

UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE PATENT TRIAL AND APPEAL BOARD

Foresight Diagnostics Inc.,
Petitioner,

v.

Personalis, Inc.,
Patent Owner.

IPR2023-00224
U.S. Patent No. 11,384,394

**DECLARATION OF JOHN QUACKENBUSH, PH.D. IN SUPPORT OF
PETITION FOR *INTER PARTES* REVIEW**

Mail Stop: PATENT BOARD
Patent Trial and Appeal Board
U.S. Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

TABLE OF CONTENTS

	<u>Page</u>
I. Introduction	1
II. Legal Standards	7
A. Claim Construction	7
B. Obviousness	8
III. The '394 Patent.....	12
IV. Person of Ordinary Skill in the Art	13
V. Overview of a Person of Ordinary Skill in the Art's Knowledge	14
A. Overview of Molecular Biology and Genetics	14
1. The Central Dogma of Molecular Biology.....	15
2. Genetics.....	18
B. Genetic Mutations	20
C. Nucleic Acid Sequencing Technologies	23
1. Sequencing Techniques	23
2. Whole-Genome Sequencing	28
3. Amplification and Enrichment	30
D. Personalized Cancer Testing and Monitoring.....	39
E. Related Applications of Next-Generation Sequencing	43
F. Providing Actionable Information to Healthcare Providers.....	46
VI. Prior Art References of the Grounds.....	49
A. Leary (Ex. 1002)	49

TABLE OF CONTENTS
Cont'd

	<u>Page</u>
B. Ley (Ex. 1012)	50
C. Chan (Ex. 1008)	50
D. Liao (Ex. 1009)	51
VII. Claim Construction.....	51
A. “generating a first subset of nucleic acid molecules”	53
B. “capture probe”	55
C. “polymorphism”	59
D. “biomedical report”	60
VIII. The Claims of the ’394 Patent Would Have Been Obvious.....	61
A. Ground 1: Leary in view of a POSA’s Knowledge	61
1. Claim 1	65
2. Claim 2	88
3. Claim 3	89
4. Claim 4	90
5. Claim 5	92
6. Claim 6	92
7. Claim 7	93
8. Claim 8	94
9. Claim 9	94
10. Claim 10	95

TABLE OF CONTENTS
Cont'd

	<u>Page</u>
11. Claim 11	95
12. Claim 12	96
13. Claim 13	96
14. Claim 14	97
15. Claim 15	97
16. Claim 16	98
17. Claim 17	98
18. Claim 18	98
19. Claim 19	100
B. Ground 2: Leary (Ex. 1002) in view of Ley (Ex. 1012) and the knowledge of a POSA.	101
1. Claim 1	101
2. Claims 4-5	105
C. Ground 3: Leary (Ex. 1002) in view of Chan (Ex. 1008), Liao (Ex. 1009) and the knowledge of a POSA.....	106
1. Claim 1	110
2. Claims 2-11, 13-19	124
3. Claim 3	124
4. Claim 12	125
IX. Conclusions	126

Exhibits Cited

Ex. 1001	U.S. Patent No. 11,384,394 to Bartha <i>et al.</i>
Ex. 1002	Rebecca J. Leary <i>et al.</i> , “Development of Personalized Tumor Biomarkers Using Massively Parallel Sequencing,” <i>Science Translational Medicine</i> (2010) (Leary)
Ex. 1003	Quake <i>et al.</i> , WO 2011/057061 (Quake)
Ex. 1004	Matthew W. Anderson & Iris Schrijver, “Next Generation DNA Sequencing and the Future of Genomic Medicine,” <i>Genes</i> (2010) (Anderson)
Ex. 1005	Andreas Gnirke <i>et al.</i> , “Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing,” <i>Nature Biotechnology</i> (2009) (Gnirke)
Ex. 1006	Florian Mertes <i>et al.</i> , “Targeted enrichment of genomic DNA regions for next-generation sequencing, Briefings in Functional Genomics,” <i>Briefings in Functional Genomics</i> (2011) (Mertes)
Ex. 1007	Lira Mamanova <i>et al.</i> , “Target-enrichment strategies for next-generation sequencing,” <i>Nature</i> (2010) (Mamanova)
Ex. 1008	K.C. Allen Chan <i>et al.</i> , “Cancer Genome Scanning in Plasma: Detection of Tumor-Associated Copy Number Aberrations, Single-Nucleotide Variants, and Tumoral Heterogeneity by Massively Parallel Sequencing,” <i>Clinical Chemistry</i> (2013) (Chan)
Ex. 1009	Gary J.W. Liao <i>et al.</i> , “Targeted Massively Parallel Sequencing of Maternal Plasma DNA Permits Efficient and Unbiased Detection of Fetal Alleles,” <i>Clinical Chemistry</i> (2011) (Liao)
Ex. 1010	<i>Personalis v. Foresight</i> , First Amended Complaint, D. Colo., Case No. 1:22-v-01913, Dkt. 9, filed August, 17, 2022 (“FAC”)
Ex. 1011	Infringement Claim Chart for U.S. Patent No. 11,384,394; Ex. I to Patent Owner’s FAC

Ex. 1012	Timothy J. Ley <i>et al.</i> , “DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome,” <i>Nature</i> (2008) (Ley)
Ex. 1013	Jay Shendure & Hanlee Ji, “Next-generation DNA sequencing,” <i>Nature Biotechnology</i> (2008) (Shendure)
Ex. 1014	Jeffrey S. Ross and Maureen Cronin, “Whole Cancer Genome Sequencing by Next-Generation Methods,” <i>American Journal of Clinical Pathology</i> (2011) (Ross)
Ex. 1015	David J. McBride, et al., “Use of Cancer-Specific Genomic Rearrangements to Quantify Disease Burden in Plasma from Patients with Solid Tumors,” <i>Genes, Chromosomes & Cancer</i> (2010) (McBride)
Ex. 1016	David M. Kurtz, et al., “Enhanced detection of minimal residual disease by targeted sequencing of phased variants in circulating tumor DNA,” <i>Nature Biotechnology</i> (2021) (Kurtz) (Ex. J to FAC)
Ex. 1017	https://www.aacc.org/science-and-research/clinical-chemistry/clinical-chemistry-podcasts/2013/cancer-genome-scanning-in-plasma-detection-of-tumor-associated-copy-number-aberrations
Ex. 1018	NIH Spreadsheet Regarding DNA Sequencing Costs
Ex. 1019	https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data
Ex. 1023	<i>Curriculum Vitae</i> of John Quackenbush
Ex. 1024	https://www.garvan.org.au/research/kinghorn-centre-for-clinical-genomics/learn-about-genomics/for-gp/genetics-refresher-1/types-of-variants
Ex. 1025	Seo Young Oh, et al., “Improved DNA Extraction Method for Molecular Diagnosis from Smaller numbers of Cells,” <i>Korean Journal of Clinical Laboratory Science</i> (2014) (Oh)
Ex. 1026	Leary Supplemental Information

Ex. 1027	Illumina, Technical Note: Sequencing
Ex. 1028	Illumina, Specification Sheet: Illumina Sequencing
Ex. 1029	Hugo Lam, et al., “Performance comparison of whole-genome sequencing platforms,” <i>Nature Biotechnology</i> (2012) (Lam)
Ex. 1030	Tim Forshew, et al., “Noninvasive Identification and Monitoring of Cancer Mutations by Targeted Deep Sequencing of Plasma DNA,” <i>Science Translational Medicine</i> (2012) (Forshew)
Ex. 1031	Matthew Meyerson, et al., “Advances in understanding cancer genomes through second-generation sequencing,” <i>Nature Reviews</i> (2010) (Meyerson)

I. Introduction

1. My name is John Quackenbush. I have been retained by Foresight Diagnostics Inc. (“Petitioner”) to provide my independent evaluation and opinions regarding U.S Patent No. 11,384,394 and related prior art.

2. For this service, I am paid my standard hourly consulting fee, and I am reimbursed for actual expenses. My compensation in no way depends on the outcome of this matter or the substance of this declaration.

3. In my testimony, I have relied upon my technical background, skills, knowledge, training, education and experience in the relevant art, and my examination of the materials utilized in preparing this declaration.

4. In providing this testimony, I have relied on the materials cited herein, as well as my knowledge and experience working in the field before the priority date of the '394 patent, from the perspective of a person of ordinary skill in the art as of the priority date.

5. I am a Henry Pickering Walcott Professor of Computational Biology and Bioinformatics and Chair of the Department of Biostatistics at the Harvard T.H. Chan School of Public Health. I am also Founding Director and co-Director of the MS Program in Computational Biology and Quantitative Genetics and Director of the Big Data to Knowledge Training Program, co-Director of the Cancer Genomics Training Program, and the Faculty Director for the Harvard Quantitative Biomedical

Research Center. In addition, I hold appointments as Professor of Biostatistics and Computational Biology at the Dana-Farber Cancer Institute and at the Channing Division of Network Medicine of Brigham and Women's Hospital.

6. In addition to my academic appointments, I co-founded Genospace LLC in 2012. Genospace produces software solutions to operationalize precision medicine through the integration of clinical and genomic data and its presentation to clinical and research users in an intuitive format. Genospace was purchased by Hospital Corporation of America in January of 2017.

7. In addition to my academic and research responsibilities, over the years I have taken significant regional and national research leadership responsibilities. I have served on the Scientific Advisory Board of the St. Jude's Children's Research Hospital and the Board of Scientific Counselors of the Environmental Protection Agency. I have also served on numerous scientific review panels for the National Institutes of Health, the National Science Foundation, the Department of Energy, the National Aeronautics and Space Administration, the Wellcome Trust, Genome British Columbia, Genome Canada, the National Sciences and Engineering Council of Canada, and the European Union. And I have served on numerous study panels for the National Research Council, the National Academy of Sciences, and the Institute of Medicine.

8. I have presented more than 350 talks, courses, and workshops since joining the Harvard School of Public Health and Dana-Farber in 2005. I have also taught numerous courses and workshops on the analysis of genomic data and served as an organizer of the EMBO course on DNA microarray analysis for more than 5 years.

9. Prior to joining the faculty at Dana-Farber and Harvard, I was a Professor of Chemical Engineering at the University of Maryland from 2003 – 2005; an Investigator with The Institute for Genomic Research from 2002 – 2006; a Professor in the Department of Biochemistry at The George Washington University from 2000 – 2005; a Lecturer in The Department of Biostatistics at The Johns Hopkins University from 1998 – 2005; an Associate Investigator at The Institute for Genomic Research from 1997 – 2001; and an Assistant Investigator at The Institute for Genomic Research in 1997.

10. I received my Ph.D. in Theoretical Particle Physics at the University of California, Los Angeles in 1990. I also hold a Master of Science in Physics from University of California, Los Angeles and a Bachelor of Science in Physics from the California Institute of Technology.

11. I have studied, taught, practiced, and conducted research involving the use of genomic technologies, including DNA sequencing and DNA microarrays, and involving computational analysis of data related to genomic technologies. I have

expertise in genomic technologies, computational analysis, and modeling of a plant, animal, and microbial systems with an emphasis on the study of human diseases, such as cancer.

12. In describing my work, I often use terms to refer to techniques and technologies, as well as specific application areas. During the course of my career, I have worked in genomics, functional genomics, epigenomics, bioinformatics, biostatistics, computational biology, systems biology, network biology, and network medicine. Although the boundaries between these various fields are often not well defined, one possible distinction is that some are directed toward data generation while others are directed toward developing and applying methods for data analysis and interpretation. For example, genomics, functional genomics, and epigenomics are generally associated with the design and conduct of experiments for generation of large-scale data on hundreds, thousands, or tens of thousands of genes, RNA transcripts, or epigenetic sites within the genome (or other alterations such as identifying regulatory interactions or structural changes in the genome). Fields including bioinformatics, biostatistics, computational biology, systems biology, network biology, and network medicine generally involve the development of new methods for data analysis and application of those methods to the biological and biomedical data sets, including those generated using techniques such as genomics, functional genomics, and epigenomics. The beauty of all of these techniques is that

their development and use is, largely although not entirely, independent of their application area. So methods developed on one area are often used in other areas.

13. I have worked in the analysis of genes in hundreds of animal, plant, unicellular, and parasitic species, including human, mouse, rat, cattle, sheep, brown bear, salmon, corn, tomato, rice, yeast, the causative agent of malaria, and countless bacteria and viruses. Despite this diversity of applications, my primary research emphasis has been in the study of human diseases, with the greatest body of my recent work in the analysis of human cancers and pulmonary diseases. However, I have also explored a wide range of other issues in health and disease, including Alzheimer's disease, Parkinson's disease, Huntington's disease, neonatal sepsis, and Down syndrome. My work has involved analysis of data generated from bulk tissue samples, small numbers of cells, single cells, and cell-free DNA.

14. I have also done work relevant to reproductive health, an area that became personally important to me as my wife and I, then in our forties, became enmeshed in the use of assisted reproductive technology. This gave me first-hand exposure to a wide range of techniques, including pre-implantation prenatal screening and the use and limitations of technologies directed toward prenatal detection of genetic abnormalities including aneuploidies. Following the birth of my son in 2006 at Brigham and Women's Hospital in Boston, I also worked with, and mentored, physician scientists at Harvard, published a genomic analysis of

diminished ovarian reserve, and presented the President's plenary lecture at the 63rd American Society for Reproductive Medicine Annual Meeting in 2007 in which I described the use of genomic technology in reproductive, prenatal, and neonatal health.

15. I have received numerous awards and honors including being recognized as a White House Open Science Champion of Change, an Australia Fellow, and a Distinguished International Advisor for the International Conference of BioInformatics and BioEngineering sponsored by the Institute of Electrical and Electronics Engineers. In 2022 I was elected to the National Academy of Medicine, one of three academies that make up the National Academies of Sciences, Engineering, and Medicine (the National Academies) in the United States. I have published more than 300 peer reviewed scientific papers that have been cited more than 79,000 times. I have had continuous grant support for my research from the National Institutes of Health and other organizations for nearly thirty years. In 2018, I received an NCI Outstanding Investigator Career Award, a seven-year research grant recognizing a lifetime record of substantial contributions to the field.

16. A copy of my Curriculum Vitae, attached as Exhibit 1023, contains further details on my education, experience, publications, and other qualifications to render an expert opinion in this matter.

II. Legal Standards

17. I am not a lawyer. However, I have been informed by counsel of general principles concerning patent law and the analysis of whether a patent claim is valid. I have followed these principles in arriving at my conclusions in this declaration. The sections below set forth my understanding of the legal principles that are relevant to my opinions in this declaration.

A. Claim Construction

18. I understand that the claims of a patent define the scope of the inventions that are protected by the patent. I understand that the language of the claims is construed as it would be understood by a hypothetical person having ordinary skill in the art (a “POSA”) at the time of the invention. Accordingly, where I refer below to a POSA—for example, where I discuss how such a person would have understood the specification or the claims, or what a POSA might have known, among other things—I am referring to such a person working at the time of the invention, even if for the sake of brevity I omit that specific timeframe from my statements.

19. I understand words in a claim will be given their ordinary or customary meaning unless it appears that the inventor used them differently.

20. I also understand that, in construing patent claims, courts seek to determine what a person skilled in the art would have understood the claims to mean primarily in light of the intrinsic evidence of record, which includes the language of the claim itself and other issued claims, the specification, the art cited during prosecution, and the prosecution history.

