Genome Medicine

LETTER TO THE EDITOR

Response to 'pervasive sequence patents cover the entire human genome' - authors' reply

Jeffrey Rosenfeld[1,2] and Christopher Mason[3,4,5*]

See related Correspondence by Rosenfeld and Mason, http://genomemedicine.com/content/5/3/27 and related letter by Tu et al., http://genomemedicine.com/content/6/2/14

Abstract

An author reply to the Letter to the Editor from Tu et al. regarding **Pervasive sequence patents cover the entire human genome** by J Rosenfeld and C Mason. *Genome Med* 2013, **5**:27.

In our previous work [1], we concluded that patents were claimed on 21 to 41% of human genes based on long (over 150 nucleotide) fragments, whereas for short (under 150 nucleotide) fragments, we showed their non-specificity meant that 100% of human genes had some portion patented. For this analysis, we relied on databases of DNA sequences included in patents provided by CAMBIA [2] and the NCBI [3] because we concluded that it is impractical to manually look through sequence files and patent applications to determine whether a specific nucleotide sequence is covered by a patent. While the CAMBIA database [2] includes all sequences present in claims, it does not distinguish between sequences that are specifically claimed from those sequences that are merely mentioned in the claims. Work by Graff et al. [4] estimated that only 8,703 patents on naturally occurring DNA sequences are still in force. Of those, they estimate that only 3,535 (41%) are human, indicating that our previous conclusions may be too broad or could lead to legal conclusions [1] that are based on an 'incorrect view of the law'. Yet, subsequent analysis with the data from CAMBIA has shown many patents are still in force and leave legal ambiguity [5].

Nevertheless, the 9–0 Supreme Court decision [6] reached effectively identical legal conclusions as we did and rendered this debate moot. The Supreme Court Justices specifically pointed out (at the bench and in the final decision) that the issue with short, patented fragments affects both DNA and cDNA molecules. They concluded that existing DNA cannot be patented, and then also concluded that cDNA is also non-patentable in cases where it has 'no intervening introns to remove when creating cDNA', or when it is the length of an exon or less. 'In that situation, a short strand of cDNA may be indistinguishable from natural DNA.' Thus, when any naturally occurring DNA or any short cDNA sequence matches the genome, at any size, it is putatively not patent-eligible, which immediately addressed the problem of short sequence non-specificity.

Also, Tu et al. claim that it should take only 60 hours to manually look through the entire set of patents containing a nucleotide sequence by only looking at the claims, but this would ignore critical definitions of terms that often appear in the text of patents. We think that the actual time required would be much higher, especially if one were trying to determine whether a sequence involved in one's research is patented. It is actually impossible to manually determine if one's sequence is close enough to a claimed sequence, especially since these claims often use 'homology' or 'similarity' measures defined beyond the claims' language. Claim non-specificity is inherent to almost all patents on genetic sequences because the precise ordering of the nucleotides for any gene is probably never the same between any two people, even for identical twins or synthesized DNA in a laboratory (owing to enzymatic errors in copying DNA). Thus, patents on DNA molecules or cDNA molecules must allow for a wide range of potential variation of a gene or sequence in order to apply to every person. To address this problem, many gene-based patents are written vaguely on purpose, including many we have previously examined [1].

* Correspondence: chm2042@med.cornell.edu
[3]Department of Physiology and Biophysics, Weill Medical College, Cornell University, New York, NY 10065, USA
[4]HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Medical College, Cornell University, New York, NY 10065, USA
Full list of author information is available at the end of the article

We have shown that this problem applies for older patents such as *BRCA1*, but also for patents granted even in the past few years [1], such as a broad patent on *JAK2* (number 7,781,199), where the patent covers:

> Claim5: "An isolated nucleic acid consisting essentially of at least 12 consecutive nucleotides of sequence SEQ ID NO 3 or 4, wherein the isolated nucleic acid comprises the nucleotide t^{261} in SEQ ID NO 3 or $t^{5}0$ in SEQ ID NO 4, wherein the isolated nucleic has the functional properties of a probe or primer."

Even though satisfying all elements of the claim seems to restrict the patent to only fragments with the specified mutations (t^{261} or rs77375493), we found that this still explicitly matches at least 145 other genes [1]. Moreover, given the loose language in other claims of this patent (such as Claim 14) that allow for 'essentially' the same molecule, a broad interpretation is possible that could easily cover thousands of other genes for this one patent.

