Published in final edited form as:
Genomics. 2010 May ; 95(5): 312–314. doi:10.1016/j.ygeno.2010.03.003.

Metastasizing patent claims on *BRCA1*

Thomas B. Kepler[1,*], **Colin Crossman**[2], and **Robert Cook-Deegan**[3]

[1]Department of Biostatistics and Bioinformatics, Duke University, Durham NC 27705

[2]762 Ninth St. #591, Durham NC 27705

[3]Institute for Genome Science & Policy, and Sanford School of Public Policy, Duke University, Box 90141, Durham NC 27708

Abstract

Many patents make claims on DNA sequences; some include claims on oligonucleotides related to the primary patented gene. We used bioinformatics to quantify the reach of one such claim from patent 4,747,282 on *BRCA1*. We find that human chromosome 1 (which does not contain BRCA1) contains over 300,000 oligonucleotides covered by this claim, and that 80% of cDNA and mRNA sequences contributed to GenBank before the patent application was filed also contain at least one claimed oligonucleotide. Any “isolated” DNA molecules that include such 15bp nucleotide sequences would fall under the claim as granted by the US Patent and Trademark Office. Anyone making, using, selling, or importing such a molecule for any purpose within the United States would thus be infringing the claim. This claim and others like it turn out, on examination, to be surprisingly broad, and if enforced would have substantial implications for medical practice and scientific research.

In 1998, the US Patent and Trademark Office granted Mark H. Skolnick and ten of his collaborators a patent on the human gene BRCA1 (US Patent 5,747,282). Mutations in BRCA1 confer a substantial risk for breast and ovarian cancers, with a cumulative risk of incidence by age 70 of 69% (breast) and 39% (ovarian) (1). Genetic tests to screen for these mutations in the United States are available exclusively through Myriad Genetics, the assignee of the patent. Women with a family history of breast and ovarian cancer may, through this genetic test, determine whether they carry one of the high-risk alleles, and if so, decide whether to take prophylactic action, including surgical removal of breasts, ovaries or both.

Human gene patents are controversial (2); *BRCA1* patents are currently the subject of litigation (3), and ’282 is among seven BRCA patents named in a complaint filed by the American Civil Liberties Union that is now being litigated in federal court. A hearing in the Southern district federal court of New York took place September 30, 2009, before Judge Robert W. Sweet (4). On November 2, he released an 88-page decision to continue the case (5), and heard oral arguments on February 2, 2010 (6).

---

*Corresponding Author: physical address: Thomas B Kepler, Center for Computational Immunology, 2424 Erwin Rd, Suite 1102, Durham NC 27710; kepler@duke.edu; voice: 919-681-0620.

The patent itself is complex and makes several different claims. One of these claims seemed to us particularly broad, so we investigated it, doing simple calculations to estimate its reach, and testing our findings by direct analysis of the extent of its reach within parts of the human genome. We find that, through this claim, the patent extends to portions of most genes in the human genome and likely to most genes in nature as well.

The patent first makes claim 1, to "An isolated DNA coding for a BRCA1 polypeptide, said polypeptide having the amino acid sequence set forth in SEQ ID NO:2." SEQ ID NO:2 is the 1863-residue amino acid sequence for the protein encoded by the BRCA1 gene. The patent further claims "5. An isolated DNA having at least 15 nucleotides of the DNA of claim 1." Note that claim 1 is DNA coding for the polypeptide, not for any specific gene. There are, of course, many polynucleotides that would encode the BRCA1 polypeptide. Claim 5, then, is a claim on any 15-mer oligonucleotide found in any such sequence. We estimate that the human genome contains over one million oligonucleotides covered by this claim, and that most human genes contain at least one and usually several oligonucleotides covered by the claim.

To estimate the breadth of this claim, one can perform a short computation. Accounting for bias in the usage of amino acids as reported, for example, in (7), the usage-weighted geometric mean codon degeneracy per amino acid is 3.107. Therefore, the mean number of 15-mers encoding a polypeptide of length 5 chosen at random from a vertebrate proteome is 3.107^5, about 290. There are 5,575 15-mers in BRCA1, so, if we consider all of the nucleotide sequences that encode the BCRA1 protein, there are about 1.6×10^6 15-mers embodied by the claim. There are $4^{15} = 1.07\times10^9$ different 15-mers altogether, so the probability that a 15-mer chosen at random will be covered by the claim is $p = 1.6\times10^6 / 1.07\times10^9 = 0.0015$ (roughly, 1 in 600 possible 15-mers). A typical human gene (before RNA editing) contains 10,000 bases, so, if human genes were random strings of nucleotides, one would expect a human gene to contain an average of 15 15-mers claimed under the patent.

