patents.

Overall, though both method claims and composition-of-matter claims have some variability in their individual drafting, they are constrained by both practice (e.g., the relatively consistent types of uses for biotechnological and diagnostic patents<sup>78</sup>) and rule (e.g., the USPTO’s Utility Examination Guidelines which pre-approve “isolated” nucleic

---

<sup>75</sup> U.S. Patent No. 5,508,167 (filed Apr. 13, 1994) (emphasis and step notation added).

<sup>76</sup> U.S. Patent No. 5,753,441 (filed Jan. 5, 1996) (emphasis and step notation added).

<sup>77</sup> *AMP II*, 653 F.3d at 1355–57.

<sup>78</sup> See *supra* notes 24–31 and accompanying text.

acid claims<sup>79</sup>) to a certain domain of potential claim scope. Although a comprehensive analysis of all gene patents is far beyond the scope of this work, the claim structures described above are enough to analyze WGS infringement at a broad level. The following text lays out that analysis for composition claims (Part II) and method claims (Part III).

## II. DOES WHOLE-GENOME SEQUENCING INFRINGE COMPOSITION-OF-MATTER CLAIMS TO ISOLATED GENES?

Gene patents are not aimed at including general sequencing methods, so their claims lack the broad scope which would easily span multiple methodological variations. Instead, as described above,<sup>80</sup> gene patent claims focus specifically on two types of inventions: composition claims on DNA molecules including the entire or partial sequence, and method claims based on sequence comparison. Whether WGS infringes the first type of claim—that is, the claims for DNA molecules—depends on the specific sequencing technology used. Whether WGS infringes the second type of claim, for drawing medical correlations, depends less on the technology itself and more on the commercial structure surrounding the sequencing. However, this general class of correlation claim is invalid under *AMP II*<sup>81</sup> based on the reasoning of *Bilski v. Kappos*<sup>82</sup> and *Prometheus Laboratories v. Mayo Collaborative Services*,<sup>83</sup> unless sequence-determining steps are explicitly included in the method.

The determination of whether WGS infringes a typical extant composition-of-matter gene patent claim depends on the precise contours of the sequencing technology employed—namely, whether the technology ever actually creates or uses the *specific* “isolated” nucleic acids claimed in the patent. If those specific claimed nucleic acids are never created or used in the sequencing process, there is no infringement. For these claims, technology determines infringement. Several different versions of sequencing technology have been developed to read entire genomes. This Article focuses first on the two sequencing techniques which have been well-developed and validated, and which are in common use today: hierarchical sequencing and shotgun sequencing. It also describes and analyzes a novel and still developing technique: nanopore sequencing. The contours of possible infringement are substantially different for the different sequencing techniques.

---

<sup>79</sup> Utility Examination Guidelines, *supra* note 22, at 1093.

<sup>80</sup> See *supra* Part I.

<sup>81</sup> *AMP II*, 653 F.3d at 1355–58.

<sup>82</sup> 130 S. Ct. 3218 (2010).

<sup>83</sup> 628 F.3d 1347 (Fed. Cir. 2010), *cert. granted*, 131 S. Ct. 3027 (2011).

An initial technical point must be made about DNA sequencing as it is generally described. The term “sequencing” is unfortunately used for two distinct processes, though both do, in fact, determine sequences of bases in a nucleic acid. On a smaller scale, sequencing generally refers to a *specific experiment* whereby the nucleotide order of a nucleic acid is determined.<sup>84</sup> This is the type of sequencing meant in the phrase “sequence a gene.” This type of sequencing, however, faces a technical limitation: it can only determine the sequence of a relatively small stretch of DNA—generally, only several hundred bases. This is the central challenge of sequencing—at most 1000 bases can be determined in a single reaction,<sup>85</sup> but the smallest human chromosome is made up of approximately forty-seven million base-pairs and contains hundreds of genes.<sup>86</sup> Enter large-scale sequencing. Used in this sense, “sequencing” refers *not* to the small-scale sequencing experiments which actually “read” bases (e.g., ACCTGTAACG . . .), but to the assembling of many such shorter sequences (determined from the smaller-scale sequencing reactions) to create accurate sequences for larger stretches of DNA. A useful (if non-linear) analogy might be a microfiche reader with poorly functioning movement controllers—on a small scale, one “reads” the microfiche based on whatever fragment is viewable at any given time, and in a large scale, one could “read” the document by recording each of those fragmentary views and then fitting them together into one coherent whole. For DNA sequencing, the traditional mode of small-scale reading is the well-established Sanger sequencing reaction; the principal differences in large-scale reading form the distinction between the main techniques of sequencing used today.

