

Filed on behalf of: Guardant Health, Inc.
By: Michael T. Rosato (mrosato@wsgr.com)
Steven W. Parmelee (sparmelee@wsgr.com)
Sonja R. Gerrard (sgerrard@wsgr.com)
WILSON SONSINI GOODRICH & ROSATI
701 Fifth Avenue, Suite 5100
Seattle, WA 98104-7036

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GUARDANT HEALTH, INC.,
Petitioner,

v.

FOUNDATION MEDICINE, INC.,
Patent Owner.

Case IPR2017-01448
Patent No. 9,340,830

**PETITION FOR INTER PARTES REVIEW OF
U.S. PATENT NO. 9,340,830**

TABLE OF CONTENTS

	<u>Page</u>
I. Introduction.....	1
A. Brief Overview of the '830 Patent	1
B. Brief Overview of the Prosecution History	4
C. Brief Overview of the Scope and Content of the Prior Art.....	5
i. Targeted Enrichment Methods and Bait Design.....	5
ii. Cited Prior Art References.....	8
D. Brief Overview of the Level of Skill in the Art	10
II. Grounds for Standing.....	10
III. Mandatory Notices under 37 C.F.R. § 42.8.....	11
IV. Statement of the Precise Relief Requested For Each Claim Challenged	12
V. Claim Construction.....	13
VI. Detailed Explanation Of Grounds For Unpatentability.....	15
A. [Ground 1] Claims 1 and 15 are Obvious under 35 U.S.C. § 103 over Levin and Shah.....	15
i. Independent Claim 1	19
ii. Dependent Claim 15	45
B. [Ground 2] Claims 17 and 18 are Obvious under 35 U.S.C. § 103 over Levin, Shah, and Stephan.....	46
i. Dependent Claim 17	46
ii. Dependent Claim 18	48
C. [Ground 3] Claims 19, 20, 32, 44, and 45 are obvious under 35 U.S.C. § 103 over Levin, Shah, and Malhis.....	50

i.	Dependent Claim 19	53
ii.	Dependent Claim 20	56
iii.	Dependent Claim 32	58
iv.	Dependent Claim 44	58
v.	Dependent Claim 45	60
D.	[Ground 4] Claims 46, 47, and 57 are obvious under 35 U.S.C. § 103 over Levin, Shah, Malhis, and Li	60
i.	Dependent Claim 46	61
ii.	Dependent Claim 47	64
iii.	Dependent Claim 57	65
VII.	Conclusion	66
VIII.	Certificate of Compliance.....	67
IX.	Payment of Fees under 37 C.F.R. §§ 42.15(a) and 42.103	68
X.	Appendix – List of Exhibits.....	69

I. INTRODUCTION

Guardant Health, Inc., (“Petitioner”) hereby requests *inter partes* review of U.S. Patent No. 9,340,830 to Doron Lipson et al. (hereinafter “the ’830 patent,” EX1001) that issued on May 17, 2016, and is currently assigned to Foundation Medicine, Inc. (“Patent Owner”). This Petition demonstrates by a preponderance of the evidence, that it is more likely than not that claims 1, 15, 17-20, 32, 44-47, and 57 of the ’830 patent are unpatentable for failing to distinguish over prior art.

Thus, the challenged claims of the ’830 patent should be found unpatentable and canceled.

A. Brief Overview of the ’830 Patent

In general, the ’830 patent claims are directed to a method for analyzing a sample of tumor cells for a somatic mutation in cellular DNA by using targeted “enrichment” of sequences, and then using basic “next generation sequencing” techniques to acquire sequence data for a mutation. EX1001, claim 1, *see also* 1:46-51, 30:40-45, Fig. 1C; EX1002 ¶¶20, 21.

While the claims of the ’830 patent are verbose, they recite nothing beyond prior art known processes. Independent claim 1, for example, can be broken into two component parts, (1) targeted enrichment followed by next generation sequencing; and (2) choice of bait sets that includes at least two bait sets from a list of five options. EX1001, claim 1; EX1002 ¶22.

Virtually every aspect of the claimed subject matter is identified in the '830 patent as being known in the art. The '830 patent acknowledges that targeted enrichment using bait hybridization and next generation sequencing were well-known “off the shelf” technologies as of the December 2010 filing date. EX1001, 61:25-43 (“...Protocols for enrichment are publicly available, e.g., *SureSelect Target Enrichment System*....”), 30:30-45 (“Next generation sequencing methods are known in the art...”). Regarding bait design and production, the '830 patent references prior art and commercial sources. EX1001, 56:26-39, 61:30-43 (“Baits can be produced by methods described in US 2010/0029498 and Gnirke, A et al. (2009) *Nat. Biotechnol.* 27(2):182-189, incorporated herein by reference....”). Indeed, targeted enrichment techniques using bait sets were well-characterized in the art at the time and predictable in terms of performance and design parameters. *See also* EX1002 ¶¶23; *see, e.g.*, “Admitted Prior Art” below.

The '830 patent also acknowledges that numerous “off the shelf” software programs were known for alignment of reads and assigning nucleotide values to preselected nucleotide positions. EX1001, 103:22-104:6 (listing algorithms/programs available for alignment of short reads derived from next generation sequencing), 100:41-47 (“For example, identifying genetic variations such as single nucleotide polymorphism and structural variants in a sample (e.g., a tumor sample) can be accomplished by aligning NGS reads to a reference sequence

....”), 119:29-31, 120:57-62. Certain of the ’830 claims, as addressed in further detail below, recite rote details associated with the use of such alignment programs in basic and conventional sequence analysis. *See also* EX1002 ¶24.

As to the specific choice of bait sets, claim 1 requires at least two bait sets from among an identified list of five options—all of which are found in the prior art. The first (i) and second (ii) bait set options each is a set that targets a somatic mutation that appears in a particular frequency of the cells of the tumor sample, 5% or less and 10% or higher, respectively. The prior art at the time disclosed how to specifically measure somatic mutation frequency in a tumor sample in conjunction with enrichment and sequencing. This is illustrated, for example, in Shah *et al.* EX1005; *see also* EX1002 ¶25.

Moreover, pre-selecting bait sets based on desired depth of sequence coverage was simply a routine optimization of prior art targeted enrichment and next generation sequencing technologies in accordance with instructions readily found in the prior art at the time, including prior art references identified in the ’830 patent. The prior art expressly instructed designing and optimizing bait sets based specifically on target abundance. This is illustrated by “Levin” (EX1010), as well as other prior art references at the time. *See also* EX1002 ¶25.

The third (iii) bait set provides for selection of one of three subgenomic intervals comprising germline single nucleotide polymorphisms (“SNPs”). The

fourth (iv) and fifth (v) bait sets target mutations comprising structural breakpoints in introns and one-copy deletions of exons, respectively. The mutations targeted by bait sets (iv) and (v) are commonly found in tumor cells and may be either germline or somatic in origin. *See also* EX1002 ¶25.

Dependent claims 15, 17-20, 32, 44-47, and 57 simply add minor limitations reciting known reporting methods for results, read alignment methods, and nucleotide calling methods. None of these limitations patentably distinguish over the prior art. *See also* EX1002 ¶26.

As described in further detail herein, a person of ordinary skill in the art at the time would not have considered any of claims 1, 15, 17-20, 32, 44-47, and 57 new or nonobvious. *See also* EX1002 ¶27.

B. Brief Overview of the Prosecution History

Application No. 13/339,986 was filed on December 29, 2011 and issued on May 17, 2016 as U.S. Patent No. 9,340,830. EX1001, front page.

The claims of the 13/339,986 application were rejected during the course of prosecution as, (1) obvious over Drmanac (US 2010/0105052), (2) obvious over Drmanac (US 2010/0105052) in view of Gnirke et al. (US 2010/0029498 [EX1024]) and Polansky (US 7,381,526), and (3) as directed to an abstract idea under § 101. EX1004 at 923-931, 846-857, 494-501.

Levin (EX1010) was cited in one of seven information disclosure statements (IDS) filed during prosecution, but was not discussed during prosecution and the corresponding data files were not submitted. Shah (EX1005), Malhis (EX1042), Li (EX1043) and Stephan (EX1044) were not submitted nor considered during ex parte prosecution. The Examiner did not have the benefit of Dr. Quackenbush’s testimony or a description of relevant aspects of the state of the art at the time, including bait design parameters and predictable optimization aspects.

C. Brief Overview of the Scope and Content of the Prior Art

As explained in detail in the corresponding Declaration of Dr. John Quackenbush (EX1002) and addressed in further detail below, the involved claims would not have been considered new or non-obvious to a person of ordinary skill in the art (“POSA”) at the relevant time.

i. Targeted Enrichment Methods and Bait Design

Targeted enrichment methods, including those relying on solution-phase hybridization, were well-known at the time.

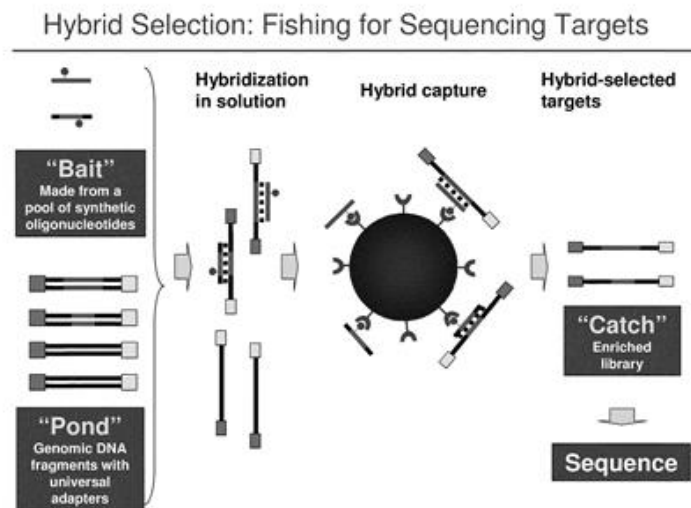


Figure 1

ADMITTED PRIOR ART

Indeed, the '830 patent identifies commercially available targeted enrichment assays using bait hybridization. EX1001, 61:25-43, 98:9-22, 206:55-59. Among the admitted prior art, Figure 1 (shown here) of Gnirke (EX1024) illustrates solution-phase hybridization, where a library is contacted with bait sets to produce a “catch” or enriched library, which is then subjected to next generation sequencing (“NGS”). A similar targeted enrichment assay using bait hybridization is also described by Levin. EX1010. Levin describes use of “bait sets” or nucleic acids for solution-phase capture of target subgenomic intervals. EX1010 at 2, 6; *see also* EX1002 ¶¶29-32.

At the relevant time, targeted enrichment techniques using bait sets were well-characterized in the art and predictable in terms of performance and design parameters. Solution-phase bait sets were commercially available from Agilent. EX1031. These bait sets, and those used by Levin (EX1010) and others for targeted enrichment, could be predictably produced according to known design parameters. EX1024 ¶216; *see also* EX1016 at 5; EX1002 ¶¶33-35.

Additionally, it was known that certain bait parameters influenced hybridization and sequence coverage. As explained by Dr. Quackenbush, such recognized parameters included, for example, amount or concentration of the bait, overlap between baits on a given target interval (*i.e.*, tiling), and the physical properties of the bait (*e.g.*, length, sequence, GC-content, type of nucleic acid,

secondary structure). As Dr. Quackenbush explains, a variety of publicly available databases and resources were available for designing baits that target the human genome. EX1010 at 6; EX1016 at 2; *see also* EX1002 ¶¶36-41.

The '830 patent identifies art-known NGS techniques as being included as a component of the claimed sequencing. EX1001, 30:30-45, 96:52-67, 97:1-28, Fig. 1C. Sequencing using the Illumina platform is described throughout the '830 patent and is also disclosed in Levin and Shah, among other references. *See also* EX1002 ¶42.

The '830 patent identifies art-known alignment and nucleotide calling methods. *See, e.g.*, EX1001, 103:22-104:6 (listing art-known alignment programs), 119:29-43 (listing art-known SNP calling programs), 120:57-121:3 (listing art-known indel calling programs). Moreover, the '830 patent states that prior expectation values (*i.e.*, likelihoods) for mutations in various cancer types may be used in mutation calling algorithms and that this information is publicly available. EX1001, 119:31-38. EX1002 ¶¶43, 44.