But human genes are not random strings, so we counted the number of claimed 15-mers in a representative sample of the human genome to test the breadth of claim 5 empirically. We examined chromosome 1 (NCBI build 37.1), and counted only a subset of claimed 15-mers. The reason for counting only a subset is that there are three amino acids (serine, leucine, and arginine) that have 6-fold degeneracy. If we neglect two of the six synonymous codons for each of these amino acids, each degenerate 15-mer can be represented as a single string of 15 letters, with degenerate positions encoded by the extended IUPAC nucleotide alphabet. This representation permits a many-fold reduction in computing time that will slightly understate the degree of redundancy and breadth of claim 5.

Examining only this subset, we find over 340,000 matches of claimed 15-mers to the 250 million base pairs of chromosome 1, for an empirical hit rate of $p_{emp} = 0.00136$ per 15-mer, close to our theoretical expectation. Using this estimate, we expect about 14 infringing sequences per human gene, just one fewer than the 15 sequences per gene predicted above based on assumption of random sequence strings.

The claims being discussed are structural, that is, claims to DNA molecules, and do not restrict acts of infringement to particular uses or contexts, but should, in theory, give the patent-holders exclusive rights to make, use, sell, or import any DNA containing the claimed 15-mers in the United States, including use in research, diagnosis or other domains. These claims are not, for example, restricted to sequences actually derived from a BRCA1 sequence, or from human chromosome 17 (where BRCA1 is located), or only those 15-mers that are unique to BRCA1, or for use only in the context of risk assessment, diagnosis, treatment or research on inherited risk of breast and ovarian cancer. That is, anyone making

an “isolated” DNA that includes any one of the 15 base-pair sequences in the United States for any purpose would be infringing US patent 5,747,282. The claim thus covers an infinite number of DNA molecules of variable length if they contain any of the 1.6 million claimed 15-mer sequences.

To test the practical significance of claim 5, we examined the 713 entries in GenBank that represent complete coding sequences for human mRNAs deposited in 1994 (the year before the patent application was filed); 568 of these 713 mRNAs (80 percent) contain at least one BRCA1-derived 15-mer using the restricted codon table. Note that these mRNA sequences (or cDNAs) are shorter than typical genes, with a median length of 1902 nucleotides.

These findings suggest that there were already many sequences in GenBank covered by claim 5 at the time the patent application was filed. This further suggests that the claim should not have been granted, based on section 102 of the Patent Act (novelty). If challenged by re-examination or in litigation, claim 5 may be deemed invalid due to readily identifiable prior art covered by the claim. The claim may also be subject to challenge due to insufficient “written description,” since the sequences are not enumerated in the patent.

It is worth noting that a 1991 patent application for Expressed Sequence Tags was rejected on several grounds, including the fact that claimed 15-mer oligonucleotides were found in existing DNA sequences. This finding that 15-mers had sequence identity to many genes was published, and so publicly known by the end of 1992 (8). That particular EST patent application was abandoned by NIH in 1994. USPTO examiner James Martinell estimated at the time that to examine the reach of the oligonucleotide claims in that patent would have taken until 2035 because of the computational time required to search for matches in over 700,000 15-mers claimed, roughly half the number of molecules covered by claim 5 of Myriad’s US Patent 5,747,282. Although still computationally intensive, such sequence comparisons can clearly be done much more rapidly now. The reason that sequence identity in prior art was not identified as a bar to claim 5 by a different examiner for US Patent 5,747,282 when it was being examined between 1995 and 1998 is not clear in the ‘282 patent’s prosecution history (i.e., there is no indication of a search on 15-mer sequences having been performed in the “file wrapper” that records the patent prosecution history) (9).

Invalidation of claim 5 would not invalidate other claims in this patent, or the very broad method claims in other patents, such as claim 1 of US Patent 5,753,441, assigned to the same parties:

> “A method for screening germline of a human subject for an alteration of a BRCA1 gene which comprises comparing germline 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.”