A. *“Traditional” Whole-Genome Sequencing Techniques: De Novo and Reference Sequence Assembly of Shotgun Sequences*

The two main well-established techniques for WGS both rely on the traditional paradigm of two-step sequencing described above: that is, essentially random and duplicative shotgun small-scale sequencing

---

<sup>84</sup> In particular, sequencing today is overwhelmingly performed by the Sanger method, first described in 1977. F. Sanger et al., *DNA Sequencing with Chain-Terminating Inhibitors*, 74 PROC. NAT'L ACAD. SCI. 5463 (1977).

<sup>85</sup> Erik Pettersson et al., *Generations of Sequencing Technologies*, 93 GENOMICS 105, 106 (2009). Note that experimental apparatuses exist to perform many such sequencing reactions in parallel—up to 384 at once, in fact—but this still generates separate sequences that then need to be non-trivially joined together. *Id.*

<sup>86</sup> See generally M. Hattori et al., *The DNA Sequence of Human Chromosome 21*, 405 NATURE 311 (2000).

followed by assembly into a larger whole.<sup>87</sup> In each technique, the small-scale sequencing follows the same procedure. First, genomic DNA is isolated from other cellular components, including not only lipid membranes and general cellular proteins, but also components more closely associated with DNA, such as histones and polymerases. In an ordinary sequencing experiment, many copies of the genomic DNA are isolated in this step.<sup>88</sup> Second, the genomic DNA is broken down into small fragments, generally ranging from twenty-five to 1000 bases in length.<sup>89</sup> Third and finally (in the small-scale sequencing stage), those fragments are sequenced base-by-base, frequently using the Sanger method, but potentially using other techniques.<sup>90</sup> At this stage, the two techniques diverge, although not in a way that is relevant for infringement analysis. In *de novo* assembly, the pieces are computationally assembled according to overlap without any external reference,<sup>91</sup> whereas in reference-sequence assembly, pieces are overlaid on a reference sequence (most typically, the first reference human genome sequence) to place them in order.<sup>92</sup> As a rough analogy, *de novo* assembly is like putting together a jigsaw puzzle just based on matching edges, whereas reference sequence assembly uses a picture that is similar, but not identical, to the puzzle as a guide for putting the pieces in place.<sup>93</sup>

---

<sup>87</sup> A historically informed reader might wonder about a different set of two WGS techniques, which were the competing two techniques in the race to sequence the human genome for the first time. In brief, one technique (used by Celera Genomics) was *de novo* assembly of shotgun sequences, which will be described *infra*. See generally J. Craig Venter et al., *The Sequence of the Human Genome*, 291 SCIENCE 1304 (2001). The other technique (used by the collaborative public effort), known as BAC-to-BAC, focused on generating a large-scale crude physical map of the genome using restriction site analysis of Bacterial Artificial Chromosomes containing large (approximately 150,000 bases) segments of DNA, and then mapping smaller sequences onto that large-scale physical map. See generally Int'l Human Genome Sequencing Consortium, *Initial Sequencing and Analysis of the Human Genome*, 409 NATURE 860 (2001). BAC-to-BAC sequencing, however, only needs to be done once for any given genome—once a single genomic sequence of relatively high quality has been completed, that sequence can serve as a rather less crude physical map for assembling small-scale sequences when sequencing additional individuals' genomes of the same species. *Id.* The analog of BAC-to-BAC sequencing in current efforts to sequence individual human genomes is reference sequence assembly of shotgun sequences—but the experimentally cumbersome process of creating the initial physical map using BAC-to-BAC is no longer necessary for human sequencing.

<sup>88</sup> Ng & Kirkness, *supra* note 10, at 216.

<sup>89</sup> *Id.*

<sup>90</sup> *Id.* at 216–17.

<sup>91</sup> *Id.* at 216.

<sup>92</sup> *Id.*; David A. Wheeler et al., *The Complete Genome of an Individual by Massively Parallel DNA Sequencing*, 452 NATURE 872, 876 (2008).

<sup>93</sup> The question might arise as to why anyone would use a *de novo* approach, which is much more computationally intensive. The answer is that *de novo* assembly is much better at dealing with novel regions that are not present in the reference sequence. Ng & Kirkness, *supra* note 10, at 216. On the other hand, *de novo* assembly can fail to properly take account of repeating segments, which are frequent in the human genome. *Id.*

The question of infringement for these two established sequencing techniques is: Do these techniques make or use an “isolated” nucleic acid as specified by one of the typical sequence specifications in gene patent composition-of-matter claims? These two parts will be analyzed separately: isolation will be analyzed first, followed by the different types of sequence specification described above.

