Preliminary Findings from a Population Level Study of DNA Patents

Prepared for SACGHS by
Lori Pressman, Mark Rohrbaugh, and Stephen Finley
February 2009

Preliminary Findings from a Population Level Study of DNA Patents

By Lori Pressman, Mark Rohrbaugh, and Stephen Finley[1]
February, 2009[2]

Purpose of Commissioning a Population Level Study

To complement the case studies, SACGHS commissioned a population level analysis of DNA Patents (patents defined by the bioinformatic algorithm described in Ref 1,), licenses to these patents, and products sold under the licenses. Population-level studies provide important input into the development of public policy as recommendations are typically directed to populations of patents and licenses, rather than applied on a case by case basis.

The population level study was conducted in two phases. The first phase explored whether and how the bioinformatic algorithm could be used to identify NIH Office of Technology Transfer "OTT" patents, and by implication, patents in general, which cover commercially-available clinical diagnostic tests. The second phase compares the licensing practices, policies and commercial outcomes for DNA Patents managed by the National Institutes of Health Office of Technology Transfer (NIH OTT) with those managed by not-for-profit academic institutions (AIs) documented in a prior study.[3] As the AIs and the NIH OTT operate under different policy frameworks (the Bayh-Dole Act[4] for the AIs and the Stevenson Wydler Technology Innovation Act[5] and the Federal Technology Transfer Act of 1986[6] for the NIH OTT), this comparison has the potential to reveal the effects of such policies on patent commercialization outcomes, including the commercial availability of clinical genetic diagnostic tests.

Predictive Value of the Bioinformatic Algorithm: The positive and negative predictive values of the i) bioninformatic algorithm and ii) bioinformatic algorithm enhanced by expert curation were explored. The term "marker" means meeting the criteria of the algorithm, and the term "refined marker" means meeting the criteria of the algorithm and *also* further selected by expert curators[7] as a patent with the potential to cover commercial clinical diagnostic tests or services. Approximately one third of patents found by the algorithm were selected by the expert curators.

---

[1] Lori Pressman is an independent consultant. Mark Rohrbaugh and Stephen Finley are affiliated with the NIH Office of Technology Transfer. Bob Cook-Deegan, Subhashini Chandrasekharan, and Carla Rydholm, all with the Duke University Center for Genome Ethics, contributed to the study, particularly regarding the expert curation patent taxonomy, but not to the report.
[2] This study is ongoing. Additional licensing data, made available to the first author in January 2009, are still undergoing analysis.
[3] Pressman, L., Burgess, R., Cook-Deegan, R.M., McCormack, S.J., Nami-Wolk, I., Soucy, M., and Walters, L. (2006). The licensing of DNA patents by U.S. academic institutions: an empirical survey. *Nat Biotechnol.* **24:** 31-9.
[4] 35 U.S.C § 200-212
[5] Public Law 96-480
[6] Public Law 99-502
[7] Dr. Subhashini Chandresekharan and Dr. Carla Rydholm.

This report was commissioned for SACGHS. The findings are preliminary. Any views expressed do not necessarily represent the views of SACGHS.

The precise numbers in this table are expected to change after the new data are integrated. The overall observations are not expected to change.

Table 1: Positive Predictive Value "PPV" and Negative Predictive Value "NPV" of the marker and the refined marker, for predicting i) how the patents will be licensed, ii) whether the patents will be associated with clinical diagnostic tests regardless of analyte, and iii) whether the patents will be associated with clinical diagnostic tests where a nucleic acid sequence is the analyte.

	marker PPV	marker NPV	refined marker PPV *(See note G)*	refined marker NPV *(See note G)*
Ability to predict which patents are in a license with a Diagnostic Sales or Testing Service Field of Use	12-33% 33/273 to 91/273	60-93% 49/81 to 76/81	23-54% 23/102 to 55/102	79-94% 135/171 to 161/171
Ability to predict which patents are associated with a royalty earning clinical diagnostic test or service	23-33% 22/93 to 31/93	66-87% 20/30 to 26/30	23-37% 11/46 to 17/46	70-76% 33/47 to 36/47
Ability to predict which patents are associated with a Nucleic Acid based clinical diagnostic test or service	8-9% 7/93 to 8/93	97-100% 29/30 to 30/30	13-15% 6/46 to 7/46	97% 46/47