Different types of somatic mutations were known in the art. Prior to the filing of the '830 patent, a POSA would have appreciated that tens of thousands of mutations associated with cancer were catalogued in publicly accessible databases. EX1022 at D945. COSMIC was one such database and is cited multiple times

within the '830 patent. EX1001, 101:9-11, 119:31-36, 120:11-17; *see also* EX1002 ¶¶45-56.

ii. Cited Prior Art References

Levin, et al. was published in the journal *Genome Biology* and included a manuscript (EX1010) along with five additional data files (EX1011; EX1012; EX1013; EX1014; EX1015). EX1011-1015 represent true and correct copies of the five additional data files obtained from Genome Biology as they were made available with the online version of the manuscript. *See* EX1010 at 7 (specifically referencing the “Additional data files”); EX1039 ¶¶6-10. The manuscript and additional data files are referred to as “Levin.” As recited on the front page, Levin published October 16, 2009. EX1010 at 1. Levin qualifies as a prior art printed publication under 35 U.S.C. § 102(a) and (b).¹

Shah, et al. was published in the journal *Nature* and included a manuscript (EX1005) along with four supplementary information files (EX1006; EX1007; EX1008; EX1009). *See* EX1010 at 812 (“Supplementary Information is linked to the online version of the paper at www.nature.com/nature.”). EX1006-1009 represent true and correct copies of the supplementary information files obtained

¹ Levin is acknowledged as prior art in the specification of the '830 patent. *See* EX1001, 206:56-59.

from *Nature*. EX1039 ¶¶2-5. Shah published October 8, 2009 and qualifies as a prior art printed publication under 35 U.S.C. § 102(a) and (b).

U.S Patent Application Publication No. US 2008/0131887 A1 to Dietrich Stephan and Melissa Floren Filippone, (“Stephan,” EX1044) was published June 5, 2008 and qualifies as a prior art printed publication under 35 U.S.C. § 102(a) and (b).

Malhis and Jones (“Malhis,” EX1042) was published online in the journal *Bioinformatics* and discloses an algorithm called Slider II. Malhis published February 26, 2010 and qualifies as a prior art printed publication under 35 U.S.C. § 102(a).² The Slider II algorithm was available for download on the Canada’s Michael Smith Genome Sciences Centre website (<http://www.bcgsc.ca/platform/bioinfo/software/SliderII>) on June 9, 2009. Along with the download, the website included documentation on capabilities, testing, and running of Slider II (“Slider II,” EX1046). Further confirming the public availability of Slider II, a webpage listing NGS alignment programs (<http://lh3lh3.users.sourceforge.net/NGSalign.shtml>) published on November 19, 2009, includes entries for both Slider and Slider II. EX1047. Slider II (EX1046) is prior art under 35 U.S.C. § 102(a) and (b).

² Malhis is acknowledged as prior art in the specification of the ’830 patent. See EX1001, 103:60-61.

Heng Li and Richard Durbin (“Li,” EX1043) was published in the journal *Bioinformatics* on May 18, 2009 and qualifies as a prior art printed publication under 35 U.S.C. § 102(a) and (b).³

D. Brief Overview of the Level of Skill in the Art

Petitioner’s technical expert, Dr. Quackenbush explains, a POSA prior to December 30, 2010 would have an understanding of cancer biology as it relates to genetic alterations associated with tumor growth, metastasis, and treatment. Dr. Quackenbush further explains that a POSA would have been aware of the various NGS methods, targeted enrichment methods, bioinformatics methods for mapping short sequence reads onto genomes, and methods for identifying genetic variants in a sample—*e.g.*, as described in Section I.C(i). *See also* EX1002 ¶¶63-68; *see generally id.* ¶¶1-10; *see also* EX1003.

II. GROUNDS FOR STANDING

Petitioner certifies that, under 37 C.F.R. § 42.104(a), the ’830 patent is available for inter partes review, and Petitioner is not barred or estopped from requesting inter partes review of the ’830 patent on the grounds identified.

³ Li is acknowledged as prior art in the specification of the ’830 patent. *See* EX1001, 103:28-29.

III. MANDATORY NOTICES UNDER 37 C.F.R. § 42.8

Real Party-in-Interest (37 C.F.R. § 42.8(b)(1)): Guardant Health, Inc. is the real party-in-interest.

Related Matters (37 C.F.R. § 42.8(b)(2)): Patent Owner has asserted the '830 patent against Petitioner in United States District Court for the Eastern District of Texas, Case No. 2-16-cv-00523-JRG. *See* EX1038. Due to the verbosity of the challenged claims and word limit constraints, Petitioner has filed three petitions (including the present petition) covering '830 patent claims asserted in district court. In particular, Petitioner has filed two other petitions for inter partes review: IPR2017-01170 challenging the patentability of claims 65, 66, 72, 73, 75-77, and 80-85 of the '830 patent, and IPR2017-01447 challenging the patentability of claims 1, 2, 8, 9, 11-14, 16, 48-53, and 62-64 of the '830 patent. Petitioner notes that independent claim 1 is challenged IPR2017-01447, and further addressed here at least to the extent challenged claims 15, 17-20, 32, 44, 46, 47 and 57 depend therefrom and include the limitations of claim 1 by virtue of said dependency.

Lead and Back-Up Counsel (37 C.F.R. § 42.8(b)(3))

Lead Counsel: Michael T. Rosato (Reg. No. 52,182)

Back-Up Counsel: Steven W. Parmelee (Reg. No. 31,990); Sonja R. Gerrard (Reg. No. 72,802)

Service Information – 37 C.F.R. § 42.8(b)(4). Petitioners hereby consent to electronic service.

Email: mrosato@wsgr.com; sparmelee@wsgr.com; sgerrard@wsgr.com

Post: WILSON SONSINI GOODRICH & ROSATI, 701 Fifth Avenue, Suite 5100,
Seattle, WA 98104-7036

Tel.: 206-883-2529

Fax: 206-883-2699

IV. STATEMENT OF THE PRECISE RELIEF REQUESTED FOR EACH CLAIM CHALLENGED

Petitioner requests review of claims 1, 15, 17-20, 32, 44-47, and 57 of the '830 patent under 35 U.S.C. § 311 and AIA § 6. The grounds for relief are as follows:

Ground	Claims	Description
1	1 and 15	Obvious under 35 U.S.C. § 103 over Levin and Shah
2	17 and 18	Obvious under 35 U.S.C. § 103 over Levin, Shah, and Stephan
3	19, 20, 32, 44, and 45	Obvious under 35 U.S.C. § 103 over Levin, Shah, and Malhis
4	46, 47, and 57	Obvious under 35 U.S.C. § 103 over Levin, Shah, Malhis, and Li

V. CLAIM CONSTRUCTION

A claim subject to inter partes review receives the broadest reasonable construction in light of the specification of the patent in which it appears. *See* 37 C.F.R. § 42.100(b); *Cuozzo Speed Technologies, LLC v. Lee*, 136 S. Ct. 2131, 2135 (2016). For the purposes of this review, claim terms are to be given their broadest reasonable interpretation, consistent with how they would be understood by a POSA. A few terms that warrant discussion are identified and discussed below.

“**tumor sample**”: The ’830 patent states that “sample” refers to “a collection of similar cells obtained from a tissue, or circulating cells, of a subject or patient.” EX1001, 31:16-19; *see also id.*, 20:59-60, 31:33-34; 31:37-38. A POSA would have appreciated that the sample types described in the ’830 patent may include premalignant, malignant, and normal cells. *See also* EX1002 ¶71.

“**library**”: The ’830 patent defines “library” as comprising “a collection of members” obtained from one or more subjects. EX1001, 30:6-14. The ’830 patent defines “member” as “a nucleic acid molecule, e.g., a DNA, RNA, or a combination thereof, that is the member of a library.” EX1001, 30:20-23; *see also* EX1002 ¶72.

“**library catch**”: The ’830 patent defines “library catch” as “refers to a subset of a library, e.g., a subset enriched for preselected subgenomic intervals,

e.g., product captured by hybridization with preselected baits.” EX1001, 30:17-19; *see also* EX1002 ¶73.

“**bait**”: The ’830 patent defines “bait” as a “type of hybrid capture reagent.” EX1001, 28:47-50. “A bait can be a nucleic acid molecule, e.g., a DNA or RNA molecule, which can hybridize to (e.g, be complementary to), and thereby allow capture of a target nucleic acid.” *Id.*; *see also* EX1002 ¶74.

“**next generation sequencing**”: The ’830 patent identifies art-known NGS techniques as being encompassed by the claims. *See, e.g.*, EX1001, 30:30-45; EX1002 ¶75.

“**subgenomic interval**”: The ’830 patent defines “subgenomic interval” as “a portion of genomic sequence.” EX1001, 32:33-34; EX1002 ¶76.

“**efficiency for selection**”: Claim 1 recites “wherein each bait set of said plurality has a unique preselected efficiency for selection for its target as compared with the other bait sets in the plurality.” The ’830 patent defines “efficiency of selection” as “the level or depth of sequence coverage as it is adjusted according to a target subgenomic interval(s).” EX1001, 11:20-35. According to the specification, “the level of sequencing depth...refers to the level of coverage of reads (e.g., unique reads), after detection and removal of duplicate reads.” EX1001, 12:43-46. Thus, bait sets exhibiting different levels of coverage following

enrichment are bait sets having different (e.g., unique) efficiency of selection per the '830 patent. *See also* EX1002 ¶77.

In a related district court litigation patent owner has taken the position that at least two baits sets must have distinct (from each other) efficiencies for selecting their targets (“unique”) and that “preselected,” in the context of “efficiency for selection,” means “chosen in advance to account for the desired level of coverage.” EX1002 ¶78.

“**nucleotide value**”: According to the '830 patent, “nucleotide value” “represents the identity of the nucleotide(s) occupying or assigned to a preselected nucleotide position.” EX1001, 30:46-61. Nucleotide values include “missing,” “additional,” or “present” (*i.e.*, A; T; C; or G). *Id.*; EX1001, 119:5-121:3 (discussing use of conventional methods for calling or assigning nucleotide value); *see also* EX1002 ¶79.

VI. DETAILED EXPLANATION OF GROUNDS FOR UNPATENTABILITY

A. [Ground 1] Claims 1 and 15 are Obvious under 35 U.S.C. § 103 over Levin and Shah

As explained in detail below, each and every feature of claims 1 and 15 is taught or suggested in Levin and Shah, and would have been obvious over the prior art. EX1002 ¶82.

Levin discloses a method for analysis of somatic mutations associated with cancer. EX1010 at 1. Levin discloses acquiring a library of nucleic acids from a

tumor sample. EX1010 at 2. Levin discloses contacting the library with at least two bait sets to enrich for target subgenomic intervals. EX1010 at 6. Levin discloses NGS of the enriched library. EX1010 at 2. Levin further discloses analysis of the resultant reads and the identification of somatic mutations and single nucleotide polymorphisms (“SNPs”) in the tumor sample. *Id.* Levin discloses designing and using a plurality of different bait sets targeting genes known to be somatically mutated in cancer. EX1010 at 2, 6; *see also* EX1002 ¶¶83-85.

Levin acknowledges that different mutations will present at different levels of abundance in a sample, a consideration Levin expressly instructs should be factored into bait design in advance to account for a desired depth of sequence coverage. EX1010 at 6; EX1002 ¶86.

Claim 1 of the ’830 patent recites bait sets that target subgenomic intervals with somatic mutations present at frequencies of less than 5% and more than 10% of cells of the tumor sample. See discussion of elements (i) and (ii) of claim 1, below. The actual frequency of a given somatic mutation in cells of a tumor sample would be unknown prior to performing the method of the claims. The ’830 patent does not disclose how to determine actual frequency of a somatic mutation prior to performing the method such that baits may be chosen in advance to account for actual frequency of a given somatic mutation in a tumor sample. The ’830 patent states that public databases, such as COSMIC, contained “likelihood” values for

somatic mutations (*e.g.*, point mutations in KRAS) in tumors of given types (*e.g.*, colon cancer). *See, e.g.*, EX1001, 120:8-14, Table 6; *see also* EX1002 ¶¶87-89.

In the related district court litigation, patent owner has taken the position that the mutation frequencies for bait sets (i) and (ii) as recited in the claim are determined after assessment—after the enrichment and sequencing method is performed and the results are analyzed. The prior art at the time disclosed how to measure somatic mutation frequency in a tumor sample following targeted enrichment and next generation sequencing. This is illustrated, for example, by Shah. Shah discloses use of targeted enrichment and NGS methods to identify and determine the frequency of somatic mutations in cancer specimens from a patient. EX1005 at 809, 811; EX1006 at 13. In particular, Shah discloses enrichment and NGS of genomic DNA derived from a metastasis and a primary lobular breast cancer tumor from the same patient. *Id.*; *see also* EX1002 ¶¶90-94.

Dr. Quackenbush explains a POSA would have been motivated to use somatic mutation information as provided by Shah to inform, and even improve, the enrichment protocol using bait hybridization described in Levin. For example, Levin expressly states that information about target signal abundance can be used during bait design to focus on low abundance signals. EX1010 at 6. Shah discloses detecting somatic mutation signal abundance. *See also* EX1002 ¶¶95-97.

As explained by Dr. Quackenbush, a POSA would have been familiar with the concept that increasing signal relative to noise enhances the ability to detect an event (*e.g.*, a mutation). Specifically, it was known that the sensitivity for detection of mutations is related to frequency of the interval comprising the mutation in the sample. This is expressly acknowledged in Levin. For instance, Levin indicates that if low frequency subgenomic intervals are a priority, then bait design should be used to increase representation. EX1010 at 6. That is, low frequency intervals need be sequenced to a greater depth than high frequency intervals for efficient detection. *See also* EX1002 ¶¶95.