“Isolated,” as discussed previously, can take a range of different plausible meanings. All of the plausible meanings of an “isolated nucleic acid” include isolation from proteins and other cellular materials, and, on this count, both of the traditional DNA sequencing techniques likely are covered.<sup>94</sup> The question of whether DNA in genetic sequencing is “isolated” from other *genetic* material—that is, from flanking sequences—is only slightly more complicated. Both *de novo* assembly and reference sequence assembly involve the creation and use of fragmented DNA sequences.<sup>95</sup> In current sequencing practice, these fragments are usually between 25 and 1,000 bases in length.<sup>96</sup> Under the Narrowest or Narrow definition, these fragments could potentially include minor flanking sequences sufficient to defeat isolation. However, under the far more likely Broad definition, as was adopted in the district court’s construction in *AMP I*,<sup>97</sup> (or, a fortiori, under the Broadest definition) as long as these flanking sequences did not include other genes, the fragment would be considered “isolated DNA.”

Since either of the main techniques likely creates and uses “isolated” DNA as claimed in composition claims, the remaining question is whether the fragments are those specified in the gene patent claims by one of the three general specification methods. This analysis first requires a close look at the size of the fragments and of human genes as differently specified.

The first two main types of sequence specification are functionally equivalent here (assuming both types are valid).<sup>98</sup> In each, infringement requires an isolated nucleic acid that is essentially the complete sequence, in some cases with some flexibility as to whether the match needs to be perfect.<sup>99</sup> Therefore, for these claims to be infringed, the

---

<sup>94</sup> It is true that if “isolated” is construed to mean completely separated from other cellular components, then DNA sequencing may fail to infringe if the particular experimental protocol fails to perfectly purify the DNA. This construction is unlikely, however—relatively little in the biological world is ever 100% pure, nor does it usually need to be. If, as seems much more likely, “isolated” means substantially or operationally isolated from other cellular components—that is, the sense of “isolated” that practically is necessary for use of DNA as DNA—then sequencing almost certainly involves DNA that is “isolated” from other cellular components.

<sup>95</sup> Ng & Kirkness, *supra* note 10, at 216.

<sup>96</sup> *Id.*

<sup>97</sup> *AMP I*, 702 F. Supp. 2d 181, 216–17 (S.D.N.Y. 2010).

<sup>98</sup> See *supra* notes 37–45 and accompanying text.

<sup>99</sup> See *supra* notes 40–42 and accompanying text.

fragment generated and used in the WGS process—that is, the fragments between twenty-five and 1000 bases used in the small-scale sequencing—must contain the entire gene sequence. The identity analysis above shows that every potential sequence is expected to occur multiple times in the process, so the remaining constraint is size. For instance, if a gene is 1200 bases long, then even a perfectly aligned 1000 base fragment will not contain the entire sequence. Therefore, whether a WGS technique infringes the whole-gene-sequence-defined composition claims of a gene patent depends on the specific size of the fragments generated, and whether the gene claimed is smaller than that size. Considering actual size distributions, a very small (but not zero) number of genes could be the subject of whole isolated gene claims which would be infringed by current sequencing methods. Human genes vary tremendously in size, from hundreds to millions of bases.<sup>100</sup> There are certainly human genes which can be found in their entirety on 1000 base fragments created as part of WGS. There are not, however, very many.<sup>101</sup> BRCA1, for instance, spreads over more than 81,000 bases of genomic DNA. Overall, it seems that few composition claims for whole isolated genes would be infringed by traditional WGS sequencing.

The third type of sequence claim is much more widely applicable, but is also significantly less likely to be valid. This type of claim claims isolated nucleic acids with a matching stretch that consists of a much smaller portion of the claimed gene—often fifteen or sixteen bases long. Effectively any fragments used in WGS that actually contain a portion of the claimed gene would meet this definition.<sup>102</sup> Since these fragments

---

<sup>100</sup> TOM STRACHAN & ANDREW P. READ, HUMAN MOLECULAR GENETICS 253–54 (3d ed. 2004). Note that most human genes contain non-coding stretches, so that the actual coding portion of the gene is much smaller than its genomic extent. *Id.* at 254. Since WGS involves sequencing *genomic* DNA, however, and not only the coding sequences of the genome, the genomic extent of genes (e.g., introns *and* exons) is the relevant quantity.