Invalidating claim 5 might have little or no effect on the availability of genetic testing for BRCA sequences in the United States, although it might prompt the evaluation of other claims in the patents. The impact of this claim on research is very difficult to assess. The claim was clearly structured to capture oligonucleotides for hybridization assays and polymerase chain reaction (PCR)-based diagnostic methods. The most commonly used diagnostic methods are based on PCR, which is applied to over 80 distinct segments of the BRCA1 and BRCA2 genes, followed by determination of the sequence of the resulting amplicons. Such a PCR-based sequencing technique might not infringe the main

independent claims 1 and 2 of this patent, although it would likely be deemed to infringe this claim 5, as well as the method claim quoted above, and claim 16 of the ‘282 patent:

> “A pair of single-stranded DNA primers for determination of a nuycleotide [sic] sequence of a BRCA1 gene by a polymerase chin [sic] reaction, the sequence of said primers being derived from human chromosomne [sic] 17q, wherein the use of said primers in a polymerase chain reaction results in the synthesis of DNA having all or part of the sequence of the BRCA1 gene.”

The fate of claim 5 has clear implications, however, if full-genome sequence analysis becomes feasible, because it would likely be deemed to be infringed by any form of genomic sequencing, while other claims of the ‘282 and ‘441 patents would not be (because they entail cDNA or mRNA steps or PCR amplicons).

The effect of this claim on research is very difficult to assess. A PubMed search for the term “BRCA1” returned 7,107 articles. This suggests a large body of research on the gene has been published in the technical literature. Myriad has not enforced its patents against most research, with the exception of laboratories engaged in clinical research that entailed giving test results to individuals beyond their home institutions (11,12,13). Any such research that entailed analysis of DNA molecules containing *BRCA1* sequences in the United States very likely infringed this claim, however, so enforcement of this claim would have substantial impact on research. (Claims to BRCA1 sequences are somewhat narrower in other English-language jurisdictions such as Canada, Australia and New Zealand, and *a fortiori* in Europe, where the claims that emerged from opposition proceedings were dramatically narrowed.) There is a very narrow “research exemption” from infringement liability in the United States under common law, and a broader exemption for research that results in data contributed to the government for a regulated medical product or service (14). Since laboratory-developed tests are not currently subject to Food and Drug Administration approval, however, this exemption may not apply.

The simplest conclusion about the effect of claim 5 and Myriad’s other *BRCA1* patents on research and clinical testing is that Myriad has only rarely enforced its patents in research, has vigorously enforced its patents against commercial genetic testing, and has selectively enforced its patents in clinical research. It is also apparent that research on BRCA1 for the past 12 years has entailed massive, pervasive infringement of this claim, even if the claim’s scope were restricted to *BRCA1* research. Any such research in the United States was thus undertaken under risk of infringement liability and its associated uncertainty. While Myriad has stated publicly that it has not enforced its patents against basic research (11,12,15), it has not stated it will not do so in the future, and therefore *BRCA* research in the United States continues only with Myriad’s indulgence.

The strategy for claiming DNA sequences exemplified by claim 5 in ‘282 is quite broad. The patent examination manual stipulates that claims use “the broadest reasonable meaning of the words in their ordinary usage as they would be understood by one of ordinary skill in the art, taking into account whatever enlightenment by way of definitions or otherwise that may be afforded by the written description contained in applicant's specification (10).” Lines 14-4 of, column 24 in the patent define “substantial homology or similarity” as “nucleotide sequence identity in at least about 60% of the nucleotide bases,” and defines “selective hybridization” “when there is at least about 55% homology over a stretch of at least about 14 nucleotides.” These definitions explain why 15-mers were chosen, but do not alter the plain meaning of any of the terms in claim 5. Our experimental sequence comparisons also meet these definitions.

Once granted, patent claims are valid until and unless they are challenged. BRCA patents were cited in enforcement letters that Myriad Genetics sent to other laboratories to cease genetic testing for BRCA mutations (11,12).

A judgment in the current BRCA lawsuit is being closely watched by many for its implications on the practice of clinical genetic testing, patent practice, and the pursuit of research. The enormously improved computational capability to examine the reach of partially ambiguous claims should provide important guidance for those seeking and granting patents claiming DNA sequences.

