# EXHIBIT 2

# Metastasizing patent claims on BRCA1

Thomas B. Kepler[1], Colin Crossman[2], 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 and Policy and Sanford Institute, Duke University, Durham NC 27705

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, an 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, generally surgical removal of breasts, ovaries or both.

Human gene patents are controversial (2); *BRCA1* patents are currently the subject of litigation (3). This particular patent is the first one named in a complaint filed by the American Civil Liberties Union and is now moving swiftly toward resolution. 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).

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 the claim in one is for 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 (6), 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 $1.6\times10^6$ 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. 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 conservative estimate, we expect about 14 infringing sequences per human gene. The claims being discussed are structural, 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 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.

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 (79 percent) contain at least one BRCA1-derived 15-mer using the restricted codon table. Note that these mRNA sequences 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.

The particular strategy for claiming DNA sequences exemplified by claim 5 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 (7)." 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 (8,9).

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 (10). That particular patent application was abandoned by NIH in 1994. Examiner James Martinell estimated 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 granted in claim 5 of US Patent 5,747,282. Although still computationally intensive, such sequence comparisons can clearly can be done much more rapidly now. The reason this was not done by a different examiner for US Patent 5,747,282 when it was being examined between 1995 and 1998 is not clear.

The judgment in this case is likely to have impact on the practice of biomedicine and the pursuit of research. The enormously improved capability to examine the reach of partially ambiguous claims should provide important guidance along the way.


1. Antoniou A, Pharoah PDP, Narod S, Risch HA, Eyfjord JE, et al. (2003) 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 72: 1117-1130.
2. Paradise J, Andrews L, Holbrook T (2005) INTELLECTUAL PROPERTY: Patents on Human Genes: An Analysis of Scope and Claims. Science 307: 1566-1567.
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. King JL, Jukes TH (1969) Non-Darwinian Evolution. Science 164: 788-798.
7. US Patent and Trademark Office, *Manual of Patent Examining Procedure*, Section 2111.
8. R. Cook-Deegan *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).
9. E. R. Gold, J. Carbone, *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).