Yet, while disagreement on the interpretation of these patents continues, both in press and on blogs (see, for example, [7,8]), it is important to remember that this is a legal debate, not a medical one. This abstract legal debate is insulated from the pragmatic medical effects of ambiguous claim language, which have directly led to cease-and-desist letters ordering scientists and clinicians to halt research. This includes the halting of scientists' research focused on *BRCA1/2* in the 1990s (as declared by the plaintiffs in the AMP v. Myriad case [6]). These vague claims are problematic not only because they are not real inventions [6], but also because their scope has a direct, negative impact on clinical research and open access to fundamental genetic information. Even beyond these vagaries, additional legal problems remain for many gene patents. First, cDNA synthesis methods, PCR-based amplification methods and gene sequencing methods are now all very obvious, dating back to the early 1970s and 1980s, meaning that remaining patents' claims may eventually fall on obviousness grounds. Second, many patents lack methodological specificity and clear enablement because they do not provide details of an actual set of molecules; rather, they have 'invented' complementary strands of existing molecules and then 'invented' any potential variation of that molecule. If drug patents were written as vaguely as gene patents, one could imagine ludicrous claims such as 'X molecule with 200 atoms, and any variation or combination of this molecule using 15 or more of its atoms.' If drug patents were written so vaguely, it would be almost impossible to develop new drugs, yet many gene patents and method claims are built on similarly vague grounds.

Finally, we note the specific novelty and impact of our k-mer analysis. Previous work in the field had only estimated the number of 'cross-matches' from short sequences or only looked at longer sequences, whereas our work has empirically detailed the clear non-specificity of both short- and long-sequence patents. Also, our work noted that many gene patents have claims 'for a linear series of nucleotides, not a specific chemical structure', and the Supreme Court explicitly agreed in the final language of their decision [6], stating that a clear problem with gene patents is that the 'claims are not expressed in terms of chemical composition', rather 'they focus on the genetic information'. Overall, the problems highlighted by our data and analysis were mostly ameliorated by the Supreme Court's decision [6]. Indeed, our previous paper's main conclusion [1], wherein we suggested that the Supreme Court should 'limit the patenting of existing nucleotide sequences because of their broad scope and non-specificity in the human genome', rather than being an 'incorrect view of the law', is precisely what happened in the Court, and it is now the law.

Competing interests

CM served as an expert witness in the case *AMP v. Myriad* [6] on behalf of the plaintiffs. CM and JR are founders of the Genome Liberty Corporation, a genetic testing company.

Author details

[1]Rutgers University OIT- High Performance and Research Computing Group, South Orange Avenue, Newark, NJ 07103, USA. [2]American Museum of Natural History, Sackler Institute for Comparative Genomics, Central Park West at 79th Street, New York, NY 10024, USA. [3]Department of Physiology and Biophysics, Weill Medical College, Cornell University, New York, NY 10065, USA. [4]HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Medical College, Cornell University, New York, NY 10065, USA. [5]The Information Society Project, Yale Law School, New Haven, CT 06520, USA.

Published: 27 February 2014

References

1. Rosenfeld JA, Mason CM: **Pervasive sequence patents cover the entire human genome.** *Genome Med* 2013, **5**:27.
2. **CAMBIA's Sequence project in patentlens is now replaced by PatSeq finder in the lens (www.Lens.org).** http://www.patentlens.net/daisy/patentlens/landscapes-tools/sequences.html.
3. 35 USC. § 157(a)(3).
4. Graff GD, Phillips D, Lei Z, Oh S, Nottenburg C, Pardey PG: **Not quite a myriad of gene patents.** *Nat Biotechnol* 2013, **31**:404–410.
5. Jefferson OA, Köllhofer D, Ehrich TH, Jefferson RA: **Transparency tools in gene patenting for informing policy and practice.** *Nat Biotechnol* 2013, **31**:1086–1093.
6. **Supreme court of the United States: syllabus: association for molecular pathology et al. v. Myriad genetics, inc., *et al.*** http://www.supremecourt.gov/opinions/12pdf/12-398_1b7d.pdf.
7. Chin A: **Artful prior art and the quality of DNA patents.** *Alabama Law Review* 2005, **57**:975.
8. McKernan KJ, Spangler J, Helbert Y, Zhang L, Tadigotla V: **DREAMing of a patent-free human genome for clinical sequencing.** *Nat Biotechnol* 2013, **31**:884–887.

doi:10.1186/gm532
Cite this article as: Rosenfeld and Mason: Response to 'pervasive sequence patents cover the entire human genome' - authors' reply. *Genome Medicine* 2014 **6**:15.