As Dr. Quackenbush further explains, a POSA would have been motivated to utilize information regarding frequency of target somatic mutations in practicing a method as in Levin because different patients, and even different tumors in an individual patient, can present different mutation profiles. As such, a POSA would have been motivated to combine the teachings of Levin and Shah, to obtain an obvious and predictable combination of complementary features. *See KSR Intern. Co. v. Teleflex Inc.*, 550 US 398, 416 (2007) (“The combination of familiar elements according to known methods is ... obvious when it does no more than yield predictable results.”); section I.C(i) above; *see also* EX1002 ¶¶97, 98. An element-by-element discussion of the claims, together with discussion illustrating

exemplary prior art disclosure and how each and every aspect of the challenged claims is found in the prior art, is provided below.

i. Independent Claim 1

Assuming that the preamble is limiting, it is disclosed by Levin and Shah:

'830 Patent	Levin and Shah
<p>“A method of analyzing a tumor sample for a somatic mutation, comprising:”</p>	<p>Levin and Shah disclose methods for analyzing a tumor sample for a somatic mutation:</p> <p>“Targeted RNA-Seq combines next-generation sequencing with capture of sequences from a relevant subset of a transcriptome. When testing by capturing sequences from a tumor cDNA library by hybridization to oligonucleotide probes specific for 467 cancer-related genes, this method showed high selectivity, improved mutation detection enabling discovery of novel chimeric transcripts,” EX1010 at 1.</p> <p>“Recent advances in next generation sequencing have made it possible to precisely characterize all somatic coding mutations that occur during the development and progression of individual cancers. ... We found 32 somatic non-synonymous coding mutations” EX1005 at 809.</p> <p><i>See also</i> EX1010 at 2 (“cDNA hybrid selection” and “Background”); at 6 (“Conclusions”); EX1002 ¶¶99-102.</p>

Levin discloses that the starting material used in the study was a “cDNA library ... constructed from the K-562 chronic myeloid leukemia (CML) cell line.” EX1010 at 2. Levin also discloses that the method is suitable “[f]or research into somatic mutations in cancer” and “has the advantage of enriching for changes in

coding sequences, which are more likely to affect function, compared with sequencing genomic DNA.” *Id.* at 1. Levin discloses its methods can be used in analyzing clinical tumor samples and “is well suited for a wide range of large-scale tumor-profiling studies in many clinical or research settings.” *Id.* at 6; *see also* EX1002 ¶¶100, 101.

Shah also discloses a method of analyzing a tumor sample for a somatic mutation. For example, Shah discloses that “advances in next generation sequencing have made it possible to precisely characterize all somatic coding mutations that occur during the development and progression of individual cancers.” EX1005 at 809. Shah further discloses using “these approaches to sequence the genomes ... and transcriptomes of an ... metastatic lobular breast cancer at depth” and identification of 32 somatic mutations. *Id.*; *see also* EX1002 ¶102.

Accordingly, Levin and Shah disclose a method of analyzing a tumor sample for a somatic mutation. EX1002 ¶103.

Levin and Shah disclose limitation 1 (a):

'830 Patent	Levin and Shah
“(a) acquiring a library comprising a plurality of tumor members from the tumor sample;”	<p>Levin and Shah disclose acquiring a library of tumor members from a tumor sample: “When testing by capturing sequences from a tumor cDNA library by hybridization to oligonucleotide probes specific for 467 cancer-related genes” EX1010 at 1.</p> <p>“[W]e created a complex pool of biotinylated</p>

	<p>oligonucleotide probes (baits) for cancer-related transcripts and used them to capture cDNAs from a library prepared for Illumina sequencing.” EX1010 at 2.</p> <p>“For sequencing library preparation, we extracted DNA from formalin-fixed paraffin embedded blocks containing the primary lobular tumour, using a modified DNA extraction protocol as described below.” EX1006 at 2.</p> <p>“Genomic DNA was sheared ... The library was prepared by following the pair end library protocol (Illumina Inc., USA).” EX1006 at 7.</p> <p><i>See also</i> EX1010 at 1 (“Abstract”); at 2 (“cDNA hybrid selection”); at 6 (“Illumina library construction and sequencing”); EX1006 at 1-2 (“Tumour characteristics and clinical information”); at 6-7 (“Illumina Genome Sequencing Library preparations”); EX1002 ¶¶104-106.</p>
--	---

Both Levin and Shah disclose acquiring a library comprising a plurality of tumor members from the tumor sample, as recited in the claim. For example, Levin discloses producing “a tumor cDNA library” from a chronic myeloid leukemia cell line using a protocol known in the art. EX1010 at 6. Levin further discloses that the library comprised “290 to 390 bp” fragments that were suitable “for Illumina sequencing.” *Id.*; *see also* EX1002 ¶¶104, 105.

Shah discloses libraries produced from “a cryopreserved pleural effusion of metastatic lobular breast carcinoma” and “formalin-fixed paraffin embedded blocks containing the primary lobular tumour.” EX1006 at 1-2. Moreover, Shah

discloses that genomic DNA was isolated from tumor samples and a nucleic acid library prepared from these samples using standard protocols. EX1006 at 6-7. Shah further discloses that the nucleic acids were “sheared” and “[t]he library was prepared” according to Illumina’s “pair end library protocol.” EX1006 at 7; *see also* EX1002 ¶106.

Accordingly, Levin and Shah disclose acquiring a library comprising a plurality of tumor members from the tumor sample. EX1002 ¶107.

Levin and Shah disclose limitation 1 (b):

'830 Patent	Levin and Shah
<p>(b) contacting the library with at least two bait sets to provide selected tumor members, wherein said bait sets hybridize with the tumor members, thereby providing a library catch;</p>	<p>Levin and Shah disclose at least two bait sets that hybridize with and select tumor members:</p> <p>“[W]e created a complex pool of biotinylated oligonucleotide probes (baits) for cancer-related transcripts and used them to capture cDNAs from a library prepared for Illumina sequencing.” EX1010 at 2.</p> <p>“We designed 11,566 bait sequences” EX1010 at 6.</p> <p>“[W]e next examined genomic DNA from the primary tumour directly, by a single molecule frequency counting experiment Twenty-eight of the 32 mutations yielded amplicons compatible with Illumina sequencing.” EX1005 at 811.</p> <p>Levin and Shah disclose its hybridization capture as providing a library catch:</p> <p>“Sequence analysis of the cDNA library after hybrid selection demonstrates that nearly all the high-quality, aligning reads derive from targeted genes.” EX1010 at 2.</p> <p>“All of the control germline and somatic amplicons were pooled together for each of the DNA sources. Paired end sequences were collected from an Illumina GA2 with a standard protocol for 75 bp reads.” EX1006 at 8.</p> <p><i>See also</i> EX1010 at 2 (“cDNA hybrid selection”); at 6 (“Bait design and synthesis” and “Hybrid selection”); EX1005 at 811; EX1006 at 8 (“Frequency analysis”); EX1002 ¶¶108-110.</p>

Levin discloses contacting the library with at least two bait sets to provide selected tumor members, as recited in the claim. For example, Levin discloses

designing “11,566 bait sequences” that “target[] the coding sequence of 887 transcripts.” EX1010 at 6. Levin further discloses the selection of “cDNAs hybridizing with these cancer cDNA baits” from the tumor library thereby providing a library catch for sequencing. EX1010 at 2. For example, Levin discloses “[s]equence analysis of the cDNA library after hybrid selection” and that the “reads derive from targeted genes.” *Id.*; *see also* EX1010 at 6 (“Illumina library construction and sequencing”); EX1002 ¶109.

Shah discloses contacting the library with PCR primers and that “[t]wenty-eight of the 32 mutations yielded amplicons compatible with Illumina sequencing.” EX1005 at 811. Shah discloses sequencing the library catch. For instance, Shah discloses that the “somatic amplicons were pooled” and “sequences were collected from an Illumina GA2.” EX1006 at 8; *see also* EX1002 ¶110.

Accordingly, Levin and Shah disclose contacting the library with at least two bait sets to provide selected tumor members, wherein said bait sets hybridize with the tumor members, thereby providing a library catch. EX1002 ¶111.

Levin and Shah disclose limitation 1 (c):

'830 Patent	Levin
<p>“(c) sequencing by a next generation sequencing method a subgenomic interval comprising the somatic mutation from a tumor member from said library or library catch, thereby acquiring a read for the subgenomic interval;”</p>	<p>Levin and Shah disclose NGS of subgenomic intervals from tumor members in a library catch: “From an aliquot of this library, we selected cDNAs hybridizing with these cancer cDNA baits. We used ...the Illumina Genome Analyzer platform.” EX1010 at 2.</p> <p>Levin and Shah disclose NGS of subgenomic intervals comprising somatic mutations: “Targeted next-generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts” EX1010 at 1.</p> <p>“For research into somatic mutations in cancer....” EX1010 at 1.</p> <p>“The PCR amplicons for known germline and somatic mutations were sequenced on an Illumina device.” EX1005 at 811.</p> <p><i>See also</i> EX1010 at 1 (“Background”); at 2 (“cDNA hybrid selection” and “Sequence enrichment”); at 5 (“Discussion”); EX1005 at 809 (“Abstract”); EX1002 ¶¶112-115.</p>

Levin discloses sequencing by a NGS method a subgenomic interval comprising the somatic mutation from a tumor member from said library catch. Indeed, “next-generation sequencing” is identified prominently in the title of paper. EX1010 at 1. Moreover, Levin discloses that “for research into somatic mutations in cancer...this method has the advantage of enriching for changes in coding sequences, which are more likely to affect function.” EX1010 at 1. Levin also

discloses “identif[ication of] nonreference bases, including SNPs and candidate mutations” in the tumor sample. EX1010 at 2; *see also* EX1002 ¶¶113, 114.

Shah also discloses NGS to identify somatic mutations from its enriched product (*i.e.*, library catch). Shah discloses that “advances in next generation sequencing have made it possible to precisely characterize all somatic coding mutations that occur during the development and progression of individual cancers.” EX1005 at 809. Moreover, Shah discloses application of “these approaches to sequence the genomes ... and transcriptomes of an oestrogen-receptor- α -positive metastatic lobular breast cancer at depth.” *Id.* Shah further discloses that “PCR amplicons for known germline and somatic mutations were sequenced on an Illumina device.” EX1005 at 811; *see also* EX1002 ¶115.

Accordingly, Levin and Shah disclose sequencing by a NGS method a subgenomic interval comprising the somatic mutation from a tumor member from said library catch. EX1002 ¶116.

Levin and Shah disclose limitation 1 (d):

'830 Patent	Levin and Shah
(d) aligning said read by an alignment method; and	<p>“Purity-filtered [13] 76-mer reads were aligned to all curated protein-coding transcripts in RefSeq (downloaded from [33] on March 1, 2009) allowing up to four mismatches, and mapped back to their genomic coordinates in the reference human genome (hg18), preserving splice junctions. Alignments were performed by using the ImperfectLookupTable (ILT) tool of the ARACHNE genome assembly suite [34].” EX1010 at 6.</p> <p>“Short read sequences obtained from the Illumina Genome Analyzer for WGSS-PE and WGSS-PRI were mapped to the reference human genome (NCBI build 36.1, hg18) using Maq in paired end mode. The WTSS-PE reads were mapped to the human reference genome plus a database of known exon junctions (Morin et al Biotechniques (2008)).” EX1006 at 9.</p> <p><i>See also</i> EX1002 ¶¶117-119.</p>

Aligning reads is a basic aspect of NGS. Levin and Shah each disclose aligning reads resulting from NGS to a reference human genome using alignment programs known in the art. EX1010 at 6; EX1006 at 9. As admitted in the '830 patent, numerous “sequence alignment algorithms/programs for short-read sequences” were known in the art. EX1001, 103:23-104:6. Both the ARACHNE (Levin) and Maq (Shah) programs are admitted in the '830 patent as well-known prior art alignment methods. EX1001, 27:48-60, 104:21-32, 117:19-22, 118:9-12. In particular, ARACHNE is described as useful for “subgenomic intervals resistant to accurate alignment” and Maq is described as being useful for aligning short

reads particularly “when accuracy is of the essence.” EX1001, 104:21-32, 117:19-22; *see also* EX1002 ¶¶117-119.

Accordingly, Levin and Shah disclose aligning said read by an alignment method. EX1002 ¶120.

