GenomeQuest

DO YOUR NGS DIAGNOSTICS VIOLATE INTELLECTUAL PROPERTY?

White Paper. Richard Resnick, CEO, GenomeQuest.

NGS stands to transform diagnostics. But the December 2012 lawsuit filed by Genetic Technologies against LabCorp and 23andMe[1] may be an example of the importance of modern search algorithms to keep pace with the exploding body of patent literature. The right patent databases need to be searched with the right algorithms to ensure that labs have freedom to operate before they launch tests, and to allow the industry to avoid a plague of lawsuits that will impede patient access to cutting edge diagnostics.

Molecular diagnostics that leverage Sanger sequencing generally price anywhere from $1,000 to $5,000 and above, and focus on small regions of the genome, typically gene-sized or smaller. Next-generation sequencing dramatically increases sequencing efficiencies, allowing for the analysis of many tens or hundreds of genes for the same price. This sea change in efficiency is rapidly morphing the diagnostics landscape, particularly for those companies selling lab-developed tests (LDTs).

With this rapid increase in efficiency, larger portions of the human genome are being measured, potentially including areas of existing intellectual property. The industry is widely aware of the Myriad patents on BRCA1 and BRCA2 genes, which they fight strongly to protect and enforce. Jensen and Murray estimated that 4,270 U.S patents claimed at least one human gene, and that nearly 20% of human genes are explicitly claimed as intellectual property in 2005.[2]

Adding genes to a next-generation sequencing gene panel test may violate the intellectual property rights of patent holders. Consider even a random selection of, say, 20 genes; the likelihood that at least one of these 20 genes is patented is over 98%. However, the gene patent landscape is not random; it is skewed towards those genes that have commercial testability or drugability. Many NGS tests currently on the market today as LDTs may be violating the IP of financially backed patent holders.

The problem is compounded when one considers the number of probes that are claimed in these patents. The GenomeQuest GQ_PAT database as of January 2013 shows over 73 million human sequences that are 50 bases or less. While not all probes are directly claimed as inventions, the number of human genes that are amplified by patented probes is surely higher than the 4,382 genes cited in the Jensen paper.

The lawsuit filed by Genetic Technologies against LabCorp and 23andMe highlights these challenges. Genetic Technologies is suing both companies for infringement of their patent on the ACTN3 gene. The patent claims just two 21-base intronic probes that amplify an exon in ACTN3 containing a SNP that predicts athletic performance[3]:

SEQ ID NO 1: CTGTTGCCTG TGGTAAGTGG G
SEQ ID NO 2: TGGTCACAGT ATGCAGGAGG G

---

[1] Zuhn, Donald, and Kevin Noonan. "Court Report." *Patent Docs*. 30 Dec. 2012. Web. 4 Feb. 2013. <http://patentdocs.typepad.com/files/gtg-v-laboratory.pdf>
[2] Jensen, Kyle and Fiona Murray. "Intellectual Property Landscape of the Human Genome." Science 14 Oct. 2005: 239-240. Print.
[3] Yang N, MacArthur, Gulbin, Hahn, Beggs, Easteal, North. "ACTN3 genotype is associated with human elite athletic performance." *Am J Hum Genetics* Sep 2003; 73(3):627-631. Print.

Shockingly, a BLAST search with default parameters[4] that compares the ACTN3 gene to our database of over 200 million patented sequences does not return these hits, even though they exist in the GenomeQuest database. Indeed, in 100,000 alignments, the shortest sequence that the BLAST algorithm found was 156 bases, a hit to a now lapsed patent by Regulome. If Labcorp or 23andMe surveyed the patent landscape using the trusted BLAST algorithm with default parameters to compare genes in their panels to patented DNA, they would have missed this hit.

GenomeQuest has developed a specific algorithm, GenePAST, for searching the body of patented sequences – one that focuses on sequence identity rather than biological homology. Where BLAST can not weed through the hundreds of millions of sequences in the patent literature to find and prioritize high identity hits, GenePAST can. It is therefore a trusted algorithm used by many patent offices worldwide as the standard of DNA search.

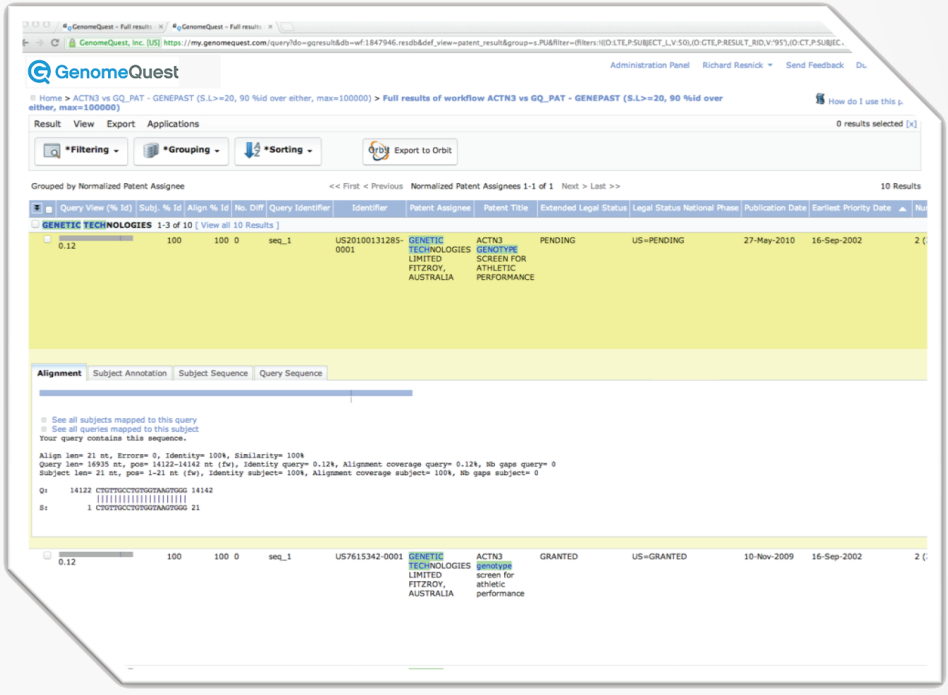

FIGURE 1: *The Genetic Technologies patent found via a GenePAST search of GQ-Pat using ACTN3 as the query.*

We applied the GenePAST algorithm to the exact same search[5]. GenePAST was able to identify exact matches between the genomic ACTN3 gene and these two probes. Had Labcorp or 23andMe used the GenePAST algorithm to determine whether they had freedom to operate, they would have discovered the prior art, and perhaps the lawsuit brought by Genetic Technologies could have been averted.

CONTACT

For more information on the GenePAST algorithm, contact GenomeQuest at:
genepast@genomequest.com

[4] BLAST search against GQ-Pat with word size = 11, e-val cutoff of 10, and the NUC 3.1 scoring matrix employed.
[5] GenePAST search against GQ-Pat with minimum subject length of 21 bases, discarding any alignments with less than 90% identity.