<sup>101</sup> This Article focuses on protein-encoding genes—that is, genes which code for a protein that will be produced by cells. The analysis of gene coverage is easily extendable to ribosomal RNA or transfer RNA genes—that is, genes which code for RNA that performs functions in the cell itself, rather than being encoded into a protein (in the common nucleic-acid-as-blueprint analogy, this would be the use of blueprints as flyswatter, window-covering, or wall art rather than as a plan to build something). What is *not* so easy to consider, however, is small interfering RNA genes (siRNA), which are only a few dozen bases in length. siRNA genes inhibit the expression of other genes, and have a wide range of potential uses. See generally Stefan Maas, *Gene Regulation Through RNA Editing*, 10 DISCOVERY MEDICINE 379 (2010); Mouldy Sioud, *Promises and Challenges in Developing RNAi as a Research Tool and Therapy*, 703 METHODS MOLECULAR BIOLOGY 173 (2011); S. Patrick Walton et al., *Designing Highly Active siRNAs for Therapeutic Applications*, 277 FEBS J. 4806 (2010). However, since siRNA patents are only just beginning to emerge, Charlie Schmidt, *Negotiating the RNAi Patent Thicket*, 25 NATURE BIOTECHNOLOGY 273 (2007), they are not considered in depth here. Should they become a significant patent presence, the two traditional WGS techniques would almost certainly infringe, since they would function essentially as valid short-sequence gene claims; however, nanopore sequencing and other single-molecule methods would remain non-infringing.

<sup>102</sup> The trivial exception consists of those cases where a fragment included less than fifteen (or so) bases of the gene at one end—for instance, a fragment which had as its last twelve bases

are almost certainly “isolated,” these claims would very likely be literally infringed by either traditional technique of WGS. However, as described above, this form of claim is very likely invalid in all or nearly all gene patents,<sup>103</sup> so this form of infringement should present little in the way of a practical problem. There are exceptions: less-sophisticated or less-funded actors who may be deterred by the mere threat of literal infringement without consideration of its vulnerability to legal challenge, and those who may believe, as described above, that “isolated” has a fundamentally different meaning in diagnostic cases.<sup>104</sup>

Overall, it appears that the two traditional WGS techniques are likely to infringe a small number of whole-gene claims and a much larger number of potentially invalid short-sequence claims. Notably, either of these processes could be modified to work around all, or almost all of the whole-gene patent claims by changing the size of the fragments generated; with fragment sizes below a few hundred bases, no whole-gene claims would be infringed. This process would not be costless—the fragment sizes used were picked for computational and experimental reasons, and changing them would require altering protocols and potentially adding additional redundancy-checking to avoid introducing sequencing errors—but is certainly possible.

One potential counterargument that could be made for infringement which obviates the size workaround is that the fragments are reassembled in the process of whole sequence generation. This argument, however, ignores the information/chemical distinction. DNA is a chemical compound, and it is as such that gene patents covering it are granted. It is also conceived as a set of information; however, this is not the subject of composition-of-matter gene patents.<sup>105</sup> The *chemical* fragments are never reassembled into larger pieces in the process of sequence generation; only the *information* fragments are rejoined. Therefore, the process of assembling the information cannot itself constitute a source of patent infringement.

---

the first twelve bases of the claimed gene. Since all possible fragments of the specified size are stochastically expected, this exception has no practical effect.

<sup>103</sup> See *supra* notes 49–58 and accompanying text.

<sup>104</sup> See *supra* note 67.

<sup>105</sup> The Federal Circuit’s description of the isolated DNA covered by the composition-of-matter claims describes them exclusively as chemicals. *AMP II*, 653 F.3d 1329, 1349–55 (Fed. Cir. 2011). DNA sequences are described as informational constructs in the patent and the opinion; however, those sequences are not compositions of matter, but the subjects of method claims covering comparison of different sequences. *Id.* at 1355–57. DNA sequences (not molecules) as covered by the method claims are analyzed *infra* Part III.

B. *A Novel Sequencing Technique: Nanopore Sequencing*

In stark contrast to the relatively familiar procedures of two-scale sequencing, the new and still developing technology of nanopore sequencing relies on completely different processes and, consequently, requires an alternate infringement analysis. It is worth noting at the outset that nanopore sequencing is perhaps years away from commercial use. However, because of the radically different technology involved, it appears to be one of the most promising options for the actual widespread advent of WGS—that is, for the popular conception of a machine which can read a person’s genome, as opposed to a relatively time-consuming and laborious process that needs to be performed offsite. Finally, nanopore sequencing, though only one of a set of new and developing third-generation sequencing technologies, exemplifies many of the common features of these new technologies, such as very long read lengths and single-molecule-based sequencing.<sup>106</sup> As such, it merits careful consideration in the WGS arena.