Levin and Shah disclose limitation 1 (e):

'830 Patent	Levin and Shah
<p>(e) assigning a nucleotide value from said read for a preselected nucleotide position,</p>	<p>“Bases that disagreed with the reference genome were classified as known SNPs if present in dbSNP [35] (build 129) or as novel variants. Novel variants were discarded if they occurred within five bases of another novel variant (to compensate for alignment artifacts produced by indels), if they were observed on Illumina reads in only one orientation, or if they fell within segmental duplications[36]. The remaining novel variants were considered high confidence....” EX1010 at 7.</p> <p>“SNVs predicted from the short read alignments become candidate germline or somatic mutations that are important to identify in order to catalogue the mutational spectrum of the genome (or transcriptome) under investigation. We developed a Bayesian mixture model called SNVmix for this purpose that takes as input the complete set of aligned reads and produces as output the probability of an SNV at each position represented in the data.” EX1009 at 1.</p> <p><i>See also</i> EX1010 at 2 (“Sequence enrichment”); at 7 (“Splice isoform identification”); EX1005 at 811 (Table 3); EX1002 ¶¶121-123.</p>

The '830 patent further admits that methods were known in the art for “calculat[ing] the genotype likelihood associated with a specific genotype and position.” EX1001, 119:29-43. That is, methods were known for assigning

nucleotide values, including assigning values to positions known in advance to be potentially variant in an individual (*e.g.*, SNPs, SNVs, indels). *Id.*; *see also* 120:63-121:3; Section V above.

Levin discloses the identification of nonreference bases in reads from subgenomic intervals targeted with the bait sets. EX1010 at 2. Specifically, Levin discloses use of SNP databases to confirm the assignment of nucleotide values for sites known to be variant in the population. EX1010 at 7; *see also* EX1002 ¶122.

Shah also discloses assigning nucleotide values to preselected nucleotide positions. For instance, Shah discloses use of *SNVMix* to assign nucleotide values to positions where the variant is present in less than all of the cells in the sample. EX1009 at 1. The *SNVMix* algorithm used by Shah, is acknowledged by the '830 patent as a prior art algorithm for assigning nucleotide values where the frequency of mutation may be less than 50% or 100% “for the analysis of cancer DNA.” EX1001, 119:55-58; *see also* EX1002 ¶123.

Accordingly, Levin and Shah disclose assigning a nucleotide value from said read for a preselected nucleotide position. EX1002 ¶124.

Claim 1 further recites “*wherein the at least two bait sets of step (b) are chosen from two of the following bait sets.*” The “following bait sets” include bait sets (i)-(v).

Levin and Shah disclose the “at least two bait sets” limitation. For example, Levin and Shah disclose bait sets at least according to categories (i) and (ii), as discussed in more detail below. EX1002 ¶125.

The '830 patent instructs that relevant information for designing the recited bait sets was available from public resources at the time. As such, at least bait sets (iii), (iv) and (v) would also have been obvious in view of the cited references (*e.g.*, Levin) and knowledge in the art at the time, as discussed in more detail below. EX1002 ¶126.

Levin and Shah disclose limitation 1 (i):

'830 Patent	Levin and Shah
<p>(i) a first bait set that selects a high-level target chosen from one or more tumor nucleic acid molecules that comprise a subgenomic interval comprising a somatic mutation that appears at a frequency of about 5% or less of the cells from the tumor sample;</p>	<p>Levin discloses bait sets targeting subgenomic intervals that comprise somatic mutations: “We targeted 467 genes in total ... and genes catalogued in the Cancer Gene Census.” EX1010 at 2.</p> <p>Levin discloses that the bait sets are suitable for detection of low frequency subgenomic intervals: “[T]argeted RNA-Seq produces an enhanced view of the molecular state of a set of ‘high-interest’ genes.” EX1010 at 1.</p> <p>“[T]his method ...enabled detection of novel fusion transcripts and isoforms thereof that would otherwise have escaped detection.” EX1010 at 2.</p> <p>“If coverage of lower abundance transcripts is a priority in a given experiment, information about transcript abundance can be used during bait design to focus on those transcripts” EX1010 at 6.</p> <p>Shah discloses measuring somatic mutation frequency in a tumor sample using enrichment and</p>

	<p>sequencing: “We found 32 somatic non-synonymous coding mutations present in the metastasis, and measured the frequency of these somatic mutations in DNA from the primary tumor of the same patient” EX1005 at 809.</p> <p>Shah discloses in Table 3 detecting somatic mutations at frequencies of about 5% or less of cells. EX1005 at 811 (see entries for MORC1, KIF1C, USP28, KIAA1468, and RNASEH2A).</p> <p><i>See also</i> EX1010 at 1 (“Abstract” and “Background”); at 2 (“Sequence enrichment”); at 5 (“Discussion”); EX1006 at 13 (“Analysis of frequency PCR amplicons”); EX1002 ¶¶127-138.</p>
--	--

Levin discloses evaluating, via its enrichment protocol using bait hybridization, “high-interest” genes, including instances of low abundance signal. EX1010 at 1, 6. Moreover, Levin discloses a method that “enabled detection of novel fusion transcripts and isoforms thereof that would otherwise have escaped detection.” EX1010 at 2. That is, Levin discloses a method with “high selectivity, improved mutation detection” for rare subgenomic intervals in a library. EX1010 at 1. A POSA would appreciate that somatic mutations are often present at low frequencies (*e.g.*, less than 5% of cells) in tumor samples and thus represent examples of rare subgenomic intervals that may benefit from the improvements in detection provided by the method disclosed by Levin. *See also* EX1002 ¶128.

Moreover, Levin discloses that different mutations will present at different levels of abundance in a sample, and instructs that signal abundance should be

factored into bait design in advance to account for a desired depth of sequence coverage. EX1010 at 6. The '830 patent does not disclose how to determine actual frequency of a somatic mutation prior to performing the method. The '830 patent states that public databases, such as COSMIC, contained “likelihood” values for somatic mutations (*e.g.*, point mutations in KRAS) in tumors of given types (*e.g.*, colon cancer). EX1001, 120:8-16, Table 6; *see also* Section C (i) above; EX1002 ¶¶129-131.⁴

Wherein frequency, as recited in the challenged claims, is referring to actual frequencies of the somatic mutation in cells of the tumor sample, Shah illustrates that it was known in the art how to measure somatic mutation frequency in a tumor sample. Shah discloses use of targeted enrichment and NGS methods to identify and determine the frequency of somatic mutations in tumor samples from a patient. EX1005 at 809, 811 (Table 3); EX1006 at 13; *see also* EX1002 ¶133.

⁴ To the extent that the bait set of (i) can be read to require a particular depth of sequence coverage, this too is provided by Levin. EX1001, 67:66-68:8. Levin discloses an average depth of sequence coverage of greater than 600×. EX1010 at 2. Also, Shah discloses that the subgenomic intervals containing somatic mutations associated with MORC1, KIF1C, USP28, KIAA1468, and RNASEH2A were sequenced with 24,572-, 8,587-, 6,654-, 719-, 6,537-fold coverage, respectively. EX1005 at 811 (Table 3) (*see* “Primary depth”); *see also* EX1002 ¶¶132, 137.

Shah discloses sequencing of genomic DNA and cDNA derived from a metastasis and a primary lobular breast cancer tumor from the same patient. EX1005 at 809. Shah also discloses targeted enrichment of subgenomic intervals comprising somatic mutations. *Id.* Shah discloses sequencing the amplicons on an Illumina NGS device to analyze the mutation and its frequency within the tumor sample. EX1005 at 811. Shah disclosed that somatic mutations identified in the primary tumor were found in a wide range of frequencies. *Id.*; *see also* EX1002 ¶134, 135.

At least some of the somatic mutations evaluated by Shah were detected as appearing at a frequency of about 5% or less of the cells from the tumor sample, as illustrated expressly in at least Table 3 (annotated). EX1005 at 811. As Dr. Quackenbush explains, the frequency—expressed as a percentage of cells—is

obtained by multiplying the Primary NR ratio by 100. See EX1002 ¶136.

Table 3 | Frequency of germline and somatic alleles in the metastatic and primary genomes

Position	Locus	R	NR	Primary depth	Primary NR ratio	Primary Pvalue	Primary status	Metastasis depth	Metastasis NR ratio	M	Copy number classification (HMM state)
4:2203607	HAUS3	C	T	5700	0.5472	0.0000	Dominant	762	0.7874	S	Neutral (2)
16:23559936	PALB2	T	G	115	0.4957	0.0000	Dominant	669	0.4350	S	Amplification (4)
2:169497197	ABCB11	C	T	506	0.3261	0.0000	Dominant	959	0.3691	S	Amplification (4)
14:92018836	SLC24A4	G	A	13347	0.2341	0.0000	Dominant	13670	0.3518	S	Amplification (4)
17:10248420	MYH8	C	G	10657	0.1353	0.0000	Subdominant	1797	0.5932	S	Neutral (2)
3:110271286	MORC1	G	A	24572	0.0468	0.0000	Subdominant	32273	0.4107	S	Gain (3)
17:4848025	KIF1C	G	A	8587	0.0107	0.0000	Subdominant	2272	0.3077	S	Neutral (2)
11:113185109	USP28	C	T	6654	0.0095	0.0000	Subdominant	1387	0.4484	S	Gain (3)
18:58076768	KIAA1468	G	A	719	0.0083	0.0020	Subdominant	1056	0.3059	S	Neutral (2)
19:12785252	RNASEH2A	G	A	6537	0.0029	0.0276	Subdominant	1497	0.2806	S	Neutral (2)
4:106982671	GSTCD	G	T	7273	0.0008	0.9885	Absent	2208	0.2174	S	Neutral (2)
17:35701114	CDC6	G	T	4894	0.0008	0.9733	Absent	4208	0.3577	S	Amplification (4)
17:7751231	CHD3	G	A	9665	0.0007	0.9981	Absent	1737	0.2671	S	Neutral (2)
4:155726802	FGA	C	T	5756	0.0007	0.9911	Absent	2287	0.2755	S	Gain (3)
17:7052251	DLG4	G	A	4383	0.0007	0.9835	Absent	706	0.3272	S	Neutral (2)
3:37267947	GOLGA4	G	T	13051	0.0006	0.9999	Absent	3262	0.2235	S	Gain (3)
9:84867250	RASEF	G	T	1690	0.0006	0.9500	Absent	796	0.3656	S	Gain (3)
17:35133783	ERBB2	C	A	3736	0.0005	0.9899	Absent	1722	0.3612	S	Amplification (4)
X:86659878	KLHL4	C	T	6561	0.0005	0.9993	Absent	977	0.3153	S	Neutral (2)
3:191172415	LPREL1	T	C	11963	0.0004	1.0000	Absent	8381	0.2148	S	Gain (3)
16:73500342	WDR59	C	T	4846	0.0004	0.9982	Absent	1396	0.2629	S	Neutral (2)
1:44650831	RNF220	G	A	8160	0.0004	0.9999	Absent	967	0.2203	S	Neutral (2)
22:45035285	PKDREJ	C	T	6674	0.0003	0.9999	Absent	1230	0.3366	S	Gain (3)
11:61313958	C11ORF10	G	A	116705	0.0003	1.0000	Absent	14354	0.4651	S	Amplification (4)
12:52063157	SP1	G	T	7732	0.0003	1.0000	Absent	2011	0.2193	S	Amplification (4)
11:77452594	THRSF	C	T	24219	0.0002	1.0000	Absent	40652	0.4750	S	Gain (3)
17:45625043	COL1A1	C	A	26343	0.0001	1.0000	Absent	32259	0.2543	S	Amplification (4)
13:77076497	SCEL	A	G	49	0.0000	1.0000	Absent	187	0.5722	S	Gain (3)
19:9314428	—	A	G	176	0.5057	0.0000	Present	321	0.4953	G	Neutral (2)
4:130144460	—	A	T	2020	0.2188	0.0000	Present	2081	0.3099	G	Neutral (2)
8:27835012	—	G	A	13587	0.8602	0.0000	Present	10781	0.6667	G	Deletion (1)
6:32908543	—	C	T	4718	0.7484	0.0000	Present	16370	0.4897	G	Amplification (4)
20:43363061	—	G	A	5950	0.5249	0.0000	Present	5540	0.5049	G	Amplification (4)
4:8672089	—	G	A	381	1.0000	0.0000	Present	2850	0.8032	G	Gain (3)
16:1331138	—	C	T	677	0.4963	0.0000	Present	554	0.6245	G	High-level amplicon (5)

Only 7 germline alleles are shown, the full list is in Supplementary Table 7. The genome positions are shown as chr:coordinate. The primary read depth represents the number of reads. Binomial exact P values were calculated using a Binomial exact test. R, reference base; NR, non-reference base. Primary status indicates whether the variant was present, subdominant or absent in the primary tumour. Column M denotes somatic (S) or germline (G) single nucleotide variants in the metastasis. HMM state refers to the metastasis.

Thus, Shah describes identifying somatic mutations and the frequency with which they occur, information that would benefit the methods of Levin. As discussed above, a POSA would have been motivated to use somatic mutation information as provided by Shah to inform, and even improve, the enrichment protocol using bait hybridization described in Levin. For example, Levin indicates that information regarding target signal abundance, information of the type derived according to the methods of Shah, would be beneficial in tailoring the bait design in Levin and improving mutation detection. See also EX1002 ¶138.