21. As discussed below in Section VII (“Claim Construction”), where relevant, I have applied the implicit interpretation of the claim language that Patent Owner applied in its complaint for infringement against Petitioner. Exs. 1010, 1011.

B. Obviousness

22. First, I understand that, for an invention claimed in a patent to be found patentable, or to be valid, it must be, among other things, new and not obvious in light of what came before it. That which came before is generally referred to as “prior art.”

23. I understand that a patent claim is unpatentable or may be found invalid as obvious if, at the time when the invention was made, the subject matter of the claim, considered as a whole, would have been obvious to a person having ordinary skill in the field of the technology (the “art”) to which the claimed subject matter belongs.

24. I understand that the obviousness standard is defined in the patent statute as follows:

A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 [in a single prior art reference], if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

35 U.S.C. § 103(a).

25. I understand that a claim in a patent is obvious when the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which the subject matter pertains.

26. I understand that obviousness may be shown by considering more than one item of prior art, and that the following factors should be considered in analyzing obviousness: (1) the scope and content of the prior art; (2) the differences between the prior art and the claims; and (3) the level of ordinary skill in the pertinent art.

27. I understand that a person of ordinary skill in the art is assumed to have knowledge of all relevant prior art. I understand that one skilled in the art can

combine the teachings of various prior art references based on the teachings of the individual prior art references, the general knowledge present in the art, or common sense. I understand that a teaching, motivation or suggestion appearing in the prior art may provide a basis for combining two (or more) prior art references or modifying a given prior art reference. A motivation to combine references may also be implicit in the prior art, and there is no requirement that there be an actual or explicit teaching to combine two references. Thus, one may take into account the inferences and creative steps that a person of ordinary skill in the art would employ to combine the known steps in the prior art in the manner claimed by the patent at issue.

28. I understand that the combination of familiar steps according to known methods is likely to be obvious when it does no more than yield predictable results. I understand that when a work is available in one field of endeavor, design incentives and other market forces can prompt variations of that work, either in the same field or a different one. Likewise, if a technique has been used to improve one device, and a person of ordinary skill in the art would recognize that it would improve similar devices in the same way, using that technique to improve the other device in a predictable manner would normally be obvious.

29. I understand that a person of ordinary skill is also a person of ordinary creativity in the field. A person of ordinary skill attempting to solve a problem will

not be restricted only to those steps of prior art designed to solve the same problem. Rather, in exercising common sense and ordinary creativity, a person of ordinary skill can draw upon prior art directed to any need or problem known in the field of endeavor at the time of the invention. In other words, the prior art does not need to be directed towards solving the same problem that is addressed in the patent. Further, the individual prior art references themselves need not all be directed towards solving the same problem.

30. I further understand that in many fields, it may be that there is little discussion of obvious techniques or combination, and it often may be the case that market demand, rather than scientific literature or knowledge, will drive design trends. When there is such a design need or market pressure to solve a problem and there are a finite number of identified, predictable solutions, a person of ordinary skill has good reason to pursue the known options within their technical grasp. If this leads to the anticipated success, it is likely the result not of innovation but of ordinary skill and common sense. In that instance, the fact that a combination was obvious to try might show that it was obvious. I understand that the fact that a particular combination of prior art steps was “obvious to try” may indicate that the combination was obvious even if no one had attempted the combination at the time of filing. Likewise, duplicating a known feature from the prior art multiple times in

order to take advantage of proven technology typically does not require a leap of inventiveness.

III. The '394 Patent

31. The '394 patent (Ex. 1001) is titled "Methods and Systems for Genetic Analysis." Ex. 1001, Code (54). The patent claims priority to "[p]rovisional application No. 61/753,828, filed on Jan. 17, 2013." *Id.*, Code (60). I have considered January 17, 2013 to be the relevant date for assessing prior art invalidity in my analysis.¹

32. In its complaint for patent infringement against Petitioner, Patent Owner characterized the subject matter claimed in the '394 patent as follows:

For example, claim 1 of the '394 Patent claims a method for analyzing nucleic acid samples isolated from an individual comprising whole genome sequencing nucleic acid molecules obtained from a first biological sample, contacting a second set of nucleic acid molecules with a set of capture probes that hybridize to polymorphisms that are based on or extracted from a database of polymorphisms or observed in a sample, or a combination thereof, sequencing the second set of nucleic acid molecules, repeating the capture probe hybridized sequencing on

¹ I am not aware of any evidence suggesting that Patent Owner is entitled to an earlier priority date. Even if Patent Owner were to establish entitlement to a priority date as early as January 17, 2012, one-year prior to the filing of the provisional application, my analysis would not materially change, and the '394 patent's claims would still be unpatentable at least as set forth in Grounds 1 and 2.

an additional sample from an additional time point, and generating a biomedical report that identifies the presence or absence of the polymorphisms.

Ex. 1010 (Personalis Complaint) at ¶30.

IV. Person of Ordinary Skill in the Art

33. I have been advised that a “person of ordinary skill in the art,” or “POSA,” is a hypothetical person and that I should consider the following factors in constructing a POSA: (1) the educational and experience level of active workers in the field; (2) the sophistication of the technology; (3) the rapidity with which innovations are made; (4) the types of problems encountered in the art; (5) the prior art solutions to those problems. I understand that not all such factors may be present in every case, and one or more of them may predominate.

34. The relevant technology field for the ’394 patent is nucleic acid sample preparation, sequencing, and analysis to detect genetic variation for clinical applications. By virtue of my education, experience, and training, I am familiar with the level of skill in the art of the ’394 patent before January 17, 2013.

35. After considering the factors above, I believe a person of ordinary skill in the relevant field before January 17, 2013 would have had an advanced degree (such as an M.D. or the equivalent of a Ph.D. in a life or physical sciences, engineering, or mathematics discipline), and through further education or

postdoctoral experience, an understanding of various sequencing and next generation sequencing methods, sample preparation methods, targeted enrichment methods, bioinformatics methods for interpreting sequencing data, and methods for identifying genetic variants in a sample. A POSA also would have had an understanding of clinical applications of sequencing technology to conditions associated with nucleic acid variation, such as cancer, non-invasive prenatal testing, and transplant rejection. Additionally, a POSA would have had an understanding of cancer biology as it relates to mutations associated with tumor diagnosis, progression, and treatment, as well as awareness of the various types of mutations that can occur in and influence the course of a patient's cancer.

36. I understand that a POSA is presumed to be aware of all relevant art.

V. Overview of a Person of Ordinary Skill in the Art's Knowledge

37. In this Section, I provide an overview of the knowledge that a POSA would have possessed as of January 17, 2013.

A. Overview of Molecular Biology and Genetics

38. I begin with an overview of the principles of biology that are relevant to understanding the state of the art, especially those pertaining to nucleic acid sequencing and cancer biology.

1. The Central Dogma of Molecular Biology

39. Genes contain the set of instructions, or blueprints, for the development and maintenance of the cells of all organisms. Genes are made of deoxyribonucleic acid (“DNA”) and are located on chromosomes in the nucleus of eukaryotic cells. DNA is composed of subunits called nucleotides, the sequences of which generate the code directing the eventual production of proteins.

40. DNA exists as a double-stranded molecule: it consists of two long strands in a double-helical structure. One end of each strand is called the 5-prime (“5’”) end, and the other is called the 3-prime (“3’”) end. The strands are bound together such that they run in opposite directions: one strand, oriented in the 5’ to 3’ direction, is attached to another strand, oriented in the 3’ to 5’ direction.

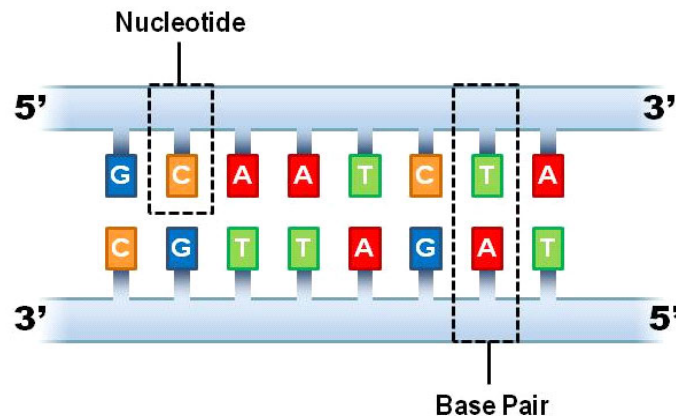


41. The smallest functional unit of DNA is the gene, which is a sequence of nucleotides that encode for ribonucleic acid (“RNA”), which in turn codes for a protein. Proteins are incredibly versatile large molecules that serve crucial functions in essentially all biological processes. As examples: some proteins, called enzymes,

serve as catalysts for the creation of other biological molecules; some transport and store other molecules, such as oxygen, throughout the body; some provide immune protection; some provide physical support and generate movement; and others transmit nerve impulses.

42. Attached at regular intervals along the backbone of each strand of DNA are “bases.” Each base, with its corresponding piece of DNA backbone, is referred to as a “nucleotide.” A strand of multiple nucleotides is called an “oligonucleotide” or “oligo.”

43. There are four natural types of bases that occur in DNA: adenine (“A”), cytosine (“C”), guanine (“G”) and thymine (“T”). Each type of base on one DNA strand binds to only one type of base on the other strand, known as its “complementary” base. When complementary bases bind together, or “hybridize,” they form what are known as “base pairs.” A hybridizes to T, and G hybridizes to C. It is the pairing, or hybridization, of the complementary bases that holds the two strands of DNA together, as depicted below.

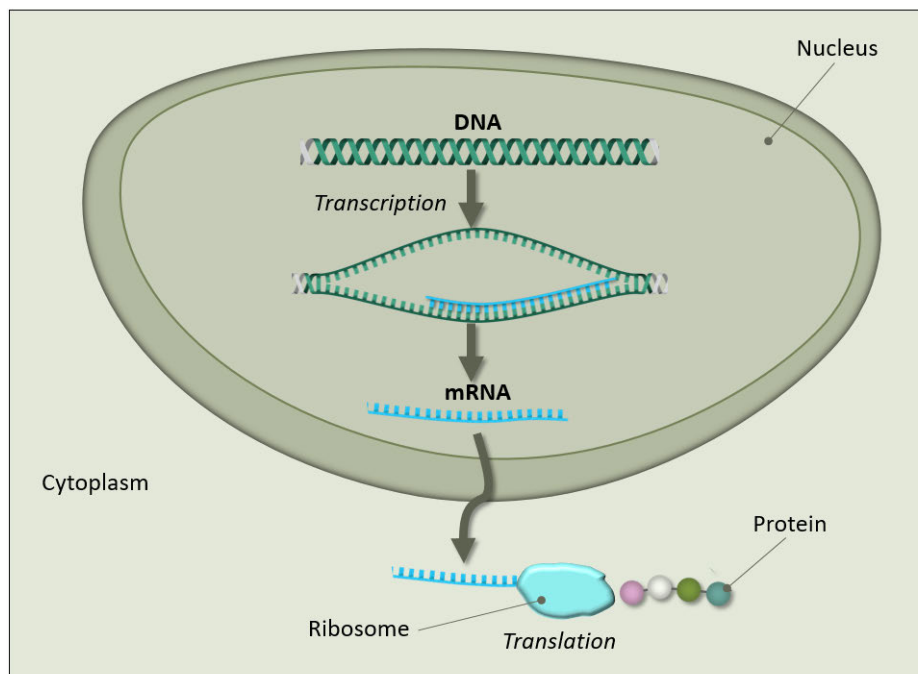


44. The A-T and G-C interactions between strands of oligonucleotides are themselves specific. However, it is possible for two strands to bind together even if not all of the bases in them are perfectly complementary. That is, as long as a sufficient number of A-T and G-C pairs can form between two oligonucleotides, they will “anneal” or “hybridize,” even if there are some mis-matches within the two strands.

45. The information encoded by the DNA in a gene is turned into a protein via two step-wise processes, called transcription and translation. In transcription, the DNA is used as a template by specialized machinery in the cell to generate another type of molecule, called messenger ribonucleic acid (mRNA, that contains a series of nucleotides “complementary” to those in the DNA-encoded gene). In translation, the mRNA is read by an organelle called the ribosome three nucleotides at a time. These three-nucleotide mRNA strings, called codons, instruct the addition of protein

building blocks called amino acids one-by-one by the ribosome's translational machinery.

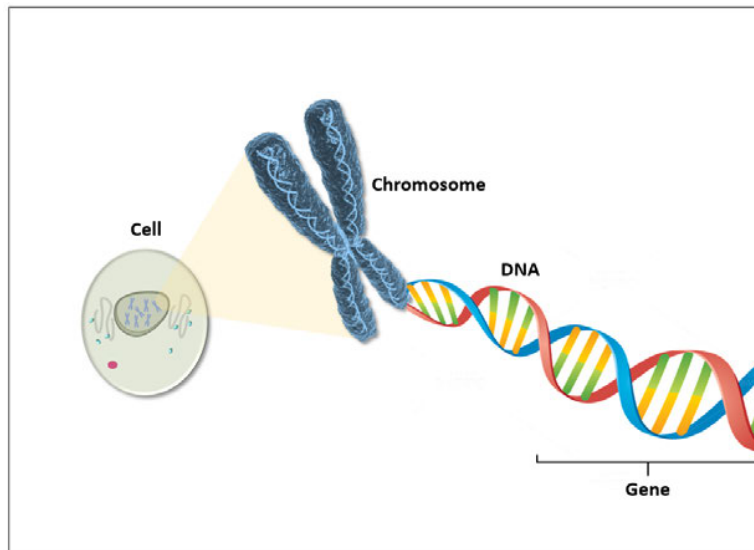
46. The flow of information from DNA to RNA to protein is referred to as “the central dogma of molecular biology.”



2. Genetics

47. The human genome consists of approximately 3.2 billion nucleotides of DNA, and (most) human cells contain two copies of the genome, one inherited from each biological parent. The DNA in the genome is divided and packaged among the chromosomes. Humans have 23 chromosome pairs: 1 pair of sex chromosomes (XX—female; XY—male), and 22 autosome pairs (the chromosomes

that are not sex chromosomes). Encoded within each copy of the genome are 25,000-30,000 genes, each located at a specific location on one of the chromosomes.



48. Although only about 1% of the overall DNA in humans is made up of protein-encoding genes, variation in the sequences of these genes is largely responsible for the differences between us. While a person's genes are important for proper biological function, no two people (apart from identical twins) have identical DNA sequences through the entirety of their genomes. Certain alterations, sometimes called "mutations," "variations," "variants," or "polymorphisms" can occur in all genes and portions of the human genome, contributing both to normal patterns of human variation as well as individual risk for developing diseases. Polymorphisms can include many types of genomic alteration, including insertions, deletions, point mutations, translocations, rearrangements, and copy number

variations. The entire collection of an individual's genetic variants is referred to as their "genotype."

49. A specific DNA sequence is said to be polymorphic if it varies between individuals, and the individual sequence variants are referred to as alleles. Each allele can have varying influences on the protein encoded by the gene. Because of redundancies in the genetic code, two slightly different alleles may produce exactly the same protein. Other alleles may cause a change in the protein encoded by a particular gene, altering the protein's function. Still others may include a mutation that instructs the cell's translational machinery to prematurely stop, resulting in the production of a truncated protein with complete loss-of-function. Other allelic variations may have resultant effects that fall somewhere in-between, with different amino acid sequences having varying levels of effect on protein function.