References

1. Antoniou A, Pharoah PDP, Narod S, Risch HA, Eyfjord JE, et al. Average Risks of Breast and Ovarian Cancer Associated with BRCA1 or BRCA2 Mutations Detected in Case Series Unselected for Family History: A Combined Analysis of 22 Studies. Am J Hum Gen 2003;72:1117–1130.
2. Paradise J, Andrews L, Holbrook T. Patents on Human Genes: An Analysis of Scope and Claims. Science 2005;307:1566–1567. [PubMed: 15761140]
3. *Association for Molecular Pathology, et al., v. United States Patent and Trademark Office, et al*., No. 09-CV-4515 (RWS) (S.D.N.Y. filed May 12, 2009) (plaintiffs' complaint).
4. *Association for Molecular Pathology, et al., v. United States Patent and Trademark Office, et al*., No. 09-CV-4515 (RWS) (S.D.N.Y. August 27, 2009) (order setting date for hearing of plaintiffs' motion for summary judgment and jurisdictional discovery).
5. *Association for Molecular Pathology, et al., v. United States Patent and Trademark Office, et al*., No. 09-CV-4515 (RWS) (S.D.N.Y. November 2, 2009) (opinion stating the suit will continue, keeping all defendants and setting dates for motions and for a December 11, 2009, oral hearing).
6. *Association for Molecular Pathology, et al., v. United States Patent and Trademark Office, et al*., No. 09-CV-4515 (RWS) (S.D.N.Y. filed May 12, 2009) (oral arguments)
7. King JL, Jukes TH. Non-Darwinian Evolution. Science 1969:788–798. [PubMed: 5767777]
8. Martinell J. Office Action: Rejection of US Patent Application 07/837,195, “Sequence Characteristics of Human Genome Transcription Product,” (J. Craig Venter, et al.) filed February 12, 1992 1992:578–596. Notice dated August 20, 1992, reprinted in part in *Biotechnology Law Report* 11 (No. 5, September-October); also on file with authors.
9. Prosecution history for US Patent 5,747,282 on file with authors, and also available through the Patent Application Information Retrieval database (US Patent and Trademark Office) for Application Number 08/483,554 (filing date June 7, 1995; issue date May 5, 1998).
10. US Patent and Trademark Office. Manual of Patent Examining Procedure. Section 2111.
11. Cook-Deegan, R., et al. Impact of Patents and Licensing Practices on Access to Genetic Testing for Inherited Susceptibility to Cancer: Comparing Breast and Ovarian Cancers to Colon Cancers. (Case study commissioned by the Secretary's Advisory Committee on Genetics, Health, and Society, 2009; http://oba.od.nih.gov/oba/SACGHS/Appendix%201%20SACGHS%20Patents%20Consultation%20Draft%20Compendium%20of%20Case%20Studies.pdf). Notification letters cited in case study.
12. Gold, ER.; Carbone, J. Myriad Genetics: In the Eye of the Policy Storm. (International Expert Group on Biotechnology, Innovation and Intellectual Property, 2008; http://www.theinnovationpartnership.org/data/ieg/documents/cases/TIP_Myriad_Report.pdf). Notification letters cited in case study.
13. Parthasarathy, S. Building Genetic Medicine: Breast Cancer, Technology, and the Comparative Politics of Health Care. Cambridge, MA: MIT Press; 2007.
14. *Madey v. Duke University*, 307 F.3d 1351 (3 October 2002)(holding that the experimental use defense is limited to infringement performed "for amusement, to satisfy idle curiosity, or for strictly philosophical inquiry" and does not include experimentation conducted in the course of a university’s normal teaching and research activities); *Merck v. Integra Lifesciences I, Ltd., et al*., 545 U.S. 193 (2005)(holding that 35 U. S. C. §271(e)(1), the Bolar amendment, exempts “from

infringement… all uses of patented inventions that are reasonably related to the development and submission of *any* information under the [Federal Food, Drug, and Cosmetic Act]").

15. Declaration of Dr. Gregory C. Critchfield in *Association for Molecular Pathology, et al., v. United States Patent and Trademark Office, et al*. No. 09-CV-4515 (RWS) (S.D.N.Y., filed December 18, 2009).