As explained by Dr. Quackenbush, a POSA would have known that bait parameters, such as tiling frequency, length, and GC-content, could be manipulated to predictably and reliably alter target hybridization and sequence coverage. EX1010 at 6; *KSR Intern. Co.*, 550 US at 416. Levin confirms the importance of these bait parameters and directly states that manipulation of bait parameters should be used to control sequence coverage. *Id.*; *see also* EX1002 ¶¶36-41, 138.

Accordingly, Levin and Shah disclose a first bait set that selects a high-level target chosen from one or more tumor nucleic acid molecules that comprise a subgenomic interval comprising a somatic mutation that appears at a frequency of about 5% or less of the cells from the tumor sample. EX1002 ¶139.

Levin and Shah disclose limitation 1 (ii):

'830 Patent	Levin and Shah
(ii) a second bait set that selects a mid-level target chosen from one or more tumor nucleic acid molecules that comprise a subgenomic interval comprising a somatic mutation that appears at a frequency of about 10% or higher of the cells from the tumor sample;	<i>See</i> discussion of limitation 1 (i) above; <i>see also</i> EX1010 at 1 (“Title,” “Abstract,” and “Background”); EX1010 at 2 (“Sequence enrichment”); EX1010 at 5 (“Discussion”); EX1010 at 6 (“Conclusions”); EX1005 at 809; EX1005 at 811; EX1006 at 13 (“Analysis of frequency PCR amplicons”); EX1002 ¶¶140-142.

Levin and Shah disclose limitation 1(ii) for at least a similar rationale as set forth above regarding limitation 1(i). Levin discloses evaluating, via its enrichment protocol using bait hybridization, a large number of different genes—including

“high-interest” genes—that present at a range of signal abundance. EX1010 at 1. Furthermore, as discussed above a POSA would have been aware of public resources for somatic mutation information available at the time, as disclosed in Levin and in the ’830 patent. *See also* EX1002 ¶¶141.

Wherein frequency, as recited in the challenged claims, is referring to actual frequencies of the somatic mutation in cells of the tumor sample, Shah illustrates that it was known in the art how to measure somatic mutation frequency in a tumor sample. EX1005 at 809, 811 (Table 3); EX1006 at 13; *see also* EX1002 ¶142.

Furthermore, it would have been apparent to a POSA that at least some of the somatic mutations evaluated by Shah were detected at frequencies of about 10% or higher of the cells from the tumor sample. EX1005 at 811 (Table 3) (*see* “Primary NR ratio” for MYH8, SLC24A4, ABCB11, PALB2, and HAUS3); *see also* EX1002 ¶142.

Accordingly, Levin and Shah disclose a second bait set that selects a high-level target chosen from one or more tumor nucleic acid molecules that comprise a subgenomic interval comprising a somatic mutation that appears at a frequency of about 10% or more of the cells from the tumor sample. EX1002 ¶143.

Levin and Shah disclose limitation 1 (iii):

'830 Patent	Levin and Shah
(iii) a third bait set that selects a low-level target chosen from one or more nucleic acid molecules	Levin discloses bait sets for identification of SNPs: “Hybrid selection enabled us to identify 257 known SNPs at high confidence (LOD > 5) in the coding sequences of target genes, compared with only 76

<p>that comprise a sub-genomic interval chosen from one or more of:</p> <p>a) a pharmacogenomic (PGx) single nucleotide polymorphism (SNP) that distinguishes the ability of a patient to metabolize different drugs,</p> <p>b) a plurality of genomic SNPs that uniquely identify (fingerprint) a patient, or</p> <p>c) a genomic SNP or locus that is used to assess copy number gains or losses of genomic DNA and loss-of-heterozygosity (LOH);</p>	<p>before hybrid selection.” EX1010 at 2.</p> <p>Shah discloses bait sets for identification of SNPs that are used to assess copy number gains or losses: “As controls we selected 36 heterozygous germline SNVs at random. The PCR amplicons for known germline and somatic mutations were sequenced on an Illumina device.” EX1005 at 811.</p> <p>Shah discloses 36 SNPs. EX1005 at 811; EX1008 at 979-980.</p> <p><i>See also</i> EX1002 ¶¶144-146.</p>
---	---

Levin and Shah disclose this limitation or at least renders it obvious in view of the state of the art. The bait set described by clause (iii) requires selection of one of the bait sets described in (a), (b), or (c). For instance, bait set (b) is directed to SNPs that could be used to “uniquely identify” a patient, which is taught or suggested by Levin. Levin discloses that “[h]ybrid selection enabled us to identify 257 known SNPs at high confidence (LOD > 5) in the coding sequences of target genes, compared with only 76 before hybrid selection.” EX1010 at 2. A POSA would have appreciated that the 257 SNPs disclosed by Levin would be suitable for patient identification as claimed. Dr. Quackenbush explains that 257 SNPs is far in excess of that needed for forensic identification of individuals and would

more than suffice to uniquely identify a patient. Levin discloses this limitation or at least renders this limitation obvious in view of the state of the art. *See* EX1002 ¶144.⁵

Additionally, the bait set of (a), the recited PGx SNP that distinguishes ability to metabolize different drugs, would have been obvious in view of the cited references. As indicated, Levin discloses bait sets targeting SNP-containing subgenomic intervals. As Dr. Quackenbush explains, various PGx SNPs that distinguish ability of a patient to metabolize different drugs were known at the time and described in the literature. Dr. Quackenbush further explains that such SNPs would have been of interest for bait targeting in a method as in Levin because medical treatment of a cancer patient would benefit from information regarding patient metabolism of different drugs. *See* EX1002 ¶146.

Accordingly, Levin and Shah disclose or render obvious a third bait set that selects a low-level target chosen from one or more nucleic acid molecules that comprise a subgenomic interval chosen from one or more of the bait sets recited in (iii)(a), (b) or (c) of claim 1. EX1002 ¶147.

⁵ Shah also discloses the identification of SNPs. EX1005 at 811. Dr. Quackenbush explains that the 36 SNPs disclosed by Shah would allow unique identification of one patient compared to another. *See* EX1002 ¶145.

Claim 1(iv): “(iv) a fourth bait set that selects a nucleic acid molecule that comprises an intron sequence that detects a structural breakpoint;”

The limitation of clause (iv) would have been obvious over Levin and Shah in view of knowledge in the art. A structural breakpoint is a position in the genome that has suffered a double-strand break (*e.g.*, translocations, indels). EX1001, 2:50-51. Levin discloses in the cDNA context bait sets that select an *exon* sequence that detects a structural breakpoint. EX1010 at 5 (Table 2. As explained by Dr. Quackenbush, it would have been apparent to a POSA that Table 2 of Levin is reflecting detection of exon sequences that detect a structural breakpoint because the first three entries in Table 2 (*i.e.*, BCR-ABL1, NUP214-XKR3, and SNHG3-RCC1-PICALM) are translocations. *See* EX1002 ¶148.

Dr. Quackenbush further explains that the selection of intron sequences from genomic DNA is not meaningfully different than selecting exon sequences from cDNA, and an intron sequence that detects a structural breakpoint would be of interest for the same reason one would target an exon sequence that detects a structural breakpoint. This would have been recognized by a POSA, and is further evidenced by Chmielecki, which discloses use of enrichment and sequencing methods virtually identical to those of Levin.⁶ Chmielecki discloses using

⁶ Chmielecki further illustrates that the methods of Levin are applicable in both the genomic and cDNA context. EX1002 ¶148.

SureSelect baits to target both introns and exons for the purpose of identifying structural breakpoints in genomic DNA. EX1036 at 6988. A POSA would consider it a simple and obvious task to design baits targeting the claimed intron sequence in view of the highly analogous exon sequence target in Levin. *Id.* Additionally, the publicly available resource for mutation sites identified in the '830 patent would provide the relevant information for corresponding bait design. EX1001, 101:9-11; *see also* EX1022 at D946; EX1002 ¶148.

Accordingly, Levin and Shah render obvious a fourth bait set that selects a nucleic acid molecule that comprises an intron sequence that detects a structural breakpoint. EX1002 ¶149.

Claim 1(v): “(v) a fifth bait set that selects a one-copy deletion of several terminal exons,”

A bait set that selects a one-copy deletion of several terminal exons is taught or suggested by the cited references. EX1002 ¶150.

Copy number variation is a well-known class of mutation, and would have been recognized as a target of interest at the relevant time for a person investigating mutations associated with cancer. A one-copy deletion of several terminal exons represents a decrease in copy number, and a mutation that would be identifiable in the publicly available databases at the time. EX1001, 101:9-11; *see also* EX1022 at D946. As Dr. Quackenbush explains, a POSA would have been motivated to design baits targeting the interval described in limitation (v), and

doing so would have been well within the ordinary skill, *e.g.*, using the publicly available resource for mutation sites identified in the '830 patent. *See also* EX1002 ¶150.

Accordingly, Levin and Shah in view of the admitted publicly available resources (*e.g.*, COSMIC), renders obvious a claim with a fifth bait set that selects a one-copy deletion of several terminal exons. EX1002 ¶151.

Levin and Shah disclose the unique preselected efficiency for selection limitation:

'830 Patent	Levin
wherein each bait set of said plurality has a unique preselected efficiency for selection for its target as compared with the other bait sets in the plurality.	<p>Levin discloses a plurality of bait sets each with a unique preselected efficiency for selection:</p> <p>“Table 2. Hybrid selection-enhanced detection of fusion transcripts” EX1010 at 5.</p> <p>Supplementary table listing validated SNPs. EX1012.</p> <p>“If coverage of lower abundance transcripts is a priority in a given experiment....” EX1010 at 6.</p> <p><i>See also</i> EX1002 ¶¶152-156.</p>

Levin discloses using a plurality of unique bait sets, with different baits having different efficiency for selection for the respective targets as compared with the other bait sets in the plurality. EX1002 ¶152.

As an initial matter, Levin discloses a unique depth of sequence coverage for different targets selected by different bait sets. As discussed above, the '830 patent equates “efficiency of selection” with the depth of sequence coverage. *See* Section V. Levin discloses in Table 2, for instance, a different depth of sequence coverage for subgenomic intervals comprising six fusion transcripts selected by different bait sets. EX1010 at 5; *see also* EX1012 (disclosing different depths of sequence coverage for SNPs). The number of sequence reads—that is, depth of sequence coverage—following targeted enrichment is shown in the column labeled “After” (annotated here).EX1010 at 5. Sequence coverage for the six fusion transcripts

ranges from 874 to 2, illustrating a unique coverage depth for each targeted subgenomic interval. *Id.*; *see also* EX1002 ¶152.

Additionally, as discussed above (*e.g.*, Section

C(i)), the prior art described predictably designing bait sets to account for a desired depth of sequence coverage, or “efficiency for selection” as it is defined in the patent. This is also reflected in Levin. EX1010 at 6. Levin discloses that its bait sets differed in nucleotide content (*e.g.*, GC-content), and indicates that this

Table 2

Hybrid selection-enhanced detection of fusion transcripts

5' Gene	5' Chr.	3' Gene	3' Chr.	Before ^a	After ^a
BCR	22	ABL1	9	13	874
NUP214 ^b	9	XKR3 ^{b, c}	22	9	152
SNHG3-RCC1 ^{b, c}	1	PICALM ^b	11	1	39
PRIMI ^{c, d}	12	NACA ^{b, d}	12	0	22
NCKIPSD ^d	3	CELSR3 ^{c, d}	3	0	5
SLC29A1 ^{c, d}	6	HSP90AB1 ^d	6	0	2

^aNumber of sequence reads before and after hybrid selection. ^bFusion transcript reads identified from more than one exon in this gene. ^cNot included in hybrid selection bait genes. ^dNot previously annotated, read-through transcripts between adjacent genes.

parameter affected the depth of sequence coverage between different bait sets. EX1010 at 3, Fig. 1, Fig. 2. Levin indicates that differences in selection efficiency can be increased or decreased (*e.g.*, introduce bias) at least by virtue of GC-content in different bait sets—a parameter that was in fact used to manipulate bias in the samples analyzed. Levin discloses different bait sets that give rise to different depths of sequencing coverage (*i.e.*, “efficiencies of selection” as defined in the patent). *See also* EX1002 ¶153.

Furthermore, Levin discloses that targeted enrichment may be used to provide roughly even enrichment or deliberately uneven enrichment depending on the priorities of the experiment. For instance, Levin discloses that when subgenomic intervals were enriched in a roughly even manner, information regarding relative expression level could be obtained. EX1010 at 2-3. But Levin also discloses that subgenomic intervals may be deliberately enriched in an uneven fashion, such as when certain low abundance intervals are a priority. EX1010 at 6. Table 2, reproduced above, illustrates circumstances that would motivate using uneven enrichment. EX1010 at 5. With uniform enrichment, the BCR-ABL1 fusion transcript was deeply sequenced with 874 reads but the SLC29A1-HSP90AB1 fusion transcript had only 2 reads. *Id.* If SLC29A1-HSP90AB1 was a priority in an experiment, 2 reads would generally be considered insufficient for sensitive detection. An enhanced enrichment approach would be desirable to a

POSA looking to manage the signal-to-noise for a low abundance signal. *See also* EX1002 ¶154.