Nanopore sequencing aligns much more closely with lay views (and legal analogies) of DNA sequencing. In particular, it actually involves sequentially reading a long stretch of DNA, one base at a time.<sup>107</sup> The concept is simple. In essence, nanopore sequencing involves feeding a long strand of DNA through a very small hole (so that only one base at a time fits in the hole) and determining the shape of each base as it passes through the hole by measuring how it blocks the hole.<sup>108</sup>

Although conceptually simple, nanopore sequencing is technically challenging.<sup>109</sup> The experimental apparatus is key. A nanopore is created in some barrier; a nanopore is simply a hole of around one nanometer in diameter, and can be protein embedded in a membrane,<sup>110</sup> an etched

---

<sup>106</sup> Eric E. Schadt et al., *A Window into Third-Generation Sequencing*, 19 HUM. MOLECULAR GENETICS R227, R228 (2010); Chandra Shekhar Pareek et al., *Sequencing Technologies and Genome Sequencing*, 52 J. APPLIED GENETICS 413, 419 (2011).

<sup>107</sup> Daniel Branton et al., *The Potential and Challenges of Nanopore Sequencing*, 26 NATURE BIOTECHNOLOGY 1146, 1146 (2008). This is in contrast to the prior forms of sequencing at each level of generality: large scale, which is the jigsaw-like assembly of smaller sequences; and small-scale, where the sequencing experiment itself does not involve sequential reading of bases, but the rather complex Sanger process. There, many, many copies of the same fragment are made with the addition of a small portion of labeled bases (labeled differently for A,C,G, and T) that cause termination—by sorting the resulting copies based on length, the sequencer can “read” the last base in each slightly longer fragment, and thereby read the sequence of the original source.

<sup>108</sup> *Id.* at 1148–49; Winston Timp et al., *Nanopore Sequencing: Electrical Measurements of the Code of Life*, 9 IEEE TRANSACTIONS NANOTECHNOLOGY 281, 283 (2010).

<sup>109</sup> See *supra* note 108.

<sup>110</sup> Branton et al., *supra* note 107, at 1146.

hole in a silicon or grapheme barrier,<sup>111</sup> or some other manifestation. The barrier with the nanopore is placed in a conductive fluid, and voltage is applied across the barrier.<sup>112</sup> The movement of ions in the fluid through the pore creates a very small current, which can be detected.<sup>113</sup> When a strand of DNA is threaded through the hole, the size of the pore means that only one base can pass through at a time (the strand cannot be doubled-up, for instance)—this ensures that the bases are encountered in sequence.<sup>114</sup> As each base passes through the pore, it blocks the pore to a certain characteristic degree, which can be read as small fluctuations in the current.<sup>115</sup> This technique, unlike traditional sequencing, can be performed on a single molecule of DNA—indeed, it can only be performed on a single molecule at a time.<sup>116</sup>

Perhaps more importantly, unlike traditional sequencing, in which long DNA strands are broken down into much smaller fragments, in nanopore sequencing no such fragmentation is necessary.<sup>117</sup> Theoretically, an entire chromosome could be sequentially read from end to end in a seamless process.<sup>118</sup> Alternatively, in some versions of the technique, the DNA strand is *initially* fed into the apparatus in one strand, but immediately before passing through the pore and being read, each nucleotide is cut off from the remaining strand and passes through the pore as a nucleotide unconnected to the strand; this reduces the potential interference from adjacent bases which can affect the current despite not actually being in the center of the pore.<sup>119</sup>

For this sequencing technique to infringe the composition claims of gene patents, as before, it would have to create or use an isolated nucleic acid which is specified by one of the three prototypical forms of sequence claims. It seems highly unlikely that either form of nanopore sequencing would infringe. In the non-cleaving variation, the DNA molecule which is used for sequencing is indeed isolated from other cellular components, but certainly not from other genetic material in the chromosome—the strands being sequenced are by design extremely long, and can be as long as an entire chromosome. This is clearly outside the broadest feasible definition of “isolated”—so even if the strand does include the specified sequence, which it would, there is no infringement. In the cleaving variation, on the other hand, the DNA exists in one of two forms—either the single long strand, which has just been shown not

---

<sup>111</sup> Timp et al., *supra* note 108, at 286.

<sup>112</sup> *Id.* at 283–84.

<sup>113</sup> *Id.* at 283.

<sup>114</sup> *Id.* at 283–84.

<sup>115</sup> *Id.* at 283–85.

<sup>116</sup> *Id.* at 284.

<sup>117</sup> *Id.* at 283.

<sup>118</sup> *Id.* at 283, 292.