50. In the autosomes (the non-sex chromosomes), there are two alleles for each gene, one coming from each parent. When each of the alleles are the same, the individual is said to be homozygous at the allele, and when the alleles are different, the individual is said to be heterozygous at the allele.

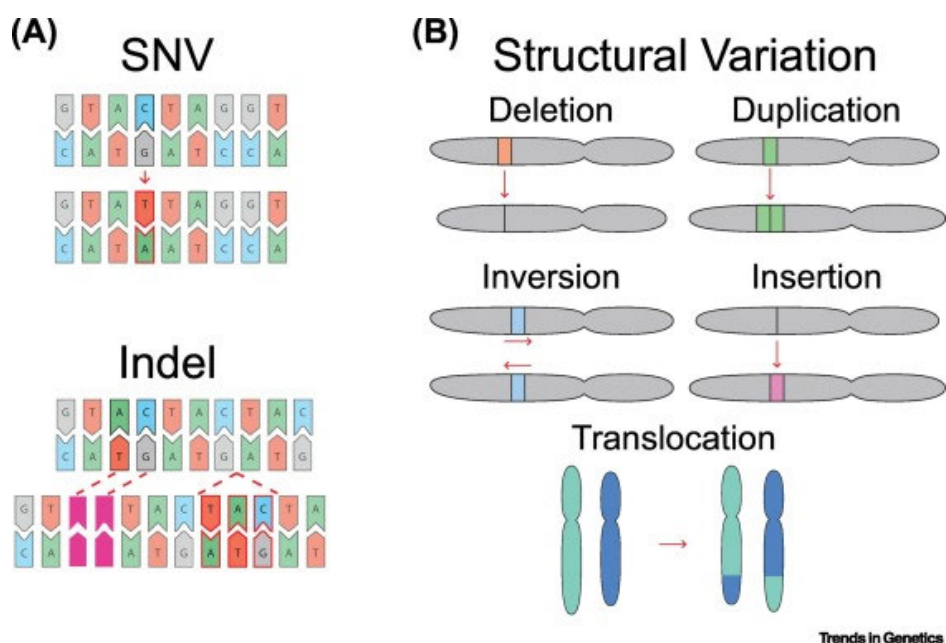
B. Genetic Mutations

51. Genetic mutations can be inherited (i.e., passed on from parents to their offspring) or acquired (i.e., developed for the first time in a new cell). Inherited mutations are called "germline" mutations. Acquired, or non-inherited, mutations

are called “somatic” mutations. These can result from chance cellular events, such as errors acquired during cell division, as well as from environmental and lifestyle factors that induce damage to the genome, such as exposure to carcinogens. Somatic mutations are not present in all of the cells in one’s body; they are only present in the progeny of the cell in which the mutation first occurred. Those skilled in the art appreciated that, unlike germline mutations that are inherited and present in every cell in the body, somatic tumor mutations arise during a person’s lifetime and occur only in the cells that have multiplied from a cell in which the somatic mutation has originated. Commonly, such mutations may first arise in cancerous (or tumor) cells, and thus propagate further as the cancer rapidly divides. Often, mutations that appear early in tumor development were likely to present in the tumor’s cells throughout the course of its development. Thus, somatic mutations unique to a particular tumor can be a useful tool for detecting, tracking, and monitoring a patient’s cancer.

52. Different types of somatic mutations were known in the prior art well before January 2013. *See* Ex. 1024 (<https://www.garvan.org.au/research/kinghorn-centre-for-clinical-genomics/learn-about-genomics/for-gp/genetics-refresher-1/types-of-variants>). These include single nucleotide polymorphisms (“SNPs”), copy number variants (“CNVs”), and insertions or deletions (“indels”). An SNP (sometimes called a single nucleotide variant (“SNV”)) is a change in one DNA base pair that may result in the substitution of one amino acid for another in the protein

that is encoded by the gene. The effect that an SNP has on the protein's function can vary widely, from having no effect at all to having a very significant effect. CNVs can result in producing too much or too little of a protein, which can have significant impacts on cell function. And indels can include the insertion of extra base pairs into a gene and/or the deletion of normal base pairs from the gene, which can also significantly affect the functioning of the proteins that the affected gene encode.



53. A POSA would have also been familiar with a type of polymorphism known as a translocation, which can move DNA from one chromosome to another chromosome (interchromosomal translocation) or to a different position on the same chromosome (intrachromosomal translocation). This type of polymorphism was known to be particularly relevant to a class of blood cancers called leukemias.

54. By 2013, there were also publicly accessible databases that a POSA would have been aware of, that could be searched for a large number of catalogued cancer-associated mutations. These included databases maintained by the International Cancer Genome Consortium (“ICGC”), The Cancer Genome Atlas (“TCGA”), International Agency for Research on Cancer (“IARC”), and the Catalogue of Somatic Mutations in Cancer (“COSMIC”).

C. Nucleic Acid Sequencing Technologies

1. Sequencing Techniques

55. Nucleic acid sequencing is the science of determining the order of the bases in a nucleic acid molecule on a base-by-base level. DNA sequencing requires a minimum amount of DNA. Thus, before it can be sequenced, a sample with too little DNA generally must be amplified.

56. One of the earliest known sequencing methods is called Sanger sequencing. Sanger sequencing is carried out by replicating a template strand of DNA, and incorporating labeled, chain-terminating bases at different locations along the different growing strands. The DNA fragments are then separated by size using electrophoresis and the various lengths of the fragments are used with the terminating bases to reconstruct the whole sequence of the template.

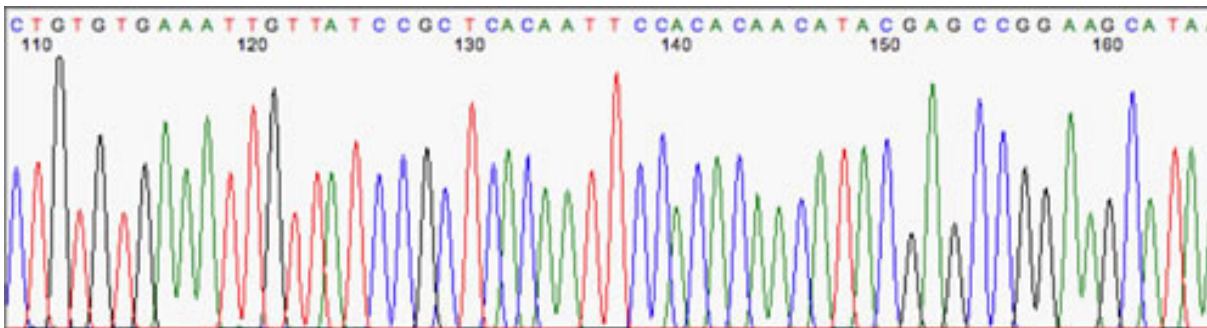
57. In Sanger sequencing, single-stranded DNA “templates” (*i.e.*, the DNA to be sequenced) are placed in solution containing a polymerase, oligonucleotide

primers (to start the reaction), and deoxynucleotides (for base pairing and primer extension). Sanger sequencing relies on the use of, in addition to regular deoxynucleotides, a small admixture of synthetic di-deoxynucleotides, which are labeled, for example, with fluorescent tags of different colors indicative of the particular base—A, G, T, or C—and serve a “stop” function as explained below.

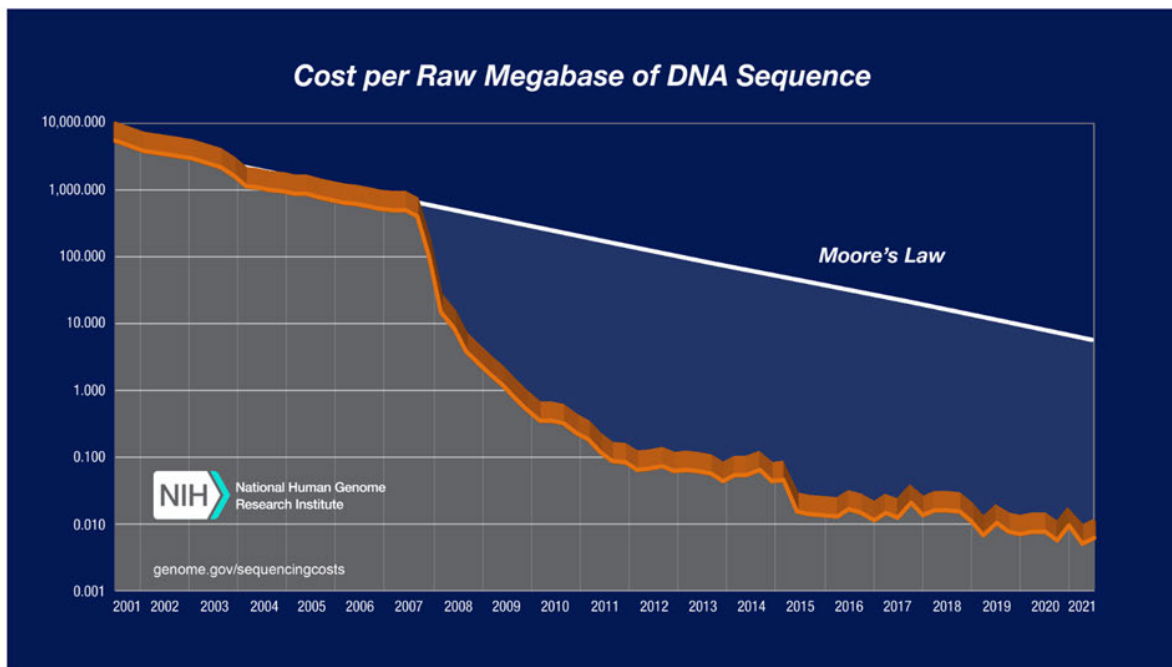
58. During strand extension in Sanger sequencing, when a normal deoxynucleotide is added to the growing strand, the next position remains “open” for another deoxynucleotide or di-deoxynucleotide to pair up to the template and elongate the copy strand. Once a di-deoxynucleotide is incorporated into the growing strand, it ends there, and no new nucleotides can be added on. Because the terminal di-deoxynucleotides are labeled, one can detect the last base that was added onto the growing strand. But if one has hundreds or thousands of copies of each strand, then during the sequencing process each copy of the same original DNA strand will see a di-deoxynucleotide incorporated randomly, leading to fluorescently labeled strands of different lengths.

59. All of the fluorescently labeled copies of different lengths are then separated based on their size by a process called gel electrophoresis. This creates a fluorescent “read” of the sequence from the shortest strands (those terminating immediately after the primer sequence) to the longest strands (those terminating after being extended all the way to the end of the template). A computer processes the

fluorescent data and produces a “trace file” that displays a representation of the fluorescent signal and the corresponding DNA sequence.



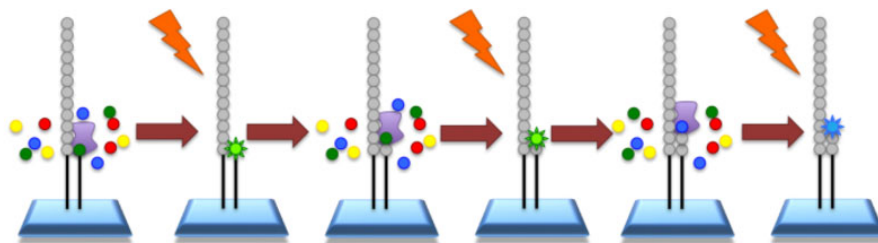
60. By 2005, a new approach to sequencing often referred to as “next generation sequencing” (“NGS”), sometimes also referred to as “massively parallel sequencing,” or “sequencing by synthesis” had been developed to allow “multiplex” sequencing of hundreds of thousands or millions of different DNA molecules simultaneously. By approximately 2008, techniques for NGS were well-known and commercially available. *See, e.g.,* Ex. 1013 (Shendure 2008). The cost of implementing NGS methods rapidly decreased, making such approaches more accessible and more desirable for further development. Ex. 1019 (<https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>).



Sequencing cost per megabase - 2021

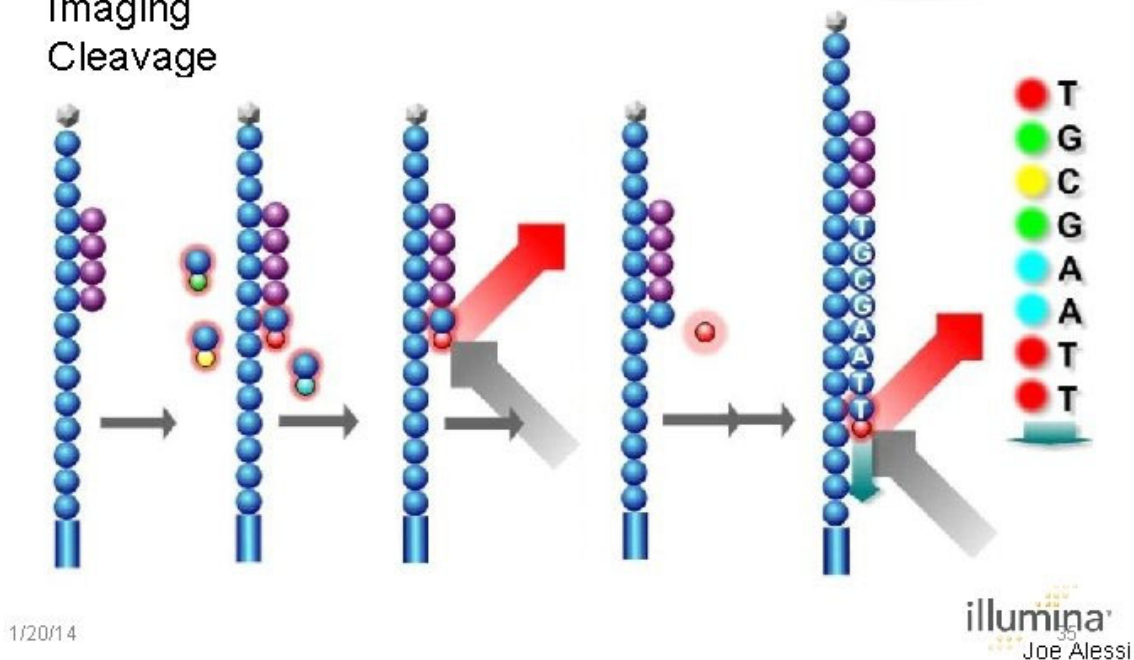
61. NGS is a high-throughput method for determining the base-by-base sequences of nucleic acid molecules, such as DNA. NGS is exponentially faster, cheaper, and more efficient than the Sanger sequencing technique that preceded it. For example, using NGS, the entire human genome can now be sequenced on a single instrument in a single day. In NGS, the DNA molecules being sequenced are first modified to attach what are often referred to as “sequencing adapters” to each end of the molecules. These modified DNA fragments are then rendered single-stranded and allowed to attach at low density to a “flow cell” in which the individual fragments can be subjected to “bridge amplification,” a PCR-like technique that creates molecular “clusters” that facilitate sequencing and detection. The DNA molecules are again rendered single-stranded so that their nucleotides can be used

as a template for DNA polymerase to make a copy. The polymerase builds a complementary strand in which single labeled nucleotides are added, illuminated with a laser and detected because they glow in a color corresponding to the base (A, G, T, or C) at that position, and then “deprotected” by releasing the fluorescent dye so the next labeled nucleotide can be added. As this process progresses and the complementary strand grows in length, the pattern of colors reveals the identity of the string of nucleotides, which is recorded by a computer that uses the series of detected bases to reconstruct the sequence. Hence the sequence is determined through the synthesis of a new (complementary) strand. This is shown in the figures below.