Levin also discloses that increasing the relative abundance of the target interval increases the number of reads and hence the sensitivity of the assay. EX1010 at 2. Therefore, to improve the sensitivity for low abundance subgenomic intervals, such as those exemplified by SLC29A1- HSP90AB1, Levin discloses that bait design can be altered to achieve deeper sequencing. EX1010 at 6. That is, Levin discloses that deeper sequencing is beneficial if low abundance transcripts are a priority. *See also* EX1002 ¶155.

In further regard to “preselecting” bait sets, Levin teaches bait design with the goal to achieve the desired depth of sequence coverage. *See* EX1041 ¶21. Levin further discloses various bait parameters, such as length, sequence, and tiling that can be manipulated to provide a preselected depth of sequence coverage for a target subgenomic interval. EX1010 at 6; *see also* Section C(i) above discussing knowledge and predictability of bait hybridization at the time. A POSA would have known how to manipulate the parameters disclosed by Levin to achieve the desired depth of sequence coverage. EX1016; EX1024 at ¶¶72, 73, 87, 104, 130, 153. Specifically, concentration of the bait, overlap between baits on a given target interval (*i.e.*, tiling), and the physical properties of the bait (*e.g.*, length, sequence,

GC-content) were all known to influence depth of sequence coverage of the selected subgenomic interval—and mentioned in Levin. *See also* EX1002 ¶156.

As such, Levin and Shah disclose “wherein each bait set of said plurality has a unique preselected efficiency for selection for its target as compared with the other bait sets in the plurality” as required by claim 1. EX1002 ¶157.

Accordingly, claim 1 would have been obvious over the prior art.

ii. **Dependent Claim 15**

“The method of claim 1, wherein the method comprises sequencing a subgenomic interval chosen from at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 or all of the following: [intervals A through N]

Claim 15 would have been obvious over the prior art. Levin discloses sequencing a subgenomic interval in at least seven of the recited 14 subgenomic intervals (*i.e.*, A-F, and N). EX1011. Specifically, Levin discloses 35 genes recited in A (“at least five...”), 63 genes recited in B (“at least five...”), 93 genes recited in C (“at least five...”), 19 genes recited in D (“at least five...”), 10 genes recited in E (“at least five...”), 29 genes recited in F (“at least five...”), and 16 genes recited in N (“at least five...”). EX1011; *see also* EX1002 ¶158.

To the extent that Levin does not expressly disclose the remaining subgenomic intervals, they were part of the background knowledge in the art. For instance, sequence information for subgenomic intervals comprising a wide variety

of somatic mutations associated with cancer was publicly available. EX1022 at D946; *see also* EX1002 ¶159.

Accordingly, claim 15 would have been obvious over Levin and Shah. EX1002 ¶160.

B. [Ground 2] Claims 17 and 18 are Obvious under 35 U.S.C. § 103 over Levin, Shah, and Stephan

i. Dependent Claim 17

“The method of claim 15, wherein the method further comprises providing a report in electronic, web-based, or paper form, to a patient or to another person or entity, a caregiver, a physician, an oncologist, a hospital, clinic, third party payor, insurance company or government office.”

Either Levin or Shah would appear to provide a “report” as required by claim 17, as each is provided in electronic, web-based or paper form and each to a person or entity. Levin, for example, is available electronically or in paper form via the journal of *Genome Biology*. EX1010 at 1 (stating that the electronic version of the article can be found at genomebiology.com/2009/10/10/R115). Levin was provided to “another person” by virtue of its publication, and to an “entity” by virtue of submission to *Genome Biology*.

To the extent something more meaningful is required, providing a report a patient or physician, *e.g.*, for the purpose of genetic counseling or personalized medicine, was known in the art at the time. For example, Stephan discloses that genomic profiles from biological samples from individuals are obtained using

methods known in the art, including NGS. EX1044 ¶20. Stephan also discloses that the genomic information for an individual may be provided to the individual or to others in the form of a report, e.g., in electronic, web-based, or paper form. EX1044 ¶¶25, 25, 83, 88, 95. Providing data in a “report” form as in Stephan represents a conventional communication means, and an obvious use of generated data. *See also* EX1002 ¶¶161-165.

Accordingly, claim 17 would have been obvious over Levin, Shah, and Stephan. EX1002 ¶166.

ii. **Dependent Claim 18**

“The method of claim 17, wherein said report comprises one or more of: (a) output from the method, comprising the identification of nucleotide values, the indication of the presence or absence of an alteration, mutation, or wild-type sequence for su[b]genomic intervals associated with a tumor of the type of the sample; (b) information on the role of a sequence, an alteration, mutation, or wild-type sequence, in a disease, wherein said information comprises information on prognosis, resistance, or potential or suggested therapeutic options; (c) information on the likely effectiveness of a therapeutic option, the acceptability of a therapeutic option, or the advisability of applying the therapeutic option to a patient having a sequence, alteration or mutation identified in the report; (d) information, or a recommendation on, the administration of a drug, the administration at a preselected dosage or in a preselected treatment regimen, in combination with other drugs, to the patient; (e) wherein not all mutations identified in the method are identified in the report, the report can be limited to mutations in genes having a preselected level of correlation with the occurrence, prognosis, stage, or sus[c]eptability of the cancer to treatment, with a preselected therapeutic option; or (f) is provided to the patient or to another person or entity within 7, 14, or 21 from receipt of the sample by the entity practicing the method.”

As an initial matter, the requirements of at least element (f) are unclear—*i.e.*, “said report comprises one or more of... (f) is provided to the patient...” To the extent that the scope of claim 18 can be determined, the claim is obvious in view of Levin, Shah and Stephan.

Levin and Shah, for example, disclose the limitation recited in (a). Levin discloses identification of nucleotide values, the indication of the presence or absence of an alteration, mutation, or wild-type sequence for su[b]genomic intervals associated with a tumor of the type of the sample. For instance, Levin

discloses the nucleotide value—that is, the sequence—for subgenomic intervals comprising NUP214-XKR3 fusions from a tumor sample. EX1010 at 4 (Figure 3). Shah discloses nucleotide values for germline and somatically mutated nucleotides from a tumor sample. EX1005 at 810 (Table 2). For instance, the subgenomic interval comprising ABCB11, has a somatic mutation at nucleotide position 2:169497197 that represents a C- to-T transition. *Id.* Similarly, Shah discloses somatic mutations in subgenomic intervals present or not present in the primary tumor and the metastasis. EX1005 at 811 (Table 3); *see also* EX1002 ¶¶167, 168.

Levin, Shah, and Stephan also disclose the limitation recited in (b) and (c), which appears to be generically directed to attempted capture of known sequence/prognosis associations. Levin discloses information on the role of a sequence, an alteration, mutation, or wild-type sequence, in a disease, wherein said information comprises information on prognosis, resistance, or potential or suggested therapeutic options. For instance, Levin discloses identification of BCR-ABL1 fusion transcripts, the presence of which was known to be predictive of successful treatment with the commercially available chemotherapeutic “Gleevec.” EX1010 at 5 (Table 2); EX1048 at 1478-1479. Shah discloses amplification of the ERBB2 locus in a tumor sample. EX1005 at 811 (Table 3). A POSA would have appreciated that ERBB2 amplification is predictive of prognosis and responsiveness to various breast cancer treatment options. *See, e.g.*, EX1035 at

6115. Stephan discloses that phenotypic information derived from the genotype of the patient can be used to inform prognosis, susceptibility to disease, therapeutic options and responsiveness to treatment. EX1044 ¶¶77, 94. Stephan further discloses that therapeutic options may change over time, and discloses how new information regarding options may be delivered to patients. EX1044 ¶99; *see also* EX1002 ¶¶169-172.

Regarding the limitation recited in (e), Stephan also discloses that the report may be limited by phenotype and may be further limited to therapeutically actionable phenotypes. EX1044 ¶¶84, 87; *see also* EX1002 ¶173.

Accordingly, claim 18 would have been obvious over Levin, Shah, and Stephan. EX1002 ¶174.

C. [Ground 3] Claims 19, 20, 32, 44, and 45 are obvious under 35 U.S.C. § 103 over Levin, Shah, and Malhis

Claims 19, 20, 32, 44, and 45 additionally recite basic aspects of alignment and nucleotide calling techniques that were well-known at the time. As Dr. Quackenbush explains, numerous alignment and nucleotide calling programs were publicly available at the relevant time and many of these programs were admitted as prior art in the '830 patent. *See* EX1001, 27:48-60, 104:21-32, 117:19-22, 118:9-12, 119:55-63, 120:50-122:3; EX1045 at 492 (Table 3). Such prior art known programs would meet the requirements of the challenged claims. To the extent that claims 19, 20, 32, 44, and 45 can be read to require alignment and

nucleotide calling methods that are optimized for detection of known variants (*e.g.*, SNVs), this is provided by Malhis. *See also* EX1002 ¶¶175, 176

Malhis discloses Slider II which is an algorithm used for alignment and nucleotide calling in NGS reads obtained from an Illumina device. EX1042 at 1029. Malhis further discloses that Slider II improves alignment by accounting for uncertainty in the reference sequence and in base calls within a read. EX1042 at 1029; *see also* EX1002 ¶177.

As explained by Dr. Quackenbush, a POSA would have appreciated that any position in a read that is different than the reference genome is seen by alignment programs as a mismatch. EX1042 at 1029. Mismatches decrease the probability that a read will properly align with a reference genome. Slider II improves alignment and nucleotide calling accuracy in two ways. First, Slider II “align[s] the probability (prb) values at each position” rather than the sequence file to the reference genome. *Id.* Second, Slider II considers known SNPs (“priors”)—that is, positions that are known to vary in the reference sequence. EX1042 at 1030. Slider II calls SNPs when the “confidence accumulated from sequence coverage is higher than the confidence in the reference sequence.” EX1042 at 1031; *see also* EX1001, 119:49-54 (“This method will compare the probability of observing the read data in the presence of a mutation with the probability of observing the read data in the presence of base-calling error alone. Mutations can be called if this comparison is

sufficiently strongly supportive of the presence of a mutation.”); EX1002 ¶¶178-180.

As discussed in detail below, Dr. Quackenbush explains that a POSA at the relevant time period would have a reasonable expectation of successfully combining the teachings of Malhis with those of Levin and Shah because Malhis explicitly teaches that Slider II was developed for analysis of reads generated by Illumina NGS devices. Indeed, Malhis employs Shah’s dataset (EX1005) for the validation of the Slider II program demonstrating that the sequencing dataset of Shah may be analyzed by the method of Malhis. In the same manner, the method of Malhis would obviously have been applied to the dataset of Levin because it too is a dataset derived from an Illumina NGS device and the methodology would have provided for more efficient and accurate analysis of SNPs. The benefits of modifying Shah or Levin according to Malhis would have been readily apparent to a POSA at that time, and the combination would have presented no significant technical challenges. *See also* EX1002 ¶181.

Additionally, a POSA at the relevant time period would have understood that the method of Malhis was readily applicable to analysis of somatic mutation in tumor samples. For instance, Malhis discloses that using data for “priors” —that is, known sequence differences that may be present in NGS reads relative to the reference sequence —in the alignment and nucleotide calling algorithm improves

the identification of SNPs. EX1042 at 1034. A POSA would have appreciated that Malhis's approach to SNPs would apply equally to the analysis of somatic mutations in cancer. As acknowledged by the '830 patent, somatic mutations found in tumor samples are catalogued in a variety of publicly accessible databases. *See, e.g.,* EX1001, 119:31-38. Therefore, the data for nucleotide positions somatically mutated in tumor samples could be used to improve identification of somatic mutations in tumor samples using Slider II. As explained by Dr. Quackenbush, a POSA would have been motivated to apply to the method of Malhis to somatic mutations in cancer because doing so would have improved the accuracy of alignment and nucleotide calling for known variants. *See also* EX1002 ¶¶182, 183.

Accordingly, a POSA would have been motivated to combine the teachings of Levin, Shah, and Malhis, to obtain an obvious and predictable combination of complementary features. *See also* EX1002 ¶184.

i. Dependent Claim 19

“The method of claim 1, wherein a nucleotide value assigned for a nucleotide position in each of X unique subgenomic intervals is assigned by a unique calling method and X is at least 2, and wherein unique subgenomic interval means different from the other X-1 subgenomic intervals, and wherein unique calling method means different from the other X-1 calling methods.”