Note G: *These calculations ignore the patents not found by the bioinformatic algorithm*

Both markers have incomplete penetrance. DNA patents associated with clinical diagnostic tests are also associated with other products. There are also patents not found by the bioinformatic algorithm, but in the same patent family [8] as a patent found by the bioinformatic algorithm, which are utilized in commercially available clinical diagnostic tests or services. With one exception, these tests are antibody based, rather than nucleic acid based.

The PPV of both markers, the bioinformatic alone, and the refined marker, is poor, both at the license level (predicting which patents are in licenses where the parties contemplated, at the time the license was being negotiated, that the licensee would make a diagnostic product or perform a diagnostic service), and at the product level (a diagnostic product is on the market, or a diagnostic service is commercially available).

The poor PPV of the markers is not surprising, as the search string clearly picks up varied group of biotechnology patents.[9] Table 2 below shows the Issued U.S. Patents assigned to companies

[8] Patents which derive their support from a shared patent "specification" (the part other than the claims) are said to be in the same patent family. The specification provides the novel and not obvious teaching which entitles the patent holder to their patent.
[9] See Table 1 Reference 1.

This report was commissioned for SACGHS. The findings are preliminary. Any views expressed do not necessarily represent the views of SACGHS.

known to be active in the development of diagnostic tests and platform technologies, and, in the case of Illumina and Helicos, patents invented by their University-based founders, David Walt and Stephen Quake, and the number of such patents also detected by the bioinformatic algorithm. The searches were run December 25, 2008. The Delphion U.S. Patent Collection starts in 1971.

Table 2

Company	Number of Issued U.S. Patents	How many are also found by the DPD Algorithm
Roche Is Patent Owner Last 10 years only	10,849 3,579	1258 777
Applied BioSystems is Patent Owner Last 10 years only	186 114	69 43
Illumina is Patent Owner David Walt, (Tufts Professor whose technology formed the basis of Illumina in 1998) is an Inventor	45 36 (30 owned by Tufts University)	27 10 (8 owned by Tufts University, 2 by Illumina)
Helicos is Patent Owner Stephen Quake, (Caltech Professor, whose technology formed the basis of Helicos in 2004) is an inventor	7 57 (all 57 owned by Caltech)	4 12 (all 12 owned by Caltech)

The improved, yet still poor, PPV of the refined marker may be due to the curators intentionally seeking high sensitivity.

The NPV of the simple bioinformatic approach could be useful when restricted to identifying issued patents that are not associated with clinical diagnostic tests which rely primarily on nucleic acids. However, this approach has several limitations. First, many licenses are executed before all patents in a patent family have issued, and some patents in the family have nucleic acid-based claims, others have protein and antibody based claims, and the order in which the claims issue is unpredictable.

More significantly, the approach is limited also because most of the clinical diagnostic tests associated with NIH OTT-licensed DNA patents are not nucleic acid based, though they are virtually all associated with at least one patent, even if through a “patent family” relationship, which has nucleic acid sequences in the claims. Thus, the phrase “clinical genetic diagnostic test” is potentially misleading, as frequently the analyte is not a nucleic acid sequence, but instead a protein product of gene expression, or an antibody to a protein product of gene expression. The phrase “clinical diagnostic test of genetic origin” may be a more accurate description of clinical diagnostic tests informed to some degree by an understanding of the underlying genetics.

Examples of clinical diagnostic tests of genetic origin, grouped by analyte type, from the NIH OTT Licensing Program, from the case studies, and from recent press releases are listed below:

This report was commissioned for SACGHS. The findings are preliminary. Any views expressed do not necessarily represent the views of SACGHS.