Sequencing By Synthesis (SBS)

Chemistry/incorporation
Imaging
Cleavage



62. A computer program can be used to align the sequence read back to a reference genome, or another sequence reference, which locates the sequence in the genome and thereby determines the identity (*i.e.* in which gene or on which chromosome) of the sequenced nucleic acid.

2. Whole-Genome Sequencing

63. The development of NGS made more widely available “[t]he ability to sequence entire human genomes.” Ex. 1027 (Lam 2012). To obtain highly accurate sequencing information of the whole genome, it is necessary to sequence each base

in the genome multiple times. Reasons for this include (i) multiple observations per base are required to have confidence that the base “call” is correct, so that errors that might occur in any one sequencing read can be rejected; (ii) so that mutation calls can be made with confidence based on multiple independent sequencing reads exhibiting the same polymorphism; and (iii) since DNA is randomly fragmented before sequencing and the amplification and library construction process may introduce some biases, sequencing reads are not evenly distributed throughout the genome, meaning that some bases will be read more than other bases. Ex. 1028 (Illumina Technical Note, “Estimating Sequencing Coverage”).

64. “Coverage,” “sequencing depth,” or “depth of sequencing” are terms that refer to the number of times a genome has been sequenced. It can be calculated by the formula:

$$C=L*(N/G)$$

Where C stands for coverage, G is the haploid genome length (~3 billion basepairs for the human genome), L is the read length, and N is the number of reads. By 2010, whole-genome sequencing to 30x coverage was routine on the commercially available Illumina HiSeq 2000 (I purchased one of these devices for us in my own lab in 2009 with funding from an NIH Shared Instrument Grant). Ex. 1029 (Illumina Specification Sheet, 2010) (“With unmatched cost-effectiveness, HiSeq 2000 is the first commercially available sequencer to enable researchers to obtain ~30x coverage

of two human genomes in a single run for under \$10,000 USD per sample.”; “The ultra-high output and speed of HiSeq 2000 now make it possible to sequence two human genomes at ~30x coverage simultaneously”). One run of the HiSeq 2000 allowed for a billion single-end reads or two billion paired-end reads (sequencing reads from both ends of a DNA fragment). *Id.* at 2. As each read was 100 bases in length, for 2 billion reads, the formula above yields 66.67x coverage for a single genome or 33.33x coverage of two genomes.

65. Whole-genome sequencing experiments that sought to ensure more complete coverage of the genome could do so by sequencing to greater depth. For example, in 2012, Lam reported performing whole-genome sequencing to an average coverage of ~76x using the HiSeq2000 and a different platform from Complete Genomics. Ex. 1027 at 78. Sequencing to even greater depth when required by an experiment also would have been routine, as it would require only running additional sequencing assays: “it is very easy to increase the coverage or sequence depth, if you later decide you need more data. Provided you still have your original sample, you can just sequence more, and combine the sequencing output from different flow cells.” Ex. 1028 at 2.

3. Amplification and Enrichment

66. Both Sanger sequencing and NGS require minimum quantities and/or relative abundances of target DNA. Thus, well before 2013 several techniques had

been developed to generate or isolate copies of DNA molecules of interest to facilitate sequencing.

67. DNA enrichment is the process of increasing the concentration of some DNA relative to other DNA in a mixture. This can be done using biochemical methods that involve physically pulling certain pieces of DNA out of a mixture for further analysis. Biochemical approaches to enrichment can involve using oligonucleotides with sequences complementary to the target DNA, which allows them to hybridize to other oligonucleotides that one seeks to analyze in a mixture. These oligonucleotides are referred to sometimes interchangeably as primers, probes, or baits.

68. Enrichment generally involves increasing the amounts of specific targets to generate sufficient quantities of a nucleic acid sequence to detect and analyze. *See, e.g.*, Ex. 1007 (Mamanova 2010) (“We have not yet reached a point at which routine sequencing of large numbers of whole eukaryotic genomes is feasible, and so it is often necessary to select genomic regions of interest and to enrich these regions before sequencing.”). I focus here on PCR and solution-based hybrid capture techniques as they are the most pertinent to the prior art references I have used in analyzing the grounds of unpatentability. However, a POSA would have been familiar with several other related techniques, such as the use of molecular inversion probes (“MIPs”), array-based systems, and solid-phase hybrid capture, among

others. *See id.* (describing PCR and solution-based hybrid capture as among “the most widely used approaches to target enrichment”); *see also id.* at Fig. 2 (reproduced below) (showing overlapping applicability of target-enrichment strategies).

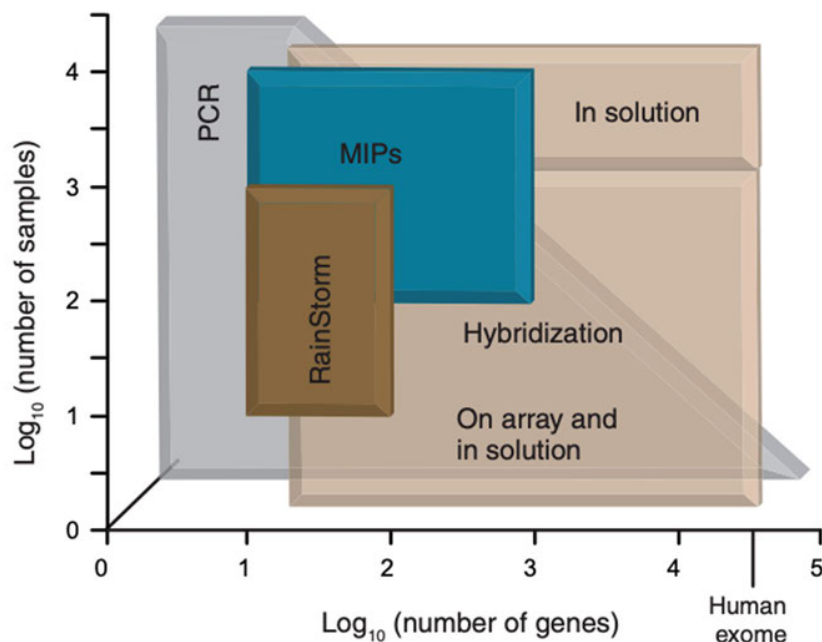


Figure 2 | Suitability of different target-enrichment strategies to different combinations of target size and sample number. Suitability was estimated from the perspective of the feasibility with which each method could be applied to the various combinations of target size and sample number, rather than the cost.

69. While these strategies have considerable overlap in their possible applications, a fair summary of the advantages of each technique is set forth in Mertes (Ex. 1006):

Overall, one can draw the following conclusions from a comparison of the currently used enrichment techniques . . . : (i) that hybrid capture

has its main advantages for medium to large target regions (10-50 Mb) in contrast to the other two approaches which typically only target small regions within the kilo base pairs and low mega base pairs range. The ability to enrich for mega base pair-sized targets is particularly advantageous in research studies where typically whole exomes or many genes are involved. Especially for clinical applications, this may be relevant for oncological applications where one would expect to sequence 100-1000's of genes. (ii) The advantage of PCR and circularization-based methods [i.e., MIPs] is that they achieve very high enrichment factors and few off-target reads, but only for small target regions. This is more suited to clinical genetics where typically only few critical loci need to be assessed.

Ex. 1006 (Mertes 2011) at 380. Thus, in assays where a POSA sought to analyze many sequences of interest, solution-based hybrid capture would have been an ideal enrichment strategy. On the other hand, in assays where a POSA intended to analyze relatively few sequences of interest, then a PCR approach would have been particularly well-suited. For intermediate numbers of potential targets, either technique might serve equally well. However, these parameters would not have been considered as strict limitations, and the techniques can and were used on a relatively interchangeable basis in many experimental protocols, and could also be used together as complements to one another. *See Ex. 1006* (Mertes 2011) (Table 1)

(describing pros, cons, and overlapping applications of “[c]urrently employed targeted enrichment techniques”).

Table 1: Currently employed targeted enrichment techniques

Enrichment technique	Vendor	Features	Pros	Cons	Number of loci (target size)	Library prep for NGS
Hybrid capture						
Solid support	Agilent SureSelect, Roche NimbleGen SeqCap EZ	Medium to large target regions, custom and preconfigured target regions (i.e. whole exome), Multiplexing possible, ready to use kits	Ease of production, large target sets	Large amount of input DNA, high-tech equipment (3–10 µg)	10^4 – 10^6 (1–50 Mb)	Before enrichment
In solution	Agilent SureSelect, FlexGen FlexSelect, Mycroarray MySelect, Roche NimbleGen SeqCap EZ		Ease of use, small amount of input DNA (<1–3 µg)			
Circularization						
Molecular inversion probes						
Selector probes	HaloGenomics	Custom kits and clinically relevant panel kits	No dedicated instruments, high specificity, input DNA (<1 µg)	exome kit not available yet	10^2 – 10^4 (0.1–5 Mb) 10^2 – 200 (0.1–1.5 Mb)	During enrichment (incorporated into hybridization probes)
PCR						
Long range	Invitrogen SequelPrep, Qiagen SeqTarget system	Smaller target regions, coverage by tiling	Relatively easy to set up and automatable, even coverage	PCR conditions largely influence effectiveness, >10 µg DNA for Large sets	10^2 – 10^4 (0.1–5 Mb)	After enrichment
Multiplex	Multiplicom, Fluidigm	Smaller target regions, coverage by tiling, multiplex PCR of 150–200 amplicons (150–450 bp)	Easy to perform, reasonably economical in terms of, even coverage			
Micro droplet	RainDance	Smaller target regions, coverage by tiling, micro droplet PCR of up to 20 000 amplicons (150–1500 bp)	Even coverage, low input DNA	Relatively expensive, specialist equipment	10^2 – 10^4 (up to 10 Mb)	

All major targeted enrichment techniques show relative pros and cons.

(a) PCR

70. In PCR, DNA molecules are replicated in a manner that takes advantage of the base pairing of complementary DNA strands.

71. In PCR, the DNA “template” strands (i.e., the strands being replicated) must first be separated. This is because DNA normally exists in a double-stranded form, but it must be single-stranded to be replicated. Thus, the strands of double-stranded DNA are first completely separated (or “denatured” using heat, or other adjustments to the conditions of the reaction), leaving two complementary single-stranded DNA template strands.

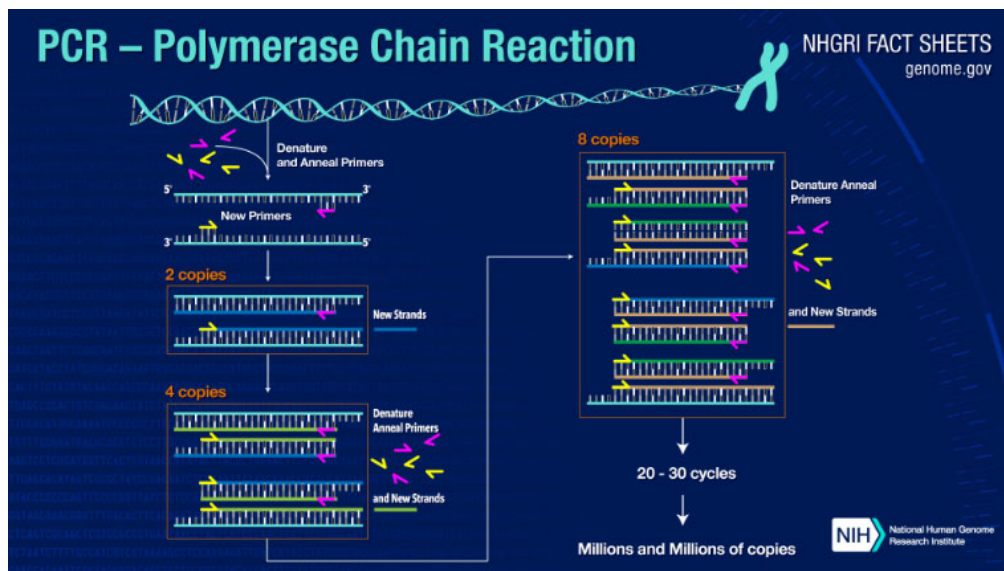
72. PCR “primers” or “probes,” which are short oligonucleotides that are complementary to a small portion of the DNA templates, can then be added to the denatured DNA at high concentrations. These primers/probes then hybridize (i.e., bind through base pairing interactions) to the DNA template strand at a particular target region of interest (specified by the order of bases) to capture that specific sequence for replication. The small portions of the template to which the primers/probes hybridize are referred to as the “priming sites.”

73. After the PCR primer hybridizes to the target sequence, an enzyme called a “polymerase” then synthesizes a new nucleic acid molecule complementary to the single-stranded DNA template starting at the 3’ end of the primer. The polymerase then adds nucleotides (one nucleotide at a time) complementary to the rest of the template until a complete copy of the template is made.

74. After a complete copy is made, the strands are then denatured (again, typically through heating), and the process starts over, each time allowing for the doubling of the copies of template strands present in a sample following hybridization of the PCR primers to the priming sites. Each round of denaturation, primer hybridization, and extension is referred to as a “cycle.”

75. PCR results in the exponential amplification of target DNA sequences. In the first cycle, two new strands of DNA are created in addition to the template strand of DNA. In the second cycle, copies are made of each of the four strands, for

a total of eight strands. As PCR continues, the number of copies increase at an exponential rate as depicted in the schematic below.²



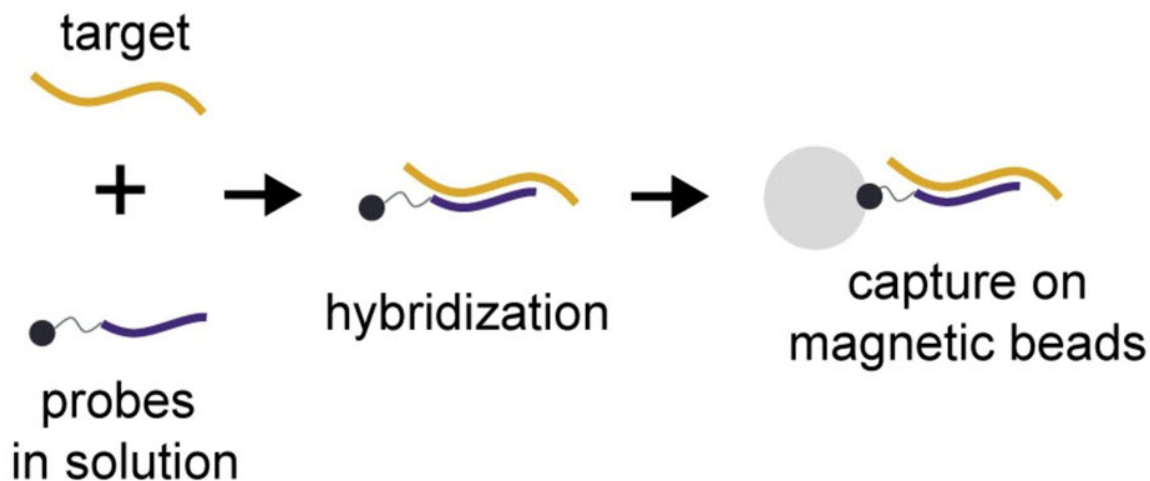
(b) Solution-Based Hybrid Capture

76. A POSA would have also been familiar with the use of solution-based hybrid capture techniques for target enrichment. In short, this technique works by mixing in solution the target DNA with an abundance of short nucleic acid “baits,” which hybridize to the target DNA. The baits also include an additional component (typically biotin) that serves as a ligand for a binding partner (typically streptavidin)³

² See <https://www.genome.gov/about-genomics/fact-sheets/Polymerase-Chain-Reaction-Fact-Sheet> (last accessed 11/15/2022).