<sup>119</sup> *See, e.g.*, Branton et al., *supra* note 107, at 1149.

to infringe, or the single nucleotides after cleavage, which are actually “used” in the sequencing step itself. Those single nucleotides clearly do not fall into any of the sequence-specifying aspects of the composition claims.

Overall, with respect to the composition-of-matter claims in gene patents, the infringement status of WGS depends on the specific sequencing technique used and even the operational parameters of that technique. The two traditional methods of sequencing likely infringe a small number of whole-gene claims, but can probably be modified to avoid infringement of most of those claims. Either of the traditional techniques almost certainly infringes many short-sequence gene claims, but those claims are very unlikely to withstand validity analysis. Finally, nanopore sequencing, though still a nascent technology, does not appear to infringe any of the composition claims of mainstream gene patents.

III. DRAWING MEDICAL CONCLUSIONS AND COMMUNICATING WITH THE PATIENT LIKELY INFRINGES CURRENT PATENTS, BUT THOSE CLAIMS ARE OF QUESTIONABLE VALIDITY UNDER *ASSOCIATION FOR MOLECULAR PATHOLOGY*

Whether whole-genome scanning infringes gene patents’ composition-of-matter claims or not, it may also infringe the diagnosis-based methods claims present in many of those patents. Unlike gene composition claims, which are generally valid, however, genetic testing diagnosis methods claims are on far shakier validity footing; one broad class is certainly invalid under current law. Therefore, this Part will briefly (though not exhaustively) analyze the serious validity concerns in diagnosis methods patents. Then, assuming validity of at least some forms of method claims, it will indicate the ways in which WGS likely violates those claims.

As described in more detail above,<sup>120</sup> these method claims can include: (1) determining the gene sequence in the patient; (2) comparing that gene sequence to the wild-type sequence (or identified mutations) in the patent; and (3) using that similarity or difference to diagnose the medical condition linked to mutations in that gene. So, for instance, to practice the method claimed in Duke’s ApoE patent, one would sequence the patient’s ApoE gene, compare those with the wild-type sequence in Duke’s patent, and then draw a conclusion of mutation from any differences. Presumably the next step is to inform the patient; how-

---

<sup>120</sup> See *supra* Part I.B.

ever, any attempts to improve patentability by including this in the claimed method are likely to be unsuccessful.<sup>121</sup>

The analysis of diagnostic methods claims is somewhat more complex due to the larger variation in precise claim-drafting terms. The claims can involve the step of actually determining the sequence of presence of a gene variant (as in the ApoE patent), or can assume that the sequence has already been determined (as in the BRCA1 patent). They can also optionally involve a step of making the actual diagnosis, or can just leave that as a “wherein” clause attached to the comparison. For the purposes of this analysis, these method claims can be divided into essentially two classes. In the first, the methods claim includes only comparison and diagnosis steps—which can be classified as mental processes. In the second type of claim, the method also explicitly includes a step of determining the sequence of the gene at issue.

The first type of methods claim, a comparison-only method, exemplified by the BRCA1 patent, has been held by the Federal Circuit to be unpatentable under section 101.<sup>122</sup> In a literal interpretation of the claims, looking at and comparing two pieces of paper with sequences, one from a patient and one from a reference sequence, would infringe the patent. In *AMP II*, Myriad argued that the comparison of genetic sequences implicitly required a first step of sequencing the DNA.<sup>123</sup> Therefore, it argued, the claim was not for unpatentable “abstract mental processes,” but for an integrated method which satisfied the machine-or-transformation test, which under *Bilski* provides a “useful and important clue” for patentable subject matter.<sup>124</sup> The Federal Circuit rejected this argument; sequence comparison claims that do not explicitly claim transformative steps are unpatentable subject material under section 101 as abstract mental processes.<sup>125</sup>

The second type of methods claim actually does involve a transformation, as long as the determination of genetic sequence is found to be transformative. *Prometheus* strongly suggests that it will be.<sup>126</sup> In *AMP II*, the Federal Circuit distinguished the Myriad claims from those in *Prometheus*: “Myriad’s claims, in contrast, do not include the step of ‘determining’ the sequence of *BRCA* genes by, *e.g.*, isolating the genes from a blood sample and sequencing them, or any other necessarily

---

<sup>121</sup> See *King Pharm., Inc. v. Eon Labs, Inc.*, 616 F.3d 1267, 1278 (Fed. Cir. 2010).

<sup>122</sup> *AMP II*, 653 F.3d 1329, 1355 (Fed. Cir. 2011).

<sup>123</sup> *Id.* at 1355–56.

<sup>124</sup> *Bilski v. Kappos*, 130 S. Ct. 3218, 3226–27 (2010).