As explained by Dr. Quackenbush, calling methods are used to assign a value (*e.g.*, A, C, T, or G) to a nucleotide position within an NGS read after it has been aligned to a reference sequence. Dr. Quackenbush further explains that

numerous calling methods were publicly available at the relevant time (*e.g.*, Maq, SNVMix, Slider II, SOAPsnp, VarScan). *See* EX1005 at 809; EX1042 at 1034; EX1045 at 494. To the extent that claim 19 can be read to require the use of at least two distinct nucleotide calling programs (*e.g.*, Maq and Slider II), a POSA would have been appreciated that using more than one nucleotide calling program was beneficial. *See, e.g.*, EX1045 at 494 (“No one tool single-handedly outperforms the others. Indeed, a combination of variant calling algorithms, each tuned to perform optimally for the dataset in hand, is likely to yield the best combination of sensitivity and specificity for variant detection in human genomes.”). Moreover, Shah and Malhis explicitly disclose using at least two unique calling programs. For example, Shah discloses using SNVMix or Maq for nucleotide calling depending on whether the subgenomic interval contains a SNP or a fusion transcript. EX1006 at 9-10, 14-15. Malhis further discloses using Maq and Slider II to call SNPs present in the dataset disclosed by Shah. EX1042 at 1034; *see also* EX1002 ¶185. As such, the cited references disclose use of at least three different calling methods (“unique calling method[s]”) for analysis of a plurality of SNPs (assigning “a nucleotide value assigned for a nucleotide position in each of X unique subgenomic intervals”), as specifically recited in claim 19.

To the extent that claim 19 is read to require a single nucleotide calling program using unique calling parameters for each unique subgenomic interval, this

too is disclosed by Malhis. *See* EX1001, 129:36-38 (“The calling methods can differ, and thereby be unique, e.g., by relying on different Bayesian prior values.”). Malhis discloses calling nucleotide values (*i.e.*, SNPs) in the NGS reads originally disclosed by Shah (EX1005) using a calling program known as Slider II. Slider II is a Bayesian calling method “using known SNPs as priors” —that is, using a dataset containing nucleotide values for known SNPs. EX1042 at 1029, 1033. Malhis discloses that Slider II calls a SNP if the nucleotide is different from the reference nucleotide and the probability that the nucleotide is a SNP exceeds the probability that the nucleotide is a sequencing error. EX1042 at 1031. That is, Malhis discloses calling methods that are unique and differ with respect to prior expectation values for SNPs. *See also* EX1002 ¶¶186, 187.

Accordingly, claim 19 would have been obvious over Levin, Shah, and Malhis. EX1002 ¶188.

ii. Dependent Claim 20

“The method of claim 19, wherein: (i) assigning said nucleotide value is a function of a value which is or represents a prior expectation of observing said read showing a preselected variant at said preselected nucleotide position in a tumor of type; or (ii) assigning said nucleotide value is a function of a set of values which represent the probabilities of observing a read showing a preselected variant at said preselected nucleotide position if the variant is present in the sample at a frequency of at least 1%, at least 5%, at least 10%, or more, or if the variant is absent.”

To the extent that the scope of claim 20 can be determined, the claim would have been obvious in view of Levin, Shah and Malhis.

Malhis discloses limitation (i), “assigning said nucleotide value is a function of a value which is or represents a prior expectation of observing said read showing a preselected variant at said preselected nucleotide position in a tumor of type.” That is, Malhis discloses that assigning a nucleotide value is a function of an expectation that a known variant will be present in a read from a tumor sample. As described above in more detail, Malhis discloses a method for assigning nucleotide values using known SNPs as “priors.” Malhis discloses that SNPs are called when the “confidence accumulated from sequence coverage is higher than the confidence in the reference sequence.” EX1042 at 1031. A POSA would have appreciated that assignment of nucleotide value is a function of the expectation of observing a SNP (or variant) at that position. *See also* EX1046 at 6; EX1002 ¶189.

Shah discloses (ii), “assigning said nucleotide value is a function of a set of values which represent the probabilities of observing a read showing a preselected variant at said preselected nucleotide position if the variant is present in the sample at a frequency of at least 1%, at least 5%, at least 10%, or more, or if the variant is absent.” That is, Shah discloses that assigning a nucleotide value is a function of an expectation that a known variant will be present at the recited frequencies in the tumor sample. For instance, Shah discloses assigning nucleotide values to variant positions using an algorithm called SNVMix. EX1005 at 812. According to the ’830 patent, SNVMix “address[es] limited deviations from frequencies of 50% or 100% for the analysis of cancer DNA.” EX1001, 119:55-58. That is, Shah discloses a method for calling variants that are found in only a small percentage of cells. Specifically, Shah discloses using SNVMix to call variants present at about 1%, at least 5%, at least 10%, and absent from the primary tumor sample. *See* EX1005 at 811 (Figure 3). For example, genes with variant frequency in the primary tumor of 1% includes USP, variant frequency of at least 5% includes MORC1, variant frequency of at least 10% includes MYH8, variant frequency of more than 10% includes HAUS3, and absent include GSTCD. *Id.*; *see also* EX1002 ¶190.

Accordingly, claim 20 would have been obvious over Levin, Shah, and Malhis. EX1002 ¶191.

iii. Dependent Claim 32

“The method of claim 19, wherein a nucleotide position in at least 20 genes from Table 1 or 1A is assigned a nucleotide value.”

Claim 32 is obvious in view of Levin, Shah and Malhis. Specifically, Levin discloses sequencing 93 genes from Table 1 and 27 genes from Table 1A on an Illumina NGS device. EX1010 at 2; EX1011. Malhis discloses that Slider II can assign nucleotide values to SNPs for reads generated on an Illumina NGS device. EX1042 at 1029; *see also* EX1002 ¶¶192-194.

Accordingly, claim 32 would have been obvious over Levin, Shah, and Malhis. EX1002 ¶195.

iv. Dependent Claim 44

“The method of claim 1, wherein: a read from each of X unique subgenomic intervals is aligned with a unique alignment method and X is at least 2, wherein unique subgenomic interval means different from the other X-1 subgenomic intervals, and wherein unique alignment method means different from the other X-1 alignment methods.”

Numerous calling methods were publicly available at the relevant time and many of these methods were admitted as prior art in the '830 patent. EX1001, 27:48-60, 104:21-32, 117:19-22, 118:9-12. To the extent that claim 44 can be read to require the use of at least two distinct alignment programs (*e.g.*, Maq and Slider II), a POSA would have been appreciated that using more than one alignment program was beneficial because each program has advantages and disadvantages with respect to speed, accuracy, and efficient detection of various mutation types

(*e.g.*, single nucleotide variants, indels, rearrangements). Moreover, Malhis discloses using Maq and Slider II to align reads from the dataset disclosed by Shah (EX1005). EX1042 at 1034; *see also* claim 19 discussion above.

To the extent that claim 44 can be read to require an alignment program that uses unique parameters for each unique subgenomic interval, this too is disclosed by Malhis. Malhis discloses alignment of reads to a reference genome using “known SNPs from SNP databases, as priors” to improve alignment. EX1042 at 1030. As explained by Dr. Quackenbush, a POSA would have appreciated that any position in a read that is different than the reference genome is seen by alignment programs as a mismatch. EX1042 at 1029. Mismatches decrease the probability that a read will properly align with a reference genome. Slider II improves alignment accuracy in two ways. First, Slider II “align[s] the probability (prb) values at each position” rather than the sequence file to the reference genome. *Id.* Second, Slider II considers known SNPs (“priors”)—that is, positions that are known to vary in the reference sequence. EX1042 at 1030. That is, Malhis discloses alignment methods that are unique and differ with respect to prior expectation values for SNPs. *See also* EX1002 ¶¶196, 197.

Accordingly, claim 44 would have been obvious over Levin, Shah, and Malhis. EX1002 ¶199.

v. Dependent Claim 45

“The method of claim 44, wherein subgenomic intervals from at least X genes from Table 1 or 1A having the priority 1 annotation, are aligned with unique alignment methods, and X is equal to 10.”

Claim 45 would have been obvious over the combination of Levin, Shah, and Malhis. Specifically, Levin discloses Illumina sequencing of 25 genes (*e.g.*, “X is equal to 10”) from Table 1 or 1A having the priority 1 annotation. EX1010 at 2; EX1011; *see also* EX1002 ¶¶200, 201.

Malhis discloses that Slider II can align reads generated on an Illumina NGS device. EX1042 at 1029, 1034. Moreover, Malhis discloses that the alignment method may utilize known sequence variation to improve alignment. EX1042 at 1030. The alignment process disclosed by Malhis is unique for each subgenomic interval containing a known variant. *See also* EX1002 ¶202.

Accordingly, claim 45 would have been obvious over Levin, Shah, and Malhis. EX1002 ¶203.

D. [Ground 4] Claims 46, 47, and 57 are obvious under 35 U.S.C. § 103 over Levin, Shah, Malhis, and Li

Claims 46, 47, and 57 are generally directed to the detection of structural rearrangements (*e.g.*, insertion and/or deletion variants). Levin and Shah disclose alignment of reads comprising structural rearrangements and include basic details of how such alignments are performed. EX1006 at 13-15; EX1010 at 7. Li

provides additional details of alignment methodology for structural rearrangements. EX1043 at 1754, 1759; *see also* EX1045 at 494.

Dr. Quackenbush explains that a POSA would have used alignment techniques optimized for detection of single nucleotide variants and structural rearrangements because both classes of mutation are relevant to progression of cancer and each represents frequently characterized somatic mutations. A POSA would have further appreciated that using more than one alignment method is often beneficial when multiple classes of mutation are likely to be present in a sample. *See, e.g.*, EX1043 at 1759; EX1045 (Koboldt) at 494; *see also* EX1002 ¶¶204-207. As explained by Dr. Quackenbush, a POSA would have had a reasonable expectation of success in combining the teachings of Li with those of Levin, Shah, and Malhis because Li discloses the BWA alignment program and explicitly teaches that it is suitable for analysis of reads generated by Illumina NGS devices. EX1043 at 1754. As explained in more detail below, a POSA would have been motivated to combine the teachings of Levin, Shah, Malhis, and Li to obtain an obvious and predictable combination of complementary features. *See also* EX1002 ¶208.

i. Dependent Claim 46

“The method of claim 44, comprising a) applying a first unique alignment method to a first genomic interval, a variant of which is associated with a tumor phenotype, wherein the variant is a point mutation from Table 6; b) applying a second unique alignment method

to a second genomic interval, a variant of which is associated with a tumor phenotype, wherein the variant is a rearrangement chosen from a deletion, insertion, or translocation on Table 5; and c) applying a third unique alignment method to a third genomic interval comprising a genomic interval in which variants are not associated with a tumor phenotype or with a tumor of the type in said sample.”

The combination of Levin, Shah, Malhis, and Li disclose the requirements of claim 46.

Table 6, recited in clause (a), lists point mutations found within the coding region of the KRAS gene. Levin discloses generation of NGS reads for the coding region of the KRAS gene. EX1013 at 31. Levin further discloses applying a first unique alignment method to a first genomic interval (*i.e.*, KRAS), a variant of which is associated with a tumor phenotype as recited in limitation (a). Levin discloses using the “ImperfectLookupTable (ILT) tool of the ARACHNE genome assembly suite” to align reads generated on an Illumina NGS device with a reference human genome. EX1010 at 6. The ARACHNE alignment method is acknowledged as a prior art alignment method in the ’830 patent. EX1001, 111:36-42; *see also* EX1002 ¶209.

Malhis discloses an alignment method that can be optimized for detection of known variants. The ’830 patent acknowledges that sites mutated in cancer genomes and the prior expectation of a site being mutated may be obtained from publicly accessible databases. EX1001, 120:10-14. The data in public databases of cancer mutations, (*e.g.*, point mutations in the KRAS gene) may be used to

optimize the alignment method as disclosed by Malhis. EX1042 at 1030. As explained by Dr. Quackenbush, a POSA would have been motivated to use the method of Malhis to align the KRAS reads of Levin because Malhis teaches that incorporation of known variants as priors improves alignment accuracy. EX1042 at 1034; *see also* EX1002 ¶210.

Table 5, recited in clause (b), lists indel mutations found within the coding region of the PTEN gene. Levin discloses generation of Illumina NGS reads for the coding region of the PTEN gene. EX1011 at 15. Li discloses an alignment method, BWA, suitable for alignment of reads comprising structural rearrangements. Specifically, Li discloses that BWA was produced to address the need for gapped alignment of short read sequences in order to identify indels. EX1043 at 1754; *see also* EX1002 ¶¶211, 212.

As described above in detail, Levin, Shah, and Malhis explicitly disclose unique alignment methods for genomic intervals in which variants are not associated with a tumor phenotype as recited in clause (c). That is, Levin, Shah, and Malhis disclose unique alignment methods for SNPs. *See also* discussion of claim 44 above; EX1002 ¶213.

Accordingly, claim 46 would have been obvious over Levin, Shah, Malhis, and Li. EX1002 ¶214.

ii. **Dependent Claim 47**

“The method of claim 44, wherein a subgenomic interval being analyzed comprises a nucleotide position with a genomic rearrangement, and the method comprises using an alignment method that includes: selecting a rearrangement reference sequence for alignment with a read, wherein said rearrangement reference sequence is preselected to align with a preselected rearrangement, wherein the preselected rearrangement reference sequence is not identical to the preselected rearrangement; comparing a read with said preselected rearrangement reference sequence; and determining if said read meets a predetermined alignment criterion, wherein the predetermined alignment criterion is an alignment to said preselected rearrangement reference sequence with less than a preselected level of mismatch or gaps; thereby analyzing a read.”