³ The streptavidin-biotin complex is well-known to be one of the strongest binding interactions occurring in nature. The binding strength is significantly greater than that between two complementary DNA strands, and is far beyond the normal protein-ligand binding strength.

that is affixed to a magnetic bead, to allow bead-probe conjugation following the hybridization step. When such beads are mixed into the solution, the target sequence is captured through bead-probe binding, and the target DNA can be isolated by washing the uncaptured DNA and probes.



The key aspects of this process were described in detail in a seminal 2009 article by Andreas Gnirke and Chad Nusbaum at the Broad Institute. Ex. 1005 (Gnirke 2009).

77. When hybrid capture baits are contacted with the nucleic acid sample (the “pool”), the baits hybridize to the sequences of interest. This hybridization results in a “catch” of nucleic acids bearing the targeted sequences, which are “caught” through the hybridization of strand-to-strand base-pairing. The catch is then pulled down or fished out using beads that bind the biotin that was incorporated into the baits, and the catch can then be amplified using PCR before sequencing, for example using NGS workflows. Ex. 1005 (Gnirke) at 183; Fig. 1 (reproduced

below). *See also* Ex. 1006 (Mertes 2011) at 374 (“By selective recover[y] and subsequent sequencing of genomic loci of interest, costs and efforts can be reduced significantly compared with whole-genome sequencing.”); 375 (“Targeted enrichment can be useful in a number of situations where particular portions of a whole genome need to be analyzed.”).

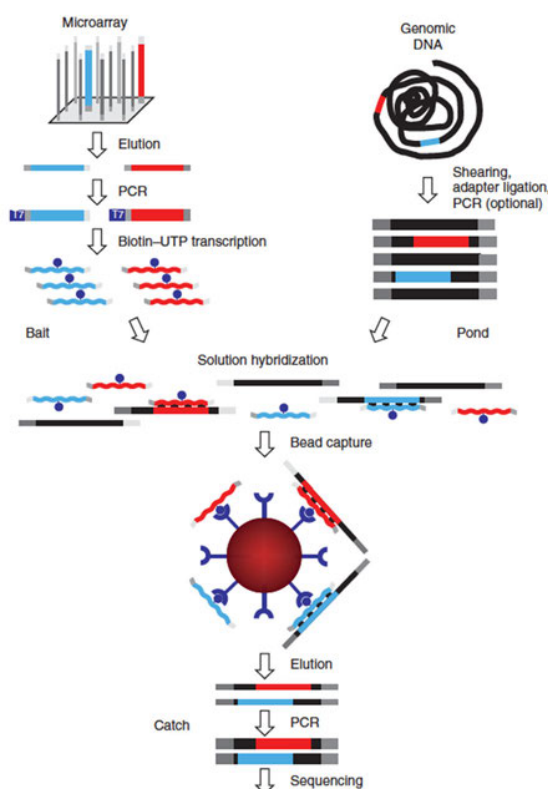


Figure 1 Overview of hybrid selection method. Illustrated are steps involved in the preparation of a complex pool of biotinylated RNA capture probes (bait; top left), whole-genome fragment input library (pond; top right) and hybrid-selected enriched output library (catch; bottom). Two sequencing targets and their respective baits are shown in red and blue. Universal adaptor sequences are gray. The excess of single-stranded nonself-complementary RNA (wavy lines) drives the hybridization.

78. Well before the priority date, “an increasing number of protocols and vendors [had] begun offering out of the box solutions for hybrid capture, meaning,

the research need not do development work but merely choose between a preset targeted enrichment regions (e.g. whole exome) or specify their own custom enrichment region.” Ex. 1006 (Mertes) at 379.

D. Personalized Cancer Testing and Monitoring

79. Once NGS techniques became available, those skilled in the art immediately sought to leverage them to develop methods to discover, monitor, and track mutations in cancer patients. For example, a 2008 *Nature* paper by Timothy Ley and others applied whole-genome sequencing and personalized monitoring techniques to improve treatment of acute myeloid leukemia, the treatment of which “ha[d] changed little in the past two decades, because most of the genetic events that initiate the disease remain undiscovered.” Ex. 1012 (Ley 2008) at 66. Once “[w]hole-genome sequencing [became] possible at reasonable cost and timeframe,” Ley “use[d] this approach for the unbiased discovery of tumour-specific somatic mutations that alter the protein-coding genes.” *Id.* Through sequencing tumor samples that were matched to normal counterparts from skin samples taken from the same patient, Ley “discovered ten genes with acquired mutations; two were previously described mutations that are thought to contribute to tumour progression, and eight were new mutations present in virtually all tumour cells at presentation and relapse” *Id.* The authors reported that “their study establishes whole-genome sequencing as an unbiased method for discovering cancer-initiating

mutations in previously unidentified genes that may respond to targeted therapies.”

Id.

80. Ley described the specifics of its protocol as follows:

Our group has pioneered the use of whole-genome re-sequencing and variant discovery approaches using the Illumina/Solexa technology with the genome of the nematode worm *Caenorhabditis elegans* as a proof-of-principle. This approach has distinct advantages in reduced cost, a markedly increased data production rate, and a low input requirement of DNA for library construction. In the present study, we used a similar approach to sequence the tumour genome of a single AML patient and the matched normal genome (derived from a skin biopsy) of the same patient. After alignment to the human reference genome, sequence variants were discovered in the tumour genome and compared to the patient’s normal sequence, to the dbSNP database, and to variants recently reported for two other human genomes; revealing new single nucleotide and small insertion/deletion (indel) variants genome-wide. Somatic mutations were detected in genes not previously implicated in AML pathogenesis, demonstrating the need for unbiased whole-genome approaches to discover all mutations associated with cancer pathogenesis.

Ex. 1012 (Ley) at 67. The whole-genome sequencing was performed using “the Illumina Genome Analyser to achieve the targeted level of 90% diploid coverage as determined by the coverage of the high-quality SNP set. Maq [a genome assembly software] was used to perform alignment, determine consensus, and identify SNVs

within the 98 billion bases generated from the tumor genome.” *Id.* They also “sequenced the genome of normal skin cells from the same patient to enable the identification of inherited sequence variants in the tumor genome.” *Id.*

81. Using the identified mutations, in order “[t]o better define the percentage of tumour cells that contained each of the discovered somatic mutations,” Ley “amplified each mutation-containing locus from non-amplified genomic DNA derived from the *de novo* and relapse tumour samples, and from the skin biopsy obtained at presentation.” *Id.* at 69. By looking at the *de novo* (initial) and relapse tumor samples, Ley described analyzing the nucleic acid levels present in the tumor at multiple time points throughout the course of the disease. The resulting enriched populations of these sequences then “were sequenced using the Roche/454 FLX⁴] platform, and the frequency of reads containing the reference and variant alleles were defined.” *Id.* The allele frequency of the variant alleles is shown in Ley’s Figure 3, which I have reproduced below, with annotations added to show the 10 somatic and 2 germline mutations that were analyzed at different stages of the disease (red and green bars), and compared to the skin sample (control, yellow bars) that was collected at the same time as the primary tumor.

⁴ The Roche/454 FLX system that used sequencing by synthesis but with a different method of detecting the nucleotides (pyrosequencing) as they were added in sequence as compared to Illumina sequencing. Pyrosequencing relies on light detection based on a chain reaction when pyrophosphate is released.

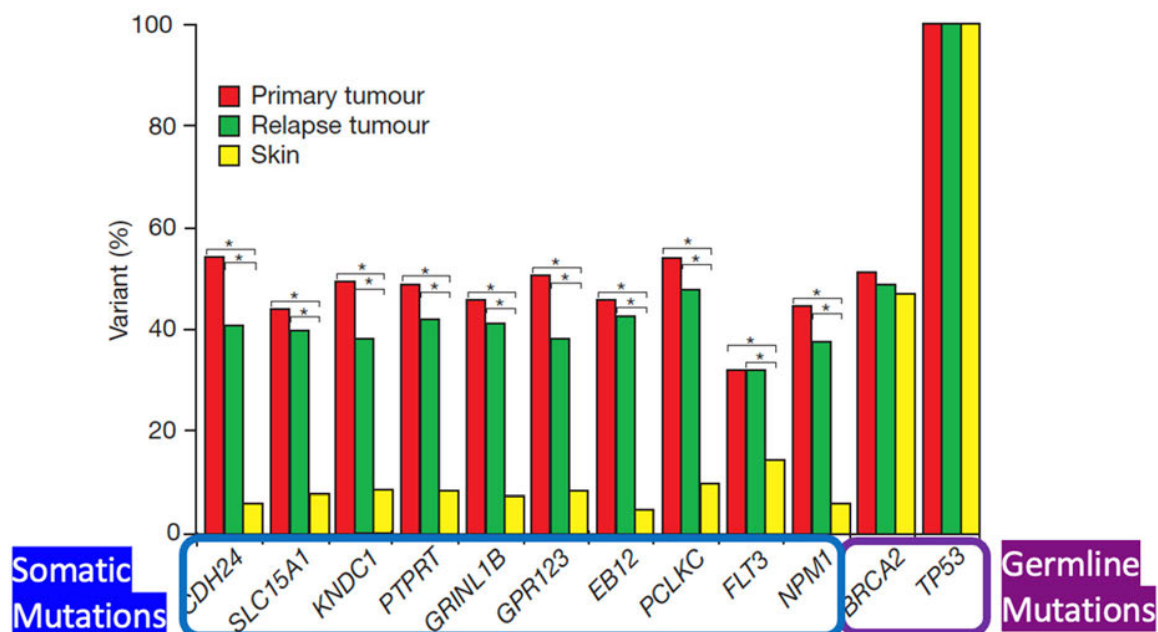


Figure 3 | Summary of Roche/454 FLX readcount data obtained for ten somatic mutations and two validated SNPs in the primary tumour, relapse tumour and skin specimens. The readcount data for the variant alleles in the primary tumour sample and relapse tumour sample are statistically different from that of the skin sample for all mutations ($P < 0.000001$ for all mutations, Fisher's exact test, denoted by a single asterisk in all cases). Note that the normal skin sample was contaminated with leukaemic cells containing the somatic mutations. The patient's white blood cell count was 105,000 (85% blasts) when the skin punch biopsy was obtained.

82. A key insight of Ley is that characterizing the patient's tumor using whole-genome sequencing to identify a patient-specific polymorphic fingerprint and then using that data to target specific polymorphisms identified in the whole-genome sequencing step to track the tumor over time is a useful approach to cancer monitoring and testing. Specifically, Ley observed that "[o]ur results strongly support the notion that hypothesis-driven (for example, candidate gene-based) examination of tumour genomes by PCR-directed or capture-based methods is

inherently limited, and will miss key mutations.” *Id.* at 70. A POSA would have appreciated that in Ley’s approach, which utilizes an initial whole-genome sequencing step to understand the nature of the patient’s tumor, the results of the whole-genome sequencing would help avoid “miss[ing] key mutations,” and only after that information is collected, a specific set of probes for further sequencing of key mutations can be created. *See* Ex. 1031 (Meyerson) at 687 (describing the use of whole-genome sequencing to discover genomic alteration events in cancer tissue in Ley and subsequent studies); 694 (“[C]urrent efforts are focused on applying deep targeted second-generation sequencing to validate the events.”); Fig. 2 (explaining use of hybrid capture for targeted sequencing in cancer genomics applications).

83. Soon after, Leary (Ex. 1002) and McBride (Ex. 1015) demonstrated the use of the same concept, by using whole-genome sequencing to identify structural variants (as opposed to SNVs as in Ley), and to subsequently use those identified markers as targets for serial monitoring of tumor DNA in the blood plasma of patients with solid tumors. *See also* Ex. 1031 (Meyerson) at 687 (Table 1) (listing Ley and subsequent whole cancer genome studies).

E. Related Applications of Next-Generation Sequencing

84. In parallel with the development of sequencing-based techniques for cancer monitoring, those skilled in the art were also applying very similar techniques to the detection and monitoring of other medical conditions. Thus, very similar target

selection, target enrichment, and target monitoring techniques were simultaneously developed in fields such as prenatal genetic testing and organ transplant rejection. This is particularly true in applications that use cell-free DNA (“cfDNA”) from the patient as a marker of disease.

85. Cell-free DNA (“cfDNA”) refers to small fragments of DNA present in the bloodstream. This generally comes from DNA released during apoptosis, or the death of cells, in the body. During apoptosis, cells undergo various cell death processes. Part of that involves cellular mechanisms that cause the DNA of the cells’ chromosomes to be cleaved and fragmented. The fragmented DNA, along with other parts of the cells, are released into the bloodstream, and eventually cleared out of the body with other bodily waste products.

86. Cells are constantly dying and being replaced throughout the body. Therefore, there is always cfDNA being released into and circulating through the bloodstream (or other body fluids, including cerebrospinal fluid and urine) before it is cleared out of the body. Thus, tumors are constantly releasing cfDNA into the patient’s bloodstream, a fetus is constantly releasing cfDNA into its mother’s bloodstream, and a transplanted organ (e.g., a liver or heart transplant) is constantly releasing cfDNA into the transplant recipient’s bloodstream. By identifying distinctions between the “foreign” DNA – whether from a tumor, fetus, or

transplanted organ – and the “host” DNA, the levels of foreign DNA present in the host can be used as powerful markers of biomedical information.

87. Before the priority date, those skilled in the art recognized the applicability of similar nucleic acid analysis workflows to each of these unique clinical settings. For example, by 2011, Dr. Stephen Quake had aptly described these connections:

Particularly for cancer, where the circulating DNA bears hallmark signs of the disease including mutations in oncogenes, microsatellite alterations, and, for certain cancers, viral genomic sequences, DNA or RNA in plasma has become increasingly studied as a potential biomarker for disease. For example, Diehl et al recently demonstrated that a quantitative assay for low levels of circulating tumor DNA in total circulating DNA could serve as a better marker for detecting the relapse of colorectal cancer compared with carcinoembryonic antigen, the standard biomarker used clinically. Maheswaran et al reported the use of genotyping of circulating cells in plasma to detect activating mutations in epidermal growth factor receptors in lung cancer patients that would affect drug treatment. *These results collectively establish both circulating DNA, either free in plasma or from circulating cells, as useful species in cancer detection and treatment. Circulating DNA has also been useful in healthy patients for fetal diagnostics, with fetal DNA circulating in maternal blood serving as a marker for gender, rhesus D status, fetal aneuploidy, and sex-linked disorders.* Fan et al recently demonstrated a strategy for detecting fetal aneuploidy

by shotgun sequencing of cell-free DNA taken from a maternal blood sample, a methodology that can replace more invasive and risky techniques such as amniocentesis or chorionic villus sampling.

Ex. 1003 (Quake) at ¶40 (citations omitted); *see also id.* at ¶41 (“In all these applications of circulating nucleic acids, the presence of sequences differing from a patient’s normal genotype has been used to detect disease. In cancer, mutations of genes are a telltale sign of the advance of the disease; in fetal diagnostics, the detection of sequences specific to the fetus compared to maternal DNA allows for analysis the health of the fetus.”). Quake then observed that “as cell-free DNA or RNA often arises from apoptotic cells, the relative amount of donor-specific sequences in circulating nucleic acids should provide a predictive measure of on-coming organ failure in transplant patients for many types of solid organ transplantation” *Id.* at ¶42.