Full Nucleic Acid Sequencing:

BRCA1 and BRCA2 breast/ovarian cancer susceptibility genes, the first case in a family (Myriad Genetics)

Detection of a partial gene sequence via hybridization:

DNA Probe for HER2 gene for predicting response to Herceptin (Invitrogen)

Detection of a protein product of gene expression:

HER2 immunoassay for predicting response to Herceptin
Prezeon® immunoassay for determining PTEN status (Myriad Genetics)
Cerebrospinal fluid tests for Apolipoprotein E or Phosphorylated Tau protein used in Alzheimer's testing. (Athena Diagnostics)

Examination of the function of a protein product of gene expression:

Hexosaminidase assay for Tay-Sachs

Detection of antibodies raised in response to infectious agents:

HIV-1 blood screening tests

The type of test is also to some degree a function of its time in history. If the human immunodeficiency virus (HIV) had been isolated before the tools of genetic engineering and antibody production were readily available, it is possible that a blood screening test could have been developed starting from collecting pooled sera of infected individuals. However, once the virus was isolated, it was easier to start from knowledge of its antigenic surface proteins, as determined precisely by the viral genome, and use that information to develop an antibody based blood screening assay.

Are antibody- and protein-based tests of equal concern to policymakers as nucleic acid-based tests? If not, is this because antibodies are perceived as less biologically fundamental than nucleic acid sequences, even if, from a pure patent point of view, a well written antibody patent could, in theory, obstruct an antibody-based test for a gene expression product? Or, if the antibody has been engineered in some way, perhaps to make a binding event easier to detect, is such a patent of less concern to policymakers because it appears to capture some more easily recognized technical contribution of the inventors?

If such tests are of equal concern, then the challenge in formulating an objective marker to identify patents claiming amino acid sequences with the potential to obstruct access to protein-based clinical diagnostic tests of genetic origin is greater than that for patents with nucleic acids in the claims, as amino acid sequences have abundant medical and commercial applications apart from clinical diagnostic tests.

The low PPV of patent-level markers, the occurrence of DNA or nucleic acid-based patents in many businesses, including relatively new companies such as Illumina and Helicos, the occurrence of protein and other biomarker-based tests in clinical diagnostic tests of genetic

This report was commissioned for SACGHS. The findings are preliminary. Any views expressed do not necessarily represent the views of SACGHS.

origin, the unpredictable order in which patent claims in a patent family issue—including those for nucleic acid sequences and those for protein products of gene expression, or for antibodies to the protein products—all suggest that public policy recommendations regarding intellectual property rights should not focus on the patents themselves. One alternative is to consider recommendations for licensing terms, discussed next.

Comparison of Licensing Practices under Two Policy Frameworks

Policy and Practice Differences

In brief, NIH OTT favors nonexclusive licensing,[10] requires a public notice period before granting licenses with exclusivity, and does not grant all-field-of-use exclusive licenses. The NIH OTT maintains more never-licensed patents as a percentage of its total (See table 3 below).

Table 3: Comparison of maintenance status for AI and NIH OTT Never-Licensed "DNA Patents"

	Did Not Pay 3.5 Year Patent Maintenance Fee	Did Not Pay 7.5 Year Patent Maintenance Fee	Did Not Pay 11.5 Year Patent Maintenance Fee
AI Never-Licensed 771 Unique	57 (about 7%)	36 (about 4.7 %)	8 (about 1%)
NIH Never-Licensed 312 Unique	11 (about 3.5%)	8 (about 2.5%)	1 (about .3%)

Based on the results of two studies,[11] [12] NIH OTT inventors appear to play a smaller role in invention marketing than AI inventors. AIs have more discretion in the scope of license grants to their patents, are more able to participate in start-up formation, their inventors appear to be more involved in the technology transfer process relative to NIH OTT inventors, and they maintain fewer never-licensed patents.