As an initial matter, it is not clear that claim 47 makes any meaningful limitation to the subject matter of claim 44. Although verbose, claim 44 appears to require nothing more than aligning a read comprising a structural rearrangement to a reference genome using conventional alignment methods. A genomic rearrangement is defined relative to a reference sequence—that is, genomic rearrangements are not identical to the reference sequence. Therefore, requiring the “reference sequence” to be “not identical to the preselected rearrangement” does not place any meaningful limitation on the alignment method. Moreover, “alignment criterion” (*e.g.*, level of mismatch or gaps) is a feature of most, if not all, alignment programs and is used to determine if a read properly aligns to a reference sequence. *See also* EX1002 ¶215.

To the extent that claim 47 can be read as limiting claim 44, Li discloses a method for alignment of subgenomic intervals comprising genomic rearrangements with a reference sequence. EX1043 at 1754. Li discloses that a subgenomic interval comprising a genomic rearrangement can be aligned to a reference genome that is not identical to the subgenomic interval comprising the genomic rearrangement. EX1043 at 1754. Li discloses that alignment criteria may be used to analyze an alignment of a subgenomic interval with a reference genome. Specifically, Li discloses that reads containing indels may be aligned to a reference genome by introducing gaps into the alignment. EX1043 at 1754, 1759; EX1045 at 494. Li discloses that the BWA alignment method can align reads to a reference genome depending on the number of differences between the read and the reference and that gaps and mismatches are both considered differences. EX1043 at 1757. In other words, Li discloses use of alignment criteria for alignment of reads to a reference genome. *See also* EX1002 ¶215.

Accordingly, claim 47 would have been obvious over Levin, Shah, Malhis, and Li. EX1002 ¶216.

iii. Dependent Claim 57

“The method of claim 47, wherein the preselected rearrangement is a preselected indel.”

Indels are a species of genomic rearrangements that Li specifically discloses may be identified from alignments produced by BWA. EX1043 at 1757.

Accordingly, claim 57 would have been obvious over Levin, Shah, Malhis, and Li.

See also EX1002 ¶¶217, 218.

VII. CONCLUSION

For the reasons set forth above, claims 1, 15, 17-20, 32, 44-47, and 57 of the '830 patent are unpatentable. Petitioners therefore request that an *inter partes* review of these claims be instituted.

Respectfully submitted,

Dated: May 17, 2017

/ Michael T. Rosato /

Michael T. Rosato, Lead Counsel

Reg. No. 52,182

VIII. CERTIFICATE OF COMPLIANCE

Pursuant to 37 C.F.R. §42.24(d), the undersigned certifies that this Petition complies with the type-volume limitation of 37 C.F.R. §42.24(a). The word count application of the word processing program used to prepare this Petition indicates that the Petition contains 13,973 words, excluding the parts of the brief exempted by 37 C.F.R. §42.24(a).

Respectfully submitted,

Dated: May 17, 2017

/ Michael T. Rosato /
Michael T. Rosato, Lead Counsel
Reg. No. 52,182

IX. PAYMENT OF FEES UNDER 37 C.F.R. §§ 42.15(A) AND 42.103

The required fees are submitted herewith. If any additional fees are due at any time during this proceeding, the Office is authorized to charge such fees to Deposit Account No. 23-2415.

X. APPENDIX – LIST OF EXHIBITS

Exhibit No.	Description
1001	U.S. Patent No. 9,340,830 to Lipson et al.
1002	Declaration of Dr. John Quackenbush
1003	<i>Curriculum Vitae</i> of Dr. John Quackenbush
1004	File history of U.S. Pat. Appl. No. 13/339,986
1005	Sohrab P. Shah, et al., “Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution” <i>Nature</i> 461: 809-813(Oct. 2009)
1006	Sohrab P. Shah, et al., “Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution” <i>Nature</i> 461: 809-813 (Oct. 2009), Supplementary Information File 1, obtained from http://www.nature.com/nature/journal/v461/n7265/supinfo/nature08489.html
1007	Sohrab P. Shah, et al., “Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution” <i>Nature</i> 461: 809-813 (Oct. 2009), Supplementary Information File 2, obtained from http://www.nature.com/nature/journal/v461/n7265/supinfo/nature08489.html
1008	Sohrab P. Shah, et al., “Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution” <i>Nature</i> 461: 809-813 (Oct. 2009), Supplementary Information File 3, obtained from http://www.nature.com/nature/journal/v461/n7265/supinfo/nature08489.html
1009	Sohrab P. Shah, et al., “Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution” <i>Nature</i> 461: 809-813 (Oct. 2009), Supplementary Information File 4, obtained from http://www.nature.com/nature/journal/v461/n7265/supinfo/nature08489.html

1010	Joshua Z. Levin, et al., “Targeted next generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts” <i>Genome Biology</i> 10(10):R115-R115.8 (Oct. 2009)
1011	Joshua Z. Levin, et al., “Targeted next generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts” <i>Genome Biology</i> 10(10):R115-R115.8 (Oct. 2009), Supplementary Information File 1, obtained from https://genomebiology.biomedcentral.com/articles/10.1186/gb-2009-10-10-r115#Sec17
1012	Joshua Z. Levin, et al., “Targeted next generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts” <i>Genome Biology</i> 10(10):R115-R115.8 (Oct. 2009), Supplementary Information File 2, obtained from https://genomebiology.biomedcentral.com/articles/10.1186/gb-2009-10-10-r115#Sec17
1013	Joshua Z. Levin, et al., “Targeted next generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts” <i>Genome Biology</i> 10(10):R115-R115.8 (Oct. 2009), Supplementary Information File 3, obtained from https://genomebiology.biomedcentral.com/articles/10.1186/gb-2009-10-10-r115#Sec17
1014	Joshua Z. Levin, et al., “Targeted next generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts” <i>Genome Biology</i> 10(10):R115-R115.8 (Oct. 2009), Supplementary Information File 4, obtained from https://genomebiology.biomedcentral.com/articles/10.1186/gb-2009-10-10-r115#Sec17
1015	Joshua Z. Levin, et al., “Targeted next generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts” <i>Genome Biology</i> 10(10):R115-R115.8 (Oct. 2009), Supplementary Information File 5, obtained from https://genomebiology.biomedcentral.com/articles/10.1186/gb-2009-10-10-r115#Sec17

1016	Ryan Tewhey, et al., “Enrichment of sequencing targets from the human genome by solution hybridization” <i>Genome Biology</i> 10(10):R116-R116.13 (Oct. 2009)
1017	Timothy J. Ley, et al., “DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome” <i>Nature</i> 456:66-72 (Nov. 2008)
1018	Timothy J. Ley, et al., “DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome” <i>Nature</i> 456:66-72 (Nov. 2008), Supplementary Information File 1, obtained from http://www.nature.com/nature/journal/v456/n7218/supinfo/nature07485.html
1019	David W. Craig, et al., “Identification of genetic variants using bar-coded multiplexed sequencing” <i>Nature</i> 5(10):887-893 (Oct. 2008)
1020	David W. Craig, et al., “Identification of genetic variants using bar-coded multiplexed sequencing” <i>Nature</i> 5(10):887-893 (Oct. 2008) Supplementary File, obtained from http://www.nature.com/nmeth/journal/v5/n10/supinfo/nmeth.1251_S1.html
1021	Simon A. Forbes, et al., “The Catalogue of Somatic Mutations in Cancer (COSMIC)” <i>Current Protocols in Human Genetics</i> , 10.11.1-10.11.26 (Mar. 2008)
1022	Simon A. Forbes, et al., “COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer” <i>Nucleic Acids Research</i> 39:D945-D950 (Oct. 2010)
1023	Andreas Gnirke, et al., “Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing” <i>Nature Biotechnology</i> 27(2):182-189 (Feb. 2009)
1024	U.S. Patent Publication 2010/0029498 (Gnirke)
1025	Ken Garber, “Fixing the front end” <i>Nature Biotechnology</i> 26(10):1101-1104 (Oct. 2008)
1026	Philip J. Stephens, et al., “Complex landscapes of somatic rearrangement in human breast cancer genomes” <i>Nature</i> 462:1005-1010 (Dec. 2009)
1027	Michael R. Stratton, et al., “The Cancer Genome” <i>Nature</i> 458:719-724 (Apr. 2009)
1028	Michael L. Metzker, “Sequencing technologies—the next generation” <i>Nature Reviews: Genetics</i> 11:31-46 (Jan. 2010)

1029	Moran Yassour, et al., “Ab initio construction of a eukaryotic transcriptome by massively parallel mRNA sequencing” <i>PNAS</i> 106(9):3264-3269 (Mar. 2009)
1030	Lira Mamanova, et al., “Target-enrichment strategies for next generation sequencing” <i>Nature Methods</i> 7(2):111-118 (Feb. 2010)
1031	Agilent Protocol “SureSelect Target Enrichment System Illumina Single-End Sequencing Platform Library Prep” Agilent Technologies, Version 1.2 (Apr. 2009)
1032	Elaine M. Kenny, et al., “Multiplex Target Enrichment Using DNA Indexing for Ultra-High Throughput SNP Detection” <i>DNA Research</i> 18(1):31-38 (Dec. 16, 2010)
1033	Nik Cummings, et al., “Combining target enrichment with barcode multiplexing for high throughput SNP discovery” <i>BMC Genomics</i> 11:641 (Nov. 18, 2010)
1034	Elaine R. Mardis, et al., “Recurring Mutations Found by Sequencing an Acute Myeloid Leukemia Genome” <i>New England Journal of Medicine</i> 361:1058-1066 (Sept. 2009)
1035	Peter J. Campbell, et al., “Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing” <i>PNAS</i> 105(35):13081-13086 (Sept. 2008)
1036	Juliann Chmielecki, et al., “Targeted next-generation sequencing of DNA regions proximal to a conserved GXGXXG signaling motif enables systematic discovery of tyrosine kinase fusions in cancer” <i>Nucleic Acids Research</i> 38(20):6985-6996 (June 29, 2010)
1037	S. Marsh, et al., “Pharmacogenetic analysis of paclitaxel transport and metabolism genes in breast cancer” <i>The Pharmacogenomics Journal</i> 7:362-365 (Jan. 2007)
1038	Complaint for Patent Infringement, Foundation Medicine, Inc. v. Guardant Health, Inc., 2:16-cv-00523 (May 17, 2016)
1039	Declaration of Dr. Patrick Medley
1040	Transcript of Deposition of Dr. Stacey Gabriel in 2 :16-cv-00523-JRG-RSP (Mar. 24, 2017)
1041	Rebuttal Declaration of Dr. Stacey Gabriel in 2 :16-cv-00523-JRG-RSP (Mar. 17, 2017)
1042	Nawar Malhis and Steven J.M. Jones, “High quality SNP calling using Illumina data at shallow coverage” <i>Bioinformatics</i> , Vol. 26(8): 1029–1035 (February 26, 2010)
1043	Heng Li and Richard Durbin, “Fast and accurate short read alignment with Burrows–Wheeler transform” <i>Bioinformatics</i> , Vol. 25(14): 1754–

	1760 (May 18, 2009)
1044	U.S Patent Application Publication No. US 2008/0131887 A1 to Dietrich Stephan and Melissa Floren Filippone (June 2008)
1045	Daniel Koboldt, et al., “Challenges of sequencing human genomes” <i>Briefings in Bioinformatics</i> , Vol. 11(5):484-498 (June 2, 2010)
1046	Slider II, “High quality SNP calling using Illumina data at minimal coverage,” Canada's Michael Smith Genome Sciences Centre (June 9, 2009) obtained from http://www.bcgsc.ca/platform/bioinfo/software/SliderII
1047	Heng Li, “NGS alignment programs,” (Nov. 19, 2009) obtained from http://lh3lh3.users.sourceforge.net/NGSalign.shtml
1048	Lydia Roy, et al., “Survival advantage from imatinib compared with the combination interferon-plus cytarabine in chronic-phase chronic myelogenous leukemia: historical comparison between two phase 3 trials” <i>Blood</i> , Vol. 108(5):1478-1484 (Sept. 2006)

CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. §§ 42.6(e) and 42.105(a), this is to certify that I caused to be served a true and correct copy of the foregoing Petition for Inter Partes Review of U.S. Patent No. 9,340,830 (and accompanying Exhibits 1001 through 1048) by overnight courier (Federal Express or UPS), on this 17th day of May, 2017, on the Patent Owner at the correspondence address of the Patent Owner as follows:

LANDO & ANASTASI, LLP
PLM – Louis Myers
One Main Street
Cambridge, MA 02142

Foundation Medicine, Inc.
One Kendall Square, Suite B6501
Cambridge, MA 02139

Respectfully submitted,

Dated: May 17, 2017

/ Michael T. Rosato /
Michael T. Rosato, Lead Counsel
Reg. No. 52,182