F. Providing Actionable Information to Healthcare Providers

88. The ability to monitor and analyze a patient’s cancer using these techniques was an important advance. Those skilled in the art recognized that it was also important to process the obtained sequencing information into a form that would be clinically useful: “Adoption of clinical NGS testing will place significant demands on laboratory infrastructure and will require extensive computational

expertise and a deep knowledge of cancer medicine and biology to generate truly useful ‘clinically actionable’ reports.” Ex. 1014 (Ross 2011) at 527.

89. Studies of cancer biology yielded discoveries of associations and relationships between certain genetic mutations and differential effects on cancer, such as response to treatment, severity of response, and more. For example, as Ross summarizes:

In late 1998, the era of personalized oncology began with the simultaneous regulatory approval of the anti-HER2-targeted monoclonal antibody therapeutic, trastuzumab, for the treatment of HER2-overexpressing breast cancer and the immunohistochemical-based diagnostic HercepTest . . . for the identification of patients who are eligible for treatment. During the ensuing 12 years, interest in these “drug-test” combinations, also referred to as “companion diagnostics,” has increased exponentially as more anticancer agents have been approved and more new drugs have entered clinical trials based on biomarker profiles. *The selection of anti-epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors in non-small cell lung cancer (NSCLC) and the use of anti-EGFR monoclonal antibody therapeutics in colorectal cancer (CRC) are now firmly based on pretreatment sequence determinations for hot spots in the EGFR (NSCLC) and KRAS (CRC) genes.*

Ex. 1014 (Ross 2011) at 527. Ross further notes that more-recently developed therapies based on “sequence detection-based companion diagnostics” had shown

impressive results.” *Id.* (discussing PLX-4032 drug candidate targeting cancers “that feature the V600E mutation of the *BRAF* gene,” and crizotinib having “impressive efficacy for the treatment of NSCLCs that feature an *EML4-ALK* translocation”).

90. “Given the potential generation of ‘actionable’ gene sequence abnormalities and abnormalities of unknown clinical importance,” Ross reported, “the successful NGS-based clinical service will have a result report that is focused on the abnormalities that can lead directly to the best available therapy, direct a patient to an ongoing clinical trial, or result in some other actionable response from the treating oncologist.” *Id.* At 537. And these sequencing-based technologies can be used with other forms of cancer cell diagnostics such that these data points can “be combined with tumor cell DNA sequencing in some form of integrated laboratory report.” *Id.* at 536.

91. Such reports were described in a number of prior art references, which depicted information about the relative abundance or a frequency of polymorphic alleles present in a patient’s sample at multiple points in time over the course of treatment. For example, Ley’s Figure 3 depicts sequence information regarding the relative abundance of particular somatic and germline mutations that were observed in samples taken at (i) primary onset, and (ii) relapse. Ex. 1012 (Ley) at 69 (Fig. 3) (reproduced below). These abundances are compared to the frequency of the same polymorphic alleles present in skin, which represents a control.

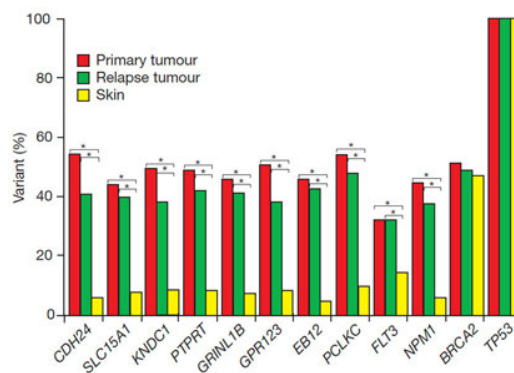


Figure 3 | Summary of Roche/454 FLX readcount data obtained for ten somatic mutations and two validated SNPs in the primary tumour, relapse tumour and skin specimens. The readcount data for the variant alleles in the primary tumour sample and relapse tumour sample are statistically different from that of the skin sample for all mutations ($P < 0.000001$ for all mutations, Fisher's exact test, denoted by a single asterisk in all cases). Note that the normal skin sample was contaminated with leukaemic cells containing the somatic mutations. The patient's white blood cell count was 105,000 (85% blasts) when the skin punch biopsy was obtained.

VI. Prior Art References of the Grounds

A. Leary (Ex. 1002)

92. Leary, et al., “Development of Personalized Tumor Biomarkers Using Massively Parallel Sequencing,” is an article that was published in *Science Translational Medicine* on February 24, 2010, and is thus prior art to the '394 patent. Leary describes methods of analyzing patient-specific translocations in solid tumors using whole genome sequencing, followed by serial monitoring of those translocations as biomarkers for disease in patient plasma samples. Ex. 1002 at 1 (Abstract). “This approach provides an exquisitely sensitive and broadly applicable approach for the development of personalized biomarkers to enhance the clinical management of cancer patients.” *Id.* The Supplementary Materials to Leary, Ex.1026, are cited in the primary manuscript and available online at <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2858564/>. In my experience, when

Supplemental Materials are cited in articles appearing in reputable journals like *Science Translational Medicine*, such materials are made available contemporaneously with the article's publication.

B. Ley (Ex. 1012)

93. Ley, et al., "DNA sequencing of a cytogenetically normal acute myeloid leukemia genome," is an article that was published in *Nature* by November 6, 2008, and is thus prior art to the '394 patent. Ley "establishes whole-genome sequencing as an unbiased method for discovering cancer-initiating mutations in previously unidentified genes that may respond to targeted therapies." Ex. 1012 at 66. See Section ¶V.D, above.

C. Chan (Ex. 1008)

94. Chan, et al., "Cancer Genome Scanning in Plasma: Detection of Tumor-Associated Copy Number Aberrations, Single-Nucleotide Variants, and Tumoral Heterogeneity by Massively Parallel Sequencing," is an article that was published in *Clinical Chemistry* by January 1, 2013. See Oxford Academic Press Website (<https://academic.oup.com/clinchem/article/59/1/211/5622098>) ("Published 01 January 2013). *Clinical Chemistry* is a well-known and respected journal in the field, and those skilled in the art routinely rely on the publication date information provided by the publishers of major, reputable journals. Moreover, Dr. Dennis Lo, a prominent researcher in the field based out of the Chinese University of Hong Kong,

and the senior author on Chan, appeared on a podcast for the American Association of Clinical Chemistry describing the significance of the Chan article that was posted online on January 3, 2013, further confirming the availability of Chan before the priority date. Ex. 1017 (<https://www.aacc.org/science-and-research/clinical-chemistry/clinical-chemistry-podcasts/2013/cancer-genome-scanning-in-plasma-detection-of-tumor-associated-copy-number-aberrations>). Thus, Chan is prior art to the '394 patent.

D. Liao (Ex. 1009)

95. Liao et al., “Targeted Massively Parallel Sequencing of Maternal Plasma DNA Permits Efficient and Unbiased Detection of Fetal Alleles,” is another article from Dennis Lo’s lab that was published in *Clinical Chemistry* by January 1, 2011. See Oxford Academic Press Website (<https://academic.oup.com/clinchem/article/57/1/92/5621222>) (Published: 01 January 2011). Liao was cited by Chan (Ex. 1008), further confirming its availability to a POSA before the priority date. Liao is therefore prior art to the '394 patent. Liao generally describes the use of target enrichment approaches for targeted sequencing of selected genomic regions from plasma DNA.

VII. Claim Construction

96. I have generally applied the plain and ordinary meaning that the claims of the '394 patent would have had to a POSA in view of the '394 patent’s

specification, and as Patent Owner has read the claims in its complaint for infringement against Petitioner. Below I set forth my understanding of several claim terms at issue, which I have applied in each of the grounds of invalidity set forth below. Each of these terms is present in claim 1 of the '394 patent, its sole independent claim:

Claim 1: A method for analyzing nucleic acid samples isolated from an individual, comprising:

(a) generating a first subset of nucleic acid molecules from a first nucleic acid sample obtained from said individual;

(b) conducting a first sequencing assay on said first subset of nucleic acid molecules to yield a first result comprising a first nucleic acid sequence wherein said first sequencing assay comprises whole genome sequencing;

(c) generating a second subset of nucleic acid molecules from a second nucleic acid sample of said individual by contacting said second nucleic acid sample with a plurality of capture probes, wherein said plurality of capture probes hybridize to one or more polymorphisms which are based on or extracted from one or more databases of polymorphisms, observed in a population of one or more samples, or a combination thereof, wherein said first sample or said second sample is obtained from said subject at a first time point;

(d) conducting a second sequencing assay on said second subset of nucleic acid molecules to yield a second result comprising a second nucleic acid sequence;

(e) using said plurality of capture probes to generate an additional subset of nucleic acid molecules from an additional nucleic acid sample isolated from an additional sample obtained from said subject at an additional time point different from said first time point;

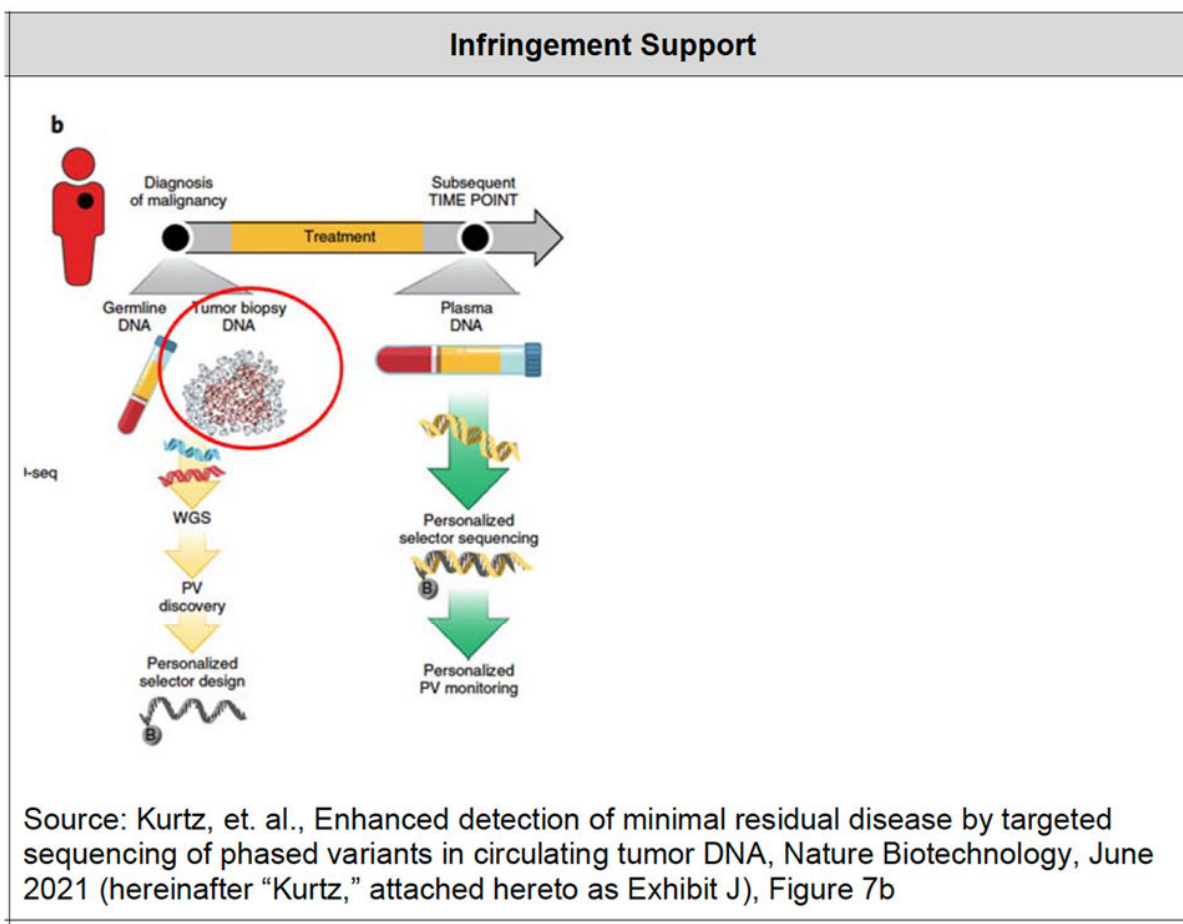
(f) conducting a third sequencing assay on said additional subset of nucleic acid molecules from said additional nucleic acid sample to yield a third result comprising a third nucleic acid sequence; and

(g) generating a biomedical report that includes biomedical information of said subject, which biomedical information identifies the presence or absence of said one or more polymorphisms identified by said second result or said third result.

A. “generating a first subset of nucleic acid molecules”

97. I have formed my opinions based on the understanding of claim limitation 1(a), “generating a first subset of nucleic acid molecules from a first nucleic acid sample obtained from said individual,” applied by Patent Owner in Exhibit I to its complaint against Petitioner. Ex. 1011 (Ex. I to Patent Owner’s Complaint). As explained below, Patent Owner’s implicit claim construction for “subset” is the term’s ordinary dictionary definition meaning of any “part of another set.”

98. Patent Owner alleged that this limitation is satisfied in Petitioner’s methods when “[t]he first subset of nucleic acid molecules is obtained from a tumor biopsy obtained from an individual.” *Id.* at PDF2-3. The evidence that Patent Owner cited in support of this contention, derived from an article authored by some of Petitioner’s scientists, Ex. 1016 (Kurtz), is depicted below, showing annotations added by Patent Owner.



99. Thus, in Patent Owner’s application, the “first nucleic acid sample” from which the claimed “first subset” is “generated” may be a tumor biopsy from an

individual. A POSA would understand that a tumor biopsy comprises nucleic acids, and in Patent Owner's application, the claimed "subset" is met when a portion of those nucleic acids from the tumor biopsy are extracted for further assaying. Moreover, a POSA would understand that known DNA extraction technologies extract less than 100% of the nucleic acids present in a tumor biopsy. *See, e.g.*, Ex. 1025 (stating that in 2014, "[i]solating total DNA from small samples using traditional methods is difficult and inefficient mainly due to loss of DNA during filtration and precipitation," and reporting "DNA extraction method" that "provides high quality and improved yields," but not 100% yields). Accordingly, a POSA would understand that routine DNA extraction from a tumor biopsy meets the "generating a first subset" step under Patent Owner's implicit claim construction.

B. "capture probe"

100. I have applied the meaning the term "capture probe" would have had to a POSA in view of the '394 patent's specification. The '394 patent explains that the main feature of a "capture probe" is that it "typically . . . comprises a nucleic acid binding site":

Capture Probes[:] The methods disclosed herein may comprise one or more capture probes, a plurality of capture probes, or one or more capture probe sets. Typically, the capture probe comprises a nucleic acid binding site. The capture probe may further comprise one or more linkers. The capture probes may further comprise one or more labels.

The one or more linkers may attach the one or more labels to the nucleic acid binding site.

Ex. 1001 ('394 Patent) at 31:4-12.

101. The nucleic acid binding sites may comprise a sequence “that hybridizes to at least a portion of the one or more nucleic acid molecules . . . in the sample or subset of nucleic acid molecules.” *Id.* at 31:24-27. Thus, the '394 patent explains, “capture probes may comprise a nucleic acid binding site that hybridizes to one or more genomic regions.” *Id.* at 31:28-31; *see also id.* at 31:35-36 (“The capture probes may comprise one or more nucleotides.”).