<sup>125</sup> *AMP II*, 653 F.3d at 1356–57. Myriad also argued that the limitation to genetic testing for its specific sequences made the claim patentable. However, the limitations to a specific field or application cannot make mental processes patentable. *Id.* at 1356 (citing *Bilski*, 130 S. Ct. at 3230).

<sup>126</sup> *Prometheus Labs. v. Mayo Collaborative Servs.*, 628 F.3d 1347, 1355–58 (Fed. Cir. 2010).

transformative step.”<sup>127</sup> This language strongly implies that method claims which *do* explicitly include a sequence-determining step are valid, at least in a patentable subject matter analysis under section 101.

Despite the likelihood of invalidity for at least some claims, if the diagnostic methods claims are broadly assumed to be valid, does WGS infringe? Obviously, the step of actually generating the sequence itself cannot infringe, since merely creating the sequence does not involve any type of comparison. However, the purpose of WGS is not merely to develop a three billion-letter piece of personal information; it is to have a sequence that can be used to determine useful information about the sequenced person, whether medical (a disease propensity, more commonly) or personal (ancestry, athletic propensity, or whatever else is linked to a particular gene).

Since comparison-only claims have been held invalid, the infringement analysis need only be performed for those method claims which include the transformative/determination step of actually determining the sequence.<sup>128</sup> For the determination/comparison claims, the crucial issue is whether any single actor actually performs all of the steps of the method, which is required for infringement of a method claim.<sup>129</sup>

---

<sup>127</sup> *AMP II*, 653 F.3d at 1357.

<sup>128</sup> The methods claims which involve only the comparative step are very easy to infringe since they involve essentially the most meaningful steps in the area of genetic sequencing. The goal of genetic sequencing is to determine the physical/medical/informational outcome by measuring the difference between a sequenced gene and the wild-type version (and any panel of known mutations)—this essential step is covered by the comparison-only claims. Therefore, a gene-sequencing company that reported the patient’s sequence and included data on which sequence characteristics were mutations would be infringing the claim; a doctor using the sequence to diagnose would be infringing; and even a patient who was given the unannotated sequence and the set of known mutations and wild-type variations would infringe by making the comparison himself or herself. It is true that a genetic testing company would likely be loath to sue the end users for making any comparison themselves—but, practically speaking, the universe of individuals who would comfortably self-diagnose a genetic disorder without medical intervention is likely small. More importantly, the promise of individualized medicine, and of integration with the medical system, demands that any sequencing comparison actually be made within the context of medical actors; therefore, comparison-only method claims almost certainly would be infringed by suit-vulnerable actors in a WGS paradigm.

<sup>129</sup> Current law requires that a single actor either perform all steps or control and direct the performance of whatever steps it does not itself perform. *Muniauction, Inc. v. Thomson Corp.*, 532 F.3d 1318, 1328–30 (Fed. Cir. 2008); *BMC Res., Inc. v. Paymentech, L.P.*, 498 F.3d 1373, 1378–79 (Fed. Cir. 2007). This requirement was upheld by a three-judge panel of the Federal Circuit in *Akamai Techs., Inc. v. Limelight Networks*, 629 F.3d 1311, 1318–21 (Fed. Cir. 2010), *vacated sub nom. Akamai Techs., Inc. v. Mass. Inst. of Tech.*, 419 F. App’x 989 (2011) (granting rehearing en banc). The en banc Federal Circuit has requested briefs on the question, “If separate entities each perform separate steps of a method claim, under what circumstances would that claim be directly infringed and to what extent would each of the parties be liable?” *Id.* at 989. How the Federal Circuit will answer this question can only be speculated; however, unless the Federal Circuit reinterprets joint infringement in an extraordinarily sweeping fashion, infringement of determination/comparison claims will still be relatively easy to avoid. Even if agency, control, or determination is no longer a requirement for multi-party infringement of a method, some link or coordination requirement seems inevitable. In that case, infringement of

It is likely that a genetic sequencing company would prefer to make available not only the sequencing results themselves but also some level of interpretation of those results. However, since such a practice would open the company to potential liability for infringing determination/comparison claims,<sup>130</sup> it might well choose to make only the unannotated sequence available to a patient or his physician, who would then compare the personal sequence to a database of wild-type and known mutant sequences. This procedure would separate the two steps of the method, and no single individual would then be using the entire method, thus avoiding infringement.

Overall, it seems that diagnostic method claims that involve only a comparative step are easy for a company, doctor, or patient to infringe, but are highly vulnerable to validity attacks based on their resemblance to purely mental processes or abstract ideas. Method claims that involve both sequence determination and comparison steps, on the other hand, fall more cleanly under the sufficient-but-not-necessary machine-or-transformation test, but are relatively easy to circumvent by having different legal actors perform the sequence-determination and comparison/diagnosis steps, as is likely to be the norm in an era of widespread WGS, where the genome need only be sequenced once, but potentially interpreted many times in different circumstances by different individuals.