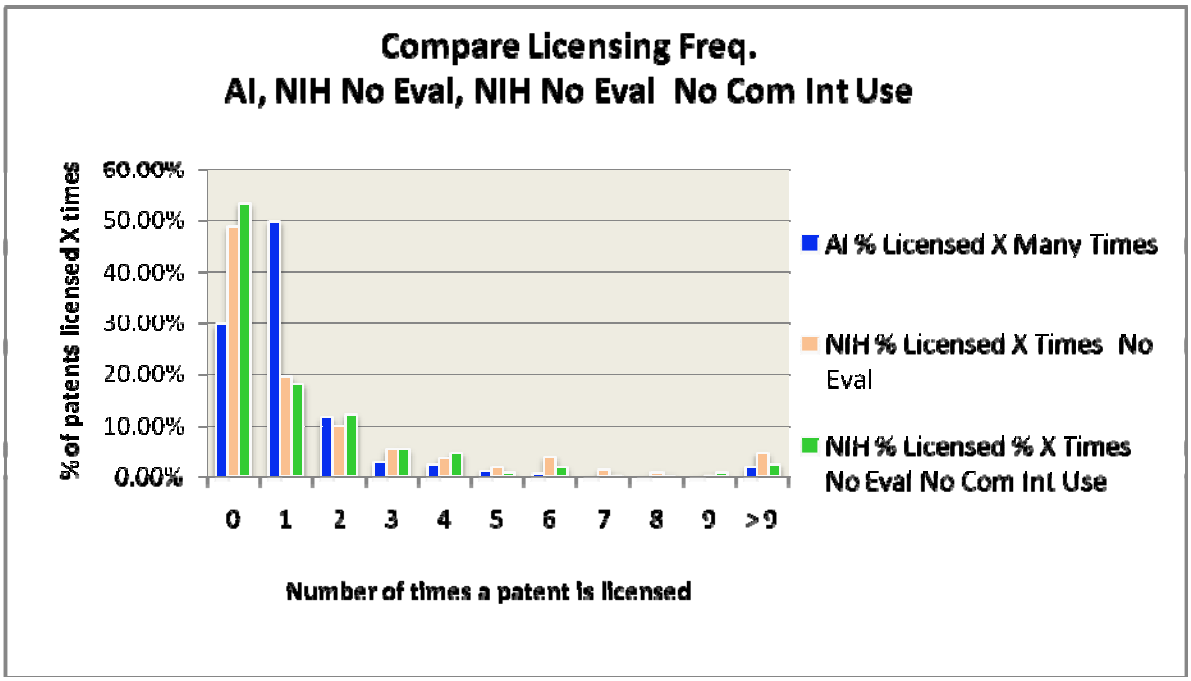
Percentage of Patents Licensed Under the Two Frameworks

More AI-managed DNA Patents are licensed overall relative to NIH OTT-managed DNA Patents. This effect remains when controlled for absolute age of the patents, age of the patent at the time the data were gathered, and is not obviously explained by patent classification codes. *The new data that will be integrated may affect these results, but probably not significantly*.

[10] http://www.ott.nih.gov/FAGs/#6 Accessed February 5, 2009
[11] Jansen, Dillon, "Where do the Leads for Licenses Come From" source Data from Six Institutions". Journal of the Association of University Technology Managers, vol. XI. p27
[12] Ramakrishnan, Chen, Balakrishnan. Effective strategies for marketing biomedical inventions: Lessons learned from NIH license leads. Journal of Medical Marketing 5(4):342-352.

This report was commissioned for SACGHS. The findings are preliminary. Any views expressed do not necessarily represent the views of SACGHS.

Figure 1. Licensing Frequency Data for the AI and NIH sets and accompanying table used to generate the graph. The NIH OTT license agreements that include neither commercial evaluation agreements nor commercial internal use agreements are the most directly comparable to the AI license agreements. Thus, the first and third bars in each group of histograms are the most directly comparable to each other. More than 50% of NIH OTT-managed DNA Patents have never been licensed, whereas only 30% of AI-managed DNA Patents have never been licensed.



This report was commissioned for SACGHS. The findings are preliminary. Any views expressed do not necessarily represent the views of SACGHS.