102. Additional features for capture probes, such as linkers and labels, are described as optional. *See id.* at 31:8-12 (“The capture probe *may* further comprise one or more labels. The capture probes *may* further comprise one or more linkers. The one or more linkers *may* attach the one or more labels to the nucleic acid binding site.”); *id.* at 31:62-32:5.

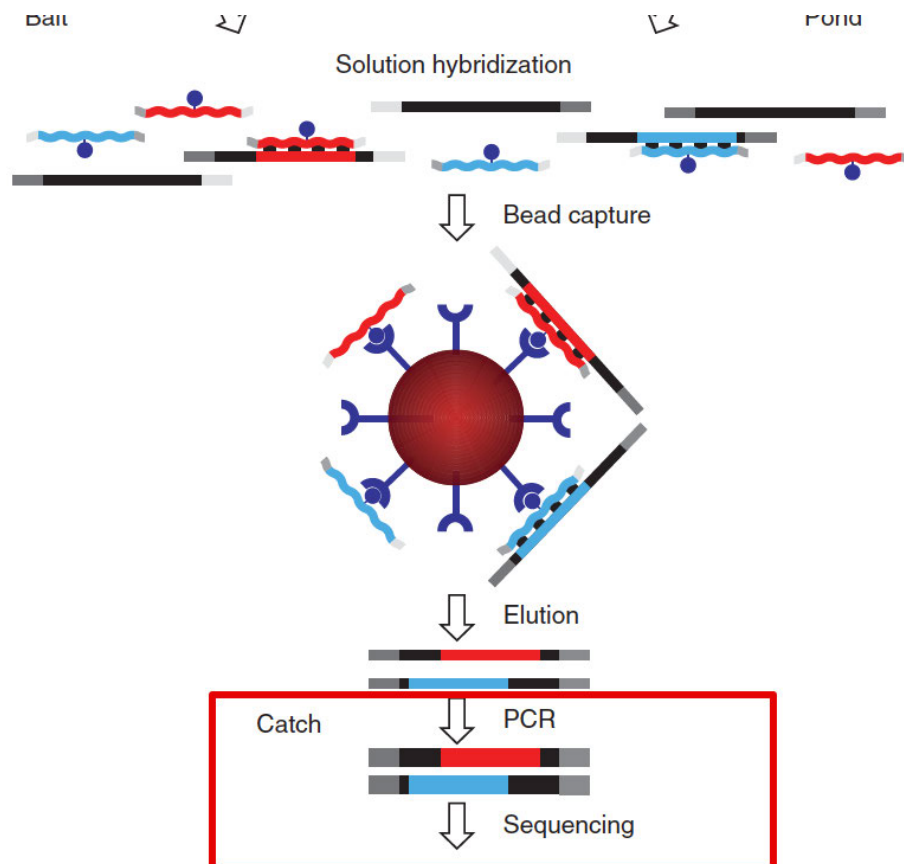
103. Accordingly, a POSA would have understood the term “capture probe” in the '394 patent to encompass at least PCR primer capture probes and solution-based hybrid capture probes (e.g., biotinylated oligonucleotide baits).

104. I further note that claim 13 recites, “[t]he method of claim 1, wherein said plurality of capture probes further comprises barcodes.” A barcode is a short oligonucleotide sequence that can be used to identify a particular nucleic acid in a

sequencing experiment. A POSA would have understood that barcodes were routinely employed in connection with PCR probes to allow the pooling of many samples in a single sequencing reaction. Because the barcode sequence would be incorporated into the PCR products, the barcodes can then be used to deconvolute the data. *See* Ex. 1007 (Mamanova) at 113 (when using PCR amplification, “[f]or the sequencing to be economical, it is necessary to barcode and pool many samples and to sequence these pools in a single lane. Several approaches to sample barcoding have been reported . . .”).

105. When hybrid capture baits are used, the capture probes themselves are not typically incorporated into the target nucleic acids that are enriched for subsequent sequencing. Taking the workflow explained in Gnirke as an example, Figure 1 shows that after the capture probes bind a DNA sequence of interest, streptavidin coated beads are introduced that bind the biotin present on the capture probes to separate out the catch. The captured nucleic acid “catch” that was bound to the capture probe is then released for amplification and sequencing, because the binding of streptavidin to biotin is much stronger than the binding between the nucleic acid of the capture probe and the target DNA. I have annotated an excerpt of Gnirke’s Figure 1 below with a red box showing that the “catch” resulting from a typical hybrid capture experiment consists of target DNA that does not incorporate any part of the capture probe. This shows that incorporating a barcode into a capture

probe would not typically be understood to effect a change in the resultant “catch,” because the capture probe (and any elements it comprises, like a barcode) remain bound to the bead while the catch is carried forward for use in subsequent steps. *See* Ex. 1005 (Gnirke) at Fig. 1.



106. Since claim 1 is broader than claim 13, this confirms my understanding that the '394 patent intended to encompass PCR primers/probes within the claimed “capture probes.”

C. “polymorphism”

107. I have applied the meaning the term “polymorphism” would have had to a POSA in view of the ’394 patent’s specification. “Polymorphism” in the context of genetics simply means that there are multiple forms of a nucleic acid sequence of interest. The ’394 patent’s specification is consistent with this understanding, explaining that a “polymorphism” may comprise essentially any mutation, alteration, or variation of a particular nucleic acid sequence:

Generally, a polymorphism refers to a mutation in a genotype. A polymorphism may comprise one or more base changes, an insertion, a repeat, or a deletion of one or more bases. Copy number variants (CNVs), transversions and other rearrangements are also forms of genetic variation. Polymorphic markers include restriction fragment length polymorphisms, variable number of tandem repeats (VNTR's), hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, simple sequence repeats, and insertion elements such as Alu. The allelic form occurring most frequently in a selected population is sometimes referred to as the wildtype form. Diploid organisms may be homozygous or heterozygous for allelic forms. A diallelic polymorphism has two forms. A triallelic polymorphism has three forms. Single nucleotide polymorphisms (SNPs) are a form of polymorphisms. *In some aspects of the disclosure, one or more polymorphisms comprise one or more single nucleotide variations, indels, small insertions, small deletions, structural variant junctions, variable length tandem repeats, flanking*

sequences, or a combination thereof. The one or more polymorphisms may be located within a coding and/or non-coding region. The one or more polymorphisms may be located within, around, or near a gene, exon, intron, splice site, untranslated region, or a combination thereof. The one or more polymorphisms may be may span at least a portion of a gene, exon, intron, untranslated region.

Ex. 1001 at 27:48-28:11.

D. “biomedical report”

108. I have applied the meaning of the term “biomedical report” would have had to a POSA in view of the ’394 patent’s specification. The ’394 patent explains that:

The methods and systems disclosed herein may further comprise producing one or more outputs based on the analysis of the nucleic acid sample. The one or more outputs may comprise one or more biomedical reports. The one or more biomedical reports may comprise biomedical information of a subject. The biomedical information of the subject predicts, prognoses, or diagnoses one or more biomedical features selected from the group, comprising disease state, genetic risk of a disease, reproductive risk, genetic risk to a fetus, risk of an adverse drug reaction, efficacy of a drug therapy, prediction of optimal drug dosage, transplant tolerance, or a combination thereof.

Ex. 1001 (’394 patent) at 8:32-43; 52:7-26.

VIII. The Claims of the '394 Patent Would Have Been Obvious

A. Ground 1: Leary in view of a POSA's Knowledge

109. Claims 1-11 and 13-19 are obvious over Leary (Ex. 1002) in view of a POSA's knowledge described in Section V, above, and as described further below.

110. Leary describes a method for analyzing nucleic acid samples isolated from an individual with cancer. Leary explains that “[a] nearly universal feature of human cancer is the widespread rearrangement of chromosomes as a result of chromosomal instability.” Ex. 1002 at 1. “Such structural alterations begin to occur at the earliest stages of tumorigenesis and persist throughout tumor development.” *Id.* Leary observes that the consequences of such mutations, i.e., polymorphic variations of the cancer genome, “can include copy number alterations (duplications, amplifications, and deletions), inversions, insertions, and translocations.” *Id.* Leary describes an approach that can leverage NGS techniques to detect and use such rearrangements “as highly sensitive biomarkers for tumor detection” and monitoring, given that “[s]uch alterations are not present in normal cells and should be exquisitely specific.” *Id.* (“Rearrangement-associated biomarkers therefore offer a reliable measure that would be useful for monitoring tumor response to specific therapies, detecting residual disease after surgery, and long-term clinical management.”). Leary's approach includes the following general steps.

111. **First**, whole-genome, paired-end sequencing is conducted on nucleic acids from samples derived from the patient's normal (germline) and cancerous cells to allow "the identification of patient-specific rearrangements in tumor samples." *Id.* at 1 (Description of the Approach). Leary "initially analyzed four tumor samples (two colon and two breast tumors) and their matched normal tissue samples using the Applied Biosystems SOLiD System." *Id.* After conventional processing and whole-genome sequencing, the resulting sequencing reads were mapped to the reference human genome. *Id.*

112. **Second**, using the whole-genome sequencing, Leary "identif[ied] somatic rearrangements from these data." *Id.* ("Identification of somatic rearrangements"). By comparing the tumor and germline genetic sequences, Leary was able to identify patient specific genetic rearrangements and copy-number variations that were specific to the tumor, including ruling out variation that was present in the patient's normal, germline sequence. *Id.* at 1-2; *see also* Fig. 1. Using their approach, Leary "identified 57 regions containing putative somatic rearrangements, with an average of 14 rearrangements per sample." *Id.* at 3; *see also* Table 2.

113. **Third**, Leary designed PCR primers that would hybridize to, and hence "capture," nucleic acid sequences spanning the junctions of the identified rearrangements. *Id.* at 3. This allowed the authors to confirm that the rearrangements

were present in the tumor samples, but not in normal tissue, because the primers confirm the presence of the rearrangement when they “yield[] PCR products of the expected size in the tumor samples but not in the normal samples.” *Id.* Sanger sequencing was then used to “confirm[] the rearrangements.” *Id.*

114. **Fourth**, with knowledge of the patient-specific rearrangements of interest from whole-genome sequencing and related analysis, Leary “evaluated circulating DNA from plasma samples of patients” whose cancer’s rearrangements were analyzed. *Id.* at 4. “PCR amplification of plasma DNA using primers spanning the breakpoints produced products of the expected sizes only in the plasma samples from patients with disease and not in plasma from healthy controls.” *Id.*; *id.* Fig. 4. “Sequencing of the PCR products from [patient’s] plasma DNA identified the identical breakpoints observed in the tumor DNA samples.” *Id.*

115. **Fifth**, the PCR analysis was then repeated to analyze plasma samples from the patients at multiple time points “throughout the course of treatment.” *Id.*; *id.* Fig. 4. By monitoring the amount of tumor DNA present in the plasma (and detected by capturing nucleic acids harboring the rearrangement using the PCR primers) over time, the patient’s disease status was monitored. *Id.* (“The fraction of mutant DNA present in plasma samples was 37% just before surgery but decreased to 14% 1 day after resection of the primary tumor. . . . [T]he fraction of mutant tumor

DNA did not reach zero (remaining at 0.3% at day 137), consistent with the fact that this patient had residual metastatic lesions in the remaining left lobe of the liver.”).

116. *Sixth*, Leary describes generating a biomedical report depicting the fraction of rearranged DNA in plasma at various time points, which conveys biomedical information such as the disease status and response to treatment, including an increase in the tumor burden post-surgery and a decrease following chemotherapy and the removal of liver metastasis. *Id.* Fig. 4 (reproduced below).

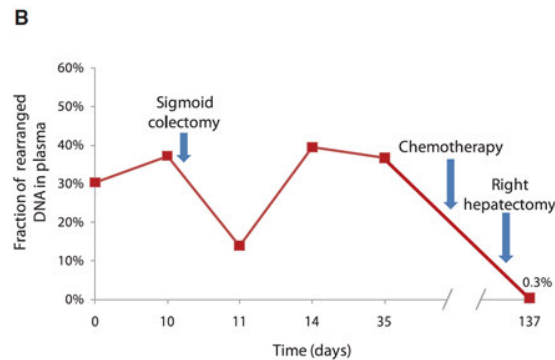


Fig. 4. Detection of tumor-specific rearrangements in plasma of cancer patients. (A) The identified chromosome 4:8 and 16 rearrangements were used to design PCR primers spanning breakpoints and to amplify rearranged DNA from tumor tissue and plasma from patients Hx402 and Hx403, respectively. A plasma sample from an unrelated healthy individual was used as a control for both rearrangements. (B) Plasma samples from patient Hx402 were analyzed at different time points using digital PCR to determine the fraction of genomic equivalents of plasma DNA containing the chromosome 4:8 rearrangement. The fraction of rearranged DNA at day 137 was 0.3%, consistent with residual metastatic lesions present in the remaining lobe of the liver.

117. As explained in further detail below, Leary in view of a POSA’s knowledge renders the claims of the ’394 patent obvious.

1. Claim 1

118. Leary expressly discloses the preamble of claim 1 and steps 1(a)-1(e) and 1(g). To the extent that Leary does not expressly disclose step 1(f), Leary in view of the knowledge of a POSA renders that step obvious, as discussed below.

119. Leary discloses the claim's preamble, "[a] method for analyzing nucleic acid samples isolated from an individual." Specifically, Leary explains:

We have developed a method, called personalized analysis of rearranged ends (PARE), which can identify translocations in solid tumors. Analysis of four colorectal and two breast cancers with massively parallel sequencing revealed an average of nine rearranged sequences (range, 4 to 15) per tumor. Polymerase chain reaction with primers spanning the breakpoints was able to detect mutant DNA molecules present at levels lower than 0.001% and readily identified mutated circulating DNA in patient plasma samples. This approach provides an exquisitely sensitive and broadly applicable approach for the development of personalized biomarkers to enhance the clinical management of cancer patients.

Ex. 1002 (Leary) at 1 (Abstract).

120. Moreover, as discussed further below, because Leary discloses and/or renders obvious the claimed method steps, Leary discloses "a method for analyzing nucleic acid samples isolated from an individual."

(a) “generating a first subset of nucleic acid molecules from a first nucleic acid sample obtained from said individual”

121. Leary discloses step 1(a), the step of “generating a first subset of nucleic acid molecules from a first nucleic acid sample obtained from said individual,”

122. Leary teaches generating a first subset of nucleic acid molecules by extracting DNA, a type of nucleic acid, from tumor samples. Specifically, Leary discloses:

The PARE (personalized analysis of rearranged ends) approach, shown schematically in Fig. 1, requires the identification of patient-specific rearrangements in tumor samples. To determine the feasibility of identifying such alterations using next-generation sequencing approaches, we initially analyzed four tumor samples (two colon and two breast tumors) and their matched normal tissues using the Applied Biosystems SOLiD System (Table 1). *Genomic DNA from each sample was purified, sheared, and used to generate libraries* with mate-paired tags ~1.4 kb apart.

Ex. 1002 (Leary) at 1 (“Description of the approach”); *see also* Fig. 1 Caption (“The method is based on next-generation mate-paired analysis of resected tumor DNA to identify individualized tumor-specific rearrangements.”).

123. Leary further discloses obtaining a nucleic acid sample from a patient’s tumor sample in its “Materials and Methods” section. Specifically, Leary’s protocol included a step in which “DNA samples were obtained from early-passage

xenografts and cell lines of breast and colorectal cancers as described. Normal DNA samples were obtained from matched normal tissue.” *Id.* at 5 (“Clinical samples and cell lines”).