#### CONCLUSION

To a certain extent, the precise contours of infringement of the thousands of extant gene patents tell only part of the story. Certain sequencing techniques are likely to infringe many composition-of-matter claims for isolated genes, although they can potentially work around that infringement by changing the parameters of traditional techniques or by using nanopore methods, which appear more suitable for widespread WGS in any case. The diagnostic methods claims are likely either invalid or avoidable. However, a crucial aspect of the infringement land-

---

a determination/comparison claim can be avoided by the widespread availability of WGS itself. Once WGS to determine a patient's genomic sequence becomes a normal background step—a one-time determination like blood-typing—then future sequence comparisons or analyses are uncoupled from the determination step, and a determination/comparison claim would lack the necessary coordination for multi-party infringement.

<sup>130</sup> A practice of actually providing diagnoses, rather than just information, might also run afoul of corporate-practice-of-medicine laws, unlicensed practice of medicine across state lines, and other various potential sources of liability, so that sequencing companies might for other reasons prefer to separate the diagnosis/interpretation functions from the sequence-providing functions—as, indeed, some sequencing companies already do. See, e.g., *Terms of Service*, 23ANDME, <https://www.23andme.com/about/tos> (last visited Feb. 15, 2012).

scape is that at least some uncertainty remains, which may itself have negative impacts. This Article has attempted to construct general analyses using the typical language of gene patents, and apply that language to generally used sequencing techniques. It is difficult to impossible, however, to rule out the possibility that among the thousands of extant human gene patents, there may be some which could be harder or unfeasible to work around. This uncertainty casts a pall of indefinite strength over the enterprise of WGS and personalized medicine; even if in actuality no or almost no valid patents are actually infringed, the looming threat of infringement lawsuits, and the costs of determining patent invalidity may deter some market actors.<sup>131</sup>

Those who argue for broad policy options with respect to diagnostic genetic testing do so based on a perception that WGS infringes many gene patents. Whether it does so or not—and whether the techniques can be or are modified to avoid infringement—policy options to shield diagnostic testing from infringement lawsuits are likely to smooth the way for personalized medicine. These policy options have been described in greater depth elsewhere, but include a research exemption for diagnostic testing (which would allow the improvement of tests, but not solve the issue of WGS), a generalized exemption from infringement for all diagnostic use (which would upset companies like Myriad but open wide the path for personalized medicine), and mandatory gene-patent clearinghouses, which could eliminate the problem of holdouts and provide for at least some revenue sharing. One final approach is the elimination of gene patents altogether, either through statutory action (which seems unlikely) or through court action through the likely vehicle of the *Association for Molecular Pathology* case. Needless to say, the elimination of gene patents would obviously allow WGS and personalized medicine without fear of infringement.

An additional policy consideration is the availability of infringing options abroad. Sequencing and sequence comparison could easily be performed abroad without infringing U.S. patents; the consumer-relevant information—that is, identification of allelic variations and drug susceptibility profiles—does not infringe gene patents. Given the ease of overseas sequencing and comparing, policy arguments could be made that attempting to strengthen gene patent protection for diagnostic testing is likely to do more than drive those tasks, jobs, and dollars overseas.

---

<sup>131</sup> See, e.g., *AMP II*, 653 F.3d at 1355 (Bryson, J., concurring in part and dissenting in part) (“Even if many of those patents include claims that are invalid for anticipation or obviousness, the costs involved in determining the scope of all of those patents could be prohibitive.”) (citing SACGHS REPORT, *supra* note 17, at 51–52; and Rebecca S. Eisenberg, *Noncompliance, Nonenforcement, Nonproblem? Rethinking the Anticommons in Biomedical Research*, 45 HOUS. L. REV. 1059, 1076–80 (2008)).

2012]

*UNBLOCKED FUTURE*

1631

Overall, it seems that infringement of intellectual property is far less of a systematic and pervasive barrier to WGS and personalized medicine than is generally assumed. Isolated gene claims may be avoidable by slightly changing the old techniques or using the new nanopore technique, which shows tremendous promise, despite its being some distance from commercial use (a characteristic of all WGS possibilities). And determination/diagnosis claims seem either invalid or avoidable as well. Perhaps policy changes would make the landscape a less risky one for innovators going forward, but it appears that, even with the status quo, WGS may be able, cautiously and mindfully, to proceed.