Data Table for Figure 1:

Number of times Licensed	AI Count	AI %	NIH Count no Eval	NIH % No Eval	NIHCount No Eval No Com Int	NIH % No Eval No Com Int
0	771	29.65%	286	48.89%	312	53.33%
1	1297	49.88%	115	19.66%	106	18.12%
2	302	11.62%	58	9.91%	71	12.14%
3	74	2.85%	32	5.47%	32	5.47%
4	58	2.23%	20	3.42%	27	4.62%
5	25	0.96%	10	1.71%	5	0.85%
6	16	0.62%	22	3.76%	10	1.71%
7	3	0.12%	7	1.20%	3	0.51%
8	3	0.12%	5	0.85%	2	0.34%
9	3	0.12%	3	0.51%	4	0.68%
> 9	48	1.85%	27	4.62%	13	2.22%

A preliminary observation is that the percent of patents licensed more than nine times, sometimes referred to as "licensed broadly" in this report, is similar under the two frameworks. Of course, not all licenses, even nonexclusive ones, are a sign that product is on the market.

Exclusivity Practice: Using the preliminary data, the exclusivity practice is more similar than generally believed, though clearly more NIH OTT licenses are nonexclusive. Generally, licenses with some degree of exclusivity also include considerably more diligence requirements than licenses granted on a non-exclusive basis. "Diligence" refers to the contractual terms that require the licensee to invest in developing and commercializing the invention. License contracts can be drafted so that meeting diligence requirements expands the scope of rights granted therein, and/or so that failure to meet the diligence requirements results in loss of rights under the contract.

Hypothetical Examples of Diligence Terms
Technical and Regulatory Requirements: i) achieve a certain sensitivity or specificity for a clinical diagnostic test, ii) obtain FDA clearance, iii) develop a multivalent monoclonal antibody to the protein product of gene expression
Requirements to Raise or Spend Money, or to support research at the patent owner's institution: i) raise no less than $5M devoted to development of the licensed technology within one year of signing the license, ii) spend [at the licensee] no less than $1M/year until receiving FDA clearance, iii) fund no less than $100,000 per year at the patent owner's institution for a minimum of 3 years after signing the license.
Requirements to demonstrate commercial traction: i) sell a certain dollar volume by a certain date, ii) sell a certain unit volume by a certain date, iii) show increasing sales over a period of years

This report was commissioned for SACGHS. The findings are preliminary. Any views expressed do not necessarily represent the views of SACGHS.

Hypothetical Examples of Consequences of failing to meet, or meeting Diligence Terms
Negative Consequences: i) failure to meet milestone xyz shall result in termination of the license under paragraph abc, ii) failure to meet milestone xyz shall result in conversion of the license to nonexclusive, iii) failure to meet milestone xyz shall enable the patent holder to grant one additional license per year to the technology in the licensee's field of use.
Positive Consequences: i) upon meeting milestone xyz, the patent holder will grant no additional licenses in licensee's field of use, ii) upon meeting milestone xyz, the patent holder will grant at most one additional license per year in licensee's field of use, and no more than 2 additional licenses total. iii), upon meeting milestone xyz, the patent holder will expand licensee's Field of Use to include jkl.

Evidence for Incentives Created by Patents: Theoretical Basis for Timeline Analysis

Timelines[13] are useful tools because they have the potential to shed light on cause and effect,[14] and because timing is itself a potential metric of availability. Product Commercialization timelines provide data on temporality, with the caveat that "home brew" or Laboratory-Developed Tests or services not reported to a patent holder are necessarily absent from this analysis. Nonetheless, it is not correct to automatically assume home brew tests are equivalent to commercial tests.

Products that appear very quickly after invention publication suggest that little commercial development is required to bring the product to market. Obviously, there was by definition another incentive. Products that appear years after publication of the invention suggest, though clearly in isolation do not prove, that some additional development was required. In the case of clinical diagnostic tests of genetic origin, additional development could include improving and/or documenting clinical utility or adapting the technology for reproducible use in a commercial manner. Both of these might be needed to obtain FDA approval for a test kit or FDA clearance for marketing claims of a kit or service.