124. A POSA would have understood that each tumor sample in Leary comprised a “nucleic acid sample,” and that extracting genomic DNA from that sample generated a “subset” of the nucleic acid molecules of the sample under Patent Owner’s implicit interpretation of step 1(a) because some but not all of the nucleic acids from the tumor sample were extracted. *See* Section VII.A (claim construction for “generating a first subset of nucleic acid molecules”), above.

(b) “conducting a first sequencing assay on said first subset of nucleic acid molecules to yield a first result comprising a first nucleic acid sequence wherein said first sequencing assay comprises whole genome sequencing”

125. Leary discloses conducting a whole genome sequencing assay on the first subset of nucleic acid molecules. Thus, Leary discloses step 1(b).

126. After extracting the first subset of DNA from a tumor sample as described above, Leary reports that such DNA was subjected to a whole genome sequencing assay. Specifically, Leary explains:

Genomic DNA from each sample was purified, sheared, and used to generate libraries with mate-paired tags ~1.4 kb apart. Libraries were digitally amplified by emulsion polymerase chain reaction (PCR) on magnetic beads (21), and 25-bp mate-paired tags were sequenced using

the sequencing-by-ligation approach (15,22). An average of 198.1 million 25-bp reads was obtained for each sample, *where each read aligned perfectly and was uniquely localized in the reference human genome (hg18), resulting in 4.95 Gb of mappable sequence per sample*. The total amount of genome base pairs covered by the mate-paired analysis (that is, distance between mate-paired tags x number of mate-paired tags) was 53.6 Gb per sample, *or an 18-fold physical coverage of the human genome*.

Ex. 1002 at 1. Leary's whole genome sequencing assay on the first subset yielded a first result, namely a nucleic acid sequence of the tumor sample.

- (c) “[i] generating a second subset of nucleic acid molecules from a second nucleic acid sample of said individual by contacting said second nucleic acid sample with a plurality of capture probes, [ii] wherein said plurality of capture probes hybridize to one or more polymorphisms which are based on or extracted from one or more databases of polymorphisms, observed in a population of one or more samples, or a combination thereof, [iii] wherein said first sample or said second sample is obtained from said subject at a first time point”

127. Leary discloses generating a second subset of nucleic acid molecules by contacting tumor DNA with a plurality of capture probes, such as PCR primer capture probes, and thus discloses step 1(c).

(i) Step 1(c)(i)

128. Leary discloses the step of “generating a second subset of nucleic acid molecules from a second nucleic acid sample by contacting said second nucleic acid sample with a plurality of capture probes.”

129. Using the results of Leary’s whole genome sequencing steps, the authors analyze the sequencing results using two approaches to identify somatic rearrangements in the tumor’s genome. Ex. 1002 at 1-2 (“Identification of somatic rearrangements”). One “involved searching for tags whose mate pairs were derived from different chromosomes,” which identified “interchromosomal rearrangements” where portions of one chromosome rearrange or relocate to a different chromosome as compared to their location on the reference human genome. *Id.* The other technique “combined mate-paired tag data with copy number alterations identified by analysis of individual 25-bp tags.” *Id.* This approach helped identify “de novo rearrangements,” the boundaries of which “would be expected to contain novel junctions not present in the human genome.” *Id.* Through the resulting analysis, the authors were able to identify “DNA sequences that have inappropriate spacing, order, or orientation on the same chromosome (intrachromosomal rearrangements), or inappropriate joining of sequences from different chromosomes (interchromosomal rearrangements).” *Id.* at 2.

130. By analyzing the whole-genome sequencing data, Leary “identified 57 regions containing putative somatic rearrangements, with an average of 14 rearrangements per sample.” *Id.* From these data, Leary then:

[D]esigned primers to 42 of the paired-end regions and used them for PCR spanning the putative breakpoints. Thirty-five of these (83%) yielded PCR products of the expected size in the tumor samples but not in the normal samples. Sanger sequencing of seven PCR Products confirmed the rearrangements in all cases tested.

Id.; *see also id.* at 6 (“Primers were used for PCR on tumor and matched normal samples as previously described. The candidate rearrangement was confirmed if a PCR product of the expected size was seen in the tumor but not in the matched normal sample. Sanger sequencing of PCR products was used to identify sequence breakpoint in a subset of cases.”).

131. Leary’s design and use of PCR primers spanning the putative breakpoints of the 42 identified rearrangements constitutes the use of capture probes to generate a second subset of DNA from the tumor sample – the PCR products. The generated PCR products are nucleic acid sequences comprising the chromosomal structural variant junction at the rearrangement. A POSA would have understood these PCR primer capture probes to be the claimed “capture probes” because they are nucleic acid sequences that “hybridize to” the complementary sequences on the polymorphic gene comprising the breakpoint. *See* Section VII.B (claim construction

of “capture probe”). Specifically, the PCR primer capture probes hybridize and anneal to DNA surrounding the polymorphism of interest, i.e., a breakpoint evidencing a translocation or rearrangement of the gene. *See* Section V.C(a) (“PCR”).

(ii) Step 1(c)(ii)

132. Leary discloses that the PCR primer capture probes utilized in step 1(c)(i) “hybridize to one or more polymorphisms which are based on or extracted from one or more databases of polymorphisms, observed in a population of one or more samples, or a combination thereof,” as claimed. Specifically, Leary discloses the use of PCR primer capture probes which hybridize to one or more polymorphisms that were observed in a population of one sample—the identified polymorphic rearrangements are “*patient specific*” and serve as “*personalized* tumor biomarkers.” Ex. 1002 (Leary) at 1 (The PARE approach “requires the identification of patient specific rearrangements in tumor samples.”); *id.* at 3 (“For confirmation, we designed primers to 42 of the paired-end regions and used them for PCR spanning the putative breakpoints.”); Ex. 1025 (Leary Supplemental Information) (showing sequences of PCR primers that hybridize to complementary sequences of the nucleic acid polymorphisms).

Table S3. Confirmed somatic rearrangements in breast and colorectal cancer samples*

Sample	Forward tag		Reverse tag		Type	Primer 1	Primer 2
	Chromosome	Position	Chromosome	Position			
B5C	3	52,638,626	3	52,573,088	AAC	AAGTTTTCAAGCTTTACCTGAAGT	TATATTGGAAGAATAGAAATGAATGG
B5C	4	93,109,700	4	-93,105,085	BAC	AGCCAAGTGCAATTCTCCAG	GCACACTGTTTGCAGGAATG
B5C	11	57,713,780	8	-48,889,516	C**	GCCACCTTCTTCTTCTCTGA	AAGCTTTGTTGGTTGTTCTCA
B5C	18	19,141,985	20	-29,591,944	C**	TGGCTTTCAAAACCCACTG	TCCTTCTGCCCATTAGGG
B5C	22	48,743,603	2	-104,047,142	C**	TCATGGTTTATCCACGGTGT	CACACCGCATTACACAAAC
B7C	1	-96,237,189	7	65,542,257	C**	TCAAAACAGAAAGCATTAGGC	CGCATCCAAAGTATTATAGCAA
B7C	2	197,428,606	2	113,761,988	AAC	AACTCCTCCACCTCAAAATC	CCAAATTGCCTGCTTAAGAGAT
B7C	2	-32,084,286	3	185,241,029	C**	TGCTACCAATACITCCCACTTG	TACCGTCTCCACGGCATGT
B7C	2	114,604,628	18	53,562,784	C**	GGAGAAAACCCCTGGTTATTTTA	TCCCTCATCAGAGCAATCA
B7C	3	-115,579,348	3	-115,651,310	AAC	AAATTGGGAAGGATCATACTGAC	TCTGAACATGCCTGATCTCATC
B7C	4	785,983	4	733,804	AAC	CTGAACCTCTGGGCTGAA	TTGCTAAGTGATGCTACCTGTG
B7C	5	107,405,959	5	107,231,803	AAC	CCTGGCCCTTAGGTAAGAT	TGAAGATCCTCTTAGTGATGGAA
B7C	5	38,284,430	10	-44,715,202	C**	TGCAGCTTTTCTGTCTTCA	CTGCCAGTCCAAACTGGTG
B7C	6	106,401,376	6	90,853,847	AAC	TGCTGTTTCAAATTCCTACAGTC	TGAAATTAGGACCTGGAGCAC
B7C	6	101,933,981	6	102,444,426	ABC	GCCAGGTAACATGCTCACTTT	GATGCAGGAAGTTGACAGCA
B7C	9	22,003,033	9	21,761,298	AAC	GGGCTAAGCTTAAGAGTCTGG	GCCATGTGCAAGTCAAGAAG
B7C	11	-6,436,033	11	-6,519,897	AAC	TCTGCCGGCATACTGGAC	TAAGGGCATTGGAACAAGG
B7C	12	65,950,588	12	65,923,399	AAC	GCCCTATTTTCAGAGAAAGTGGA	AACATCTCTCTTTTGAAGATCC
B7C	13	60,438,525	13	52,159,979	AAC	AATTTGCTCTCATCGTATTGTGT	AGCTGAATCAAAATTTCCAATG
B7C	X	31,583,118	X	31,179,704	AAC	CTGAATCTCTTCCAGCAAAAT	AATGGGTTAAGCAGTTTAGGG
Co108C	2	191,184,628	5	-104,930,827	C**	TAGCATGCACCACTTTTAGGC	AAAGGTTAAAGGACTGTTTAAAGTTG
Co108C	2	78,849,963	6	-13,299,323	C**	GGTTCTGGAGGTTGGAGA	GTAAAGTCAACATTTTGTTCAGG
Co108C	2	-7,268,710	6	13,299,385	C**	TATGCCACCATCCTCTAGGT	TCCCAAGTCAAAATCAAAACCA
Co108C	2	-141,266,018	13	96,916,170	C**	GGTGTCTCTCTCCCATACCA	CGATCTATACACCAACCCACA
Co108C	3	-60,400,269	3	-60,437,489	AAC	TGCTTTTAGTTTGGGTACGG	GCTGATTTGTTATACCCAGTGC
Co108C	3	-60,365,933	3	-60,498,861	AAC	ATCCTCGGACTGGACTGAGA	AACCCCATCTGAAGCTACC
Co108C	3	60,573,034	3	60,472,593	AAC	GGGTTATCTCAAAGGGCAGA	GCTCTCAATTTGTGATTGG
Co108C	4	81,934,151	15	54,039,041	C**	TGTGTCTCTCTCTTAAGCAT	GACTACAAATGGCCAGACTC
Co108C	6	-13,299,291	5	157,523,537	C**	ATCCCCACATTCCCAACC	CCCAGCCATATGTTTGGTTTA
Co108C	6	13,299,271	2	-20,956,947	C**	GTAATTTGTTATGTTTGTAGGTGT	TCAATGGGGAGAGAGAGC
Co108C	13	34,581,537	10	67,756,452	C**	ACGTGTGATTGGGGGTAGC	CCAGATGGCTGGGTTAAATAA
Co84C	8	128,442,121	19	49,144,200	C**	AGCTAGGTGGAGAAATTTGTCTG	GGCTTCTGTAGAGTGACATGA
Co84C	11	34,790,251	13	109,267,462	C**	AAGGAGATTGGTTATTGTGGA	CTGCAGGAACCTGTCTCATCTT
Co84C	11	-34,405,644	15	88,736,701	C**	TGCTGAATCATTCTCCCAACT	TGGTGAATCCACTGAGGTGA
Co84C	15	-89,096,347	8	127,747,412	C**	GCATTCTAAAGATGAAGTCCCAT	GGAAACCCTGTAGTGAAAAGTC
Hx402x	8	96,971,644	4	156,043,548	C**	CAGGTGATATACCAAGAAATAGG	TTTGGGTTCAAGTTCTATTTGAAGA
Hx402x	5	-100,413,406	5	-137,521,052	AAC	AGTCAACGCCCTAGCATGG	TGGGCATGAGCAAGATATTC
Hx402x	8	-144,771,376	8	-144,787,051	AAC	AATCACGTTGGGTGACTGTG	GTGACAGGCTGGGTGTCC
Hx402x	14	-85,526,541	14	-85,560,400	AAC	TGAAGGTTGAGTTGCCAGTG	TGTATGAACATTGTAGAGGCTGT
Hx403x	1	119,547,240	1	-119,550,445	BBC	AGGAGGAAAGCAACACATAGAG	GGTGATTTTCAATGCATATTTCA
Hx403x	5	-27,160,637	5	27,150,736	BBC	AATTACCACAACCTCCACGAG	CAAAAGATTCTCAAATGCAGGT
Hx403x	11	66,674,459	11	66,662,814	AAC	TGAATCAGAAAGTCTGGCAGT	CACCTTGAGAACTAATGATGCAG
Hx403x	16	6,343,641	16	6,727,736	AAC	CCTAGCCCTTTGTCCCTGT	TTTGTGTACCTAGACATTATCCAA
Hx403x	16	6,574,321	16	6,759,729	AAC	GCAGAGAACAGCAGAAAGTTG	AGCCAAAGATCAAGCCACAGA
Hx403x	16	26,579,136	16	-26,582,595	BBC	TTCTCTTCTCTGCTTCACTG	TTGATGATTGAAGAACTAGCCGTG
Hx403x	17	34,622,352	17	-34,624,284	BBA	GGCTCCCTCTCCATTCC	CTGCTGACGTGCTGCTCTT

* A single representative mate pair is shown for each rearrangement. Forward and reverse tags and their genomic coordinates correspond to F3 and R3 SOLiD mate pair tags. The type of rearrangement corresponds to the categories described in http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_058717.pdf. AAC corresponds to mate pairs spanning deletions; codes starting with B denote incorrect strand orientation; codes containing a B at the middle position denote incorrect ordering; and C** corresponds to interchromosomal translocations. Primers 1 and 2 correspond to primers used for confirming tumor-specific rearranged sequences.

LEARY ET AL. SUPPLEMENTARY MATERIAL, PAGE 3

133. The polymorphisms analyzed in Leary include chromosomal rearrangements, which comprise at least deletions (the rearranged DNA is deleted from its prior chromosomal location), insertions (the rearranged DNA is inserted into its new chromosomal location), and structural variant junctions (the new junctions created as a result of the rearrangement). *See* Section VII.C (claim

construction for “polymorphism”) (e.g., “Generally, a polymorphism refers to a mutation in a genotype. A polymorphism may comprise one or more base changes, an *insertion*, a repeat, or a *deletion* of one or more bases. Copy number variants (CNVs), transversions and other *rearrangements* are also forms of genetic variation.”). *See also* Claim 3 (dependent claim expressly providing that the claimed polymorphisms include “deletions”). A capture probe that hybridizes to a deletion necessarily hybridizes to sequences that flank where the deleted nucleotide existed (since the deleted sequence is no longer present). Thus, Leary’s capture probes, which hybridize to the nucleic acid sequences spanning chromosomal rearrangements (which themselves comprise deletions, insertions and structural variant junctions), “hybridize to” such polymorphisms within the meaning of the claim.

134. Leary further reported that 83% of the PCR primer capture probes designed for its study “yielded PCR products of the expected size in the tumor samples but not in the normal samples (Fig. 2 and table S3).” Ex. 1002 at 3. This means that the PCR primer capture probes hybridized to, and in fact captured, the genetic polymorphism of interest – in this case, the genetic variant carrying the rearrangement – but not the wild type, unrearranged version of that same gene (hence the absence of PCR products observed when the PCR primer capture probes are contacted with normal DNA