Products that appear on the market soon after the patent license was signed again suggest that the patent itself was not a significant incentive for the company producing the product, particularly when the license is nonexclusive with little diligence required from the licensee. Products which appear years after the patent license was signed are consistent with the patent serving as an incentive for the company to invest in producing a product. The argument that such a license is a development incentive is stronger when the license is at least partially exclusive and supported by the presence of diligence in the license agreements, as there is a clearer and stronger contractual obligation on the part of the licensee and cost to that licensee in advance of the anticipated benefits. The licensee will likely only accept these contractual obligations costs when the license protects them from the risks that others might enter the market later without similar costs of developing the technology.

It is not yet clear how the new data could affect this analysis.

[13] Relative timing of the invention publication, patent filing, license execution, product availability.
[14] Hill, A.B. (1965). The Environment and Disease: Association or Causation? *Proceedings of the Royal Society of Medicine* 58:295-300. As a convenience, they are listed here: Strength, Consistency, Specificity, Temporality, Gradient, Plausibility, Coherence, Experiment, and Analogy.

This report was commissioned for SACGHS. The findings are preliminary. Any views expressed do not necessarily represent the views of SACGHS.

Timeline Data: The available timeline data are presented in tabular form below. The limitations of this data for the AI set include: i) a small number of data points, ii) values distribution, and thus the standard deviations are quite large, and iii) survey respondents submitted license effective dates and product introduction dates by fiscal year, rather than providing the actual day and year the licenses were executed. There are apparently only a handful of NIH OTT licenses with exclusivity so it is not possible to compare exclusivity types within the NIH OTT set. The AIs did not describe the products in their licenses—only that they existed.

The only set of licenses with an average multi-year delay between signing the license and the appearance of product are the AI Licenses with some exclusivity. One of the NIH OTT licenses with exclusivity has a significant time delay, the others do not. The other NIH OTT Licenses associated with products, all of which are essentially nonexclusive, and the AI nonexclusive licenses tend to be signed close to the time products are introduced, and not, on average, in advance of their introduction.

In isolation, these results do not prove that exclusive licensing creates incentives. However, combined with the documented contractual diligence in the AI licenses with exclusivity, the timeline data provide evidence that such licenses create development incentives.

There is also some evidence that NIH OTT Diagnostic Products from commercial providers are on the market later than reagents, possibly showing the effect of a certification or regulatory delay.

Finally, there is some evidence that DNA Patents licensed under an AI practice and policy framework may be associated with commercial products, as indicated by the presence of an earned royalty report, sooner than DNA Patents licensed under the NIH OTT policy framework.

This report was commissioned for SACGHS. The findings are preliminary. Any views expressed do not necessarily represent the views of SACGHS.

Academic Institution Summary Product Timeline Data

	Number of Data Points	AI Delta 1st pat Priority Date[15] to Product availability years	Standard Deviation	AI Delta License Effective Date to Product Availability years	Standard Deviation
Nonexclusive Mean	11	5.68	4.84	0.00	2.94
Nonexclusive Median	11	4.66		0.00	
Some Exclusivity Mean	20	4.47	3.26	3.10	.90
Some Exclusivity Median	20	4.14		2.75	

NIH OTT Summary Product Timeline Data

NIH Delta 1st pat Priority Date to Product availability	Number of Data Points	NIH Delta 1st pat priority Date to Product Availability years	Standard Deviation	NIH Delta License Effective Date to Product Availability	Standard Deviation
Dx Product Sales Mean	13	8.45	4.15	0.00	3.63
Dx Product Sales Median	13	9.5		0.00	
Reagent Product Sales Mean	42	7.74	4.7	1.13	2.44
Reagent Product Sales Median	42	6.41		.31	

These results suggest that automatic and default nonexclusivity could have a cost, especially given the apparent impossibility of *a priori* identifying groups of patents "needing" to be licensed nonexclusively because they are certain or likely to be associated with clinical diagnostic tests of genetic origin.

Given the apparent impossibility of identifying patents "needing" to be licensed nonexclusively, and the potential unintended removal of incentives where they could be beneficial, nuanced exclusivity with prudent diligence is an attractive policy option.

[15] Reasonable proxy for invention publication in biotech sector.

This report was commissioned for SACGHS. The findings are preliminary. Any views expressed do not necessarily represent the views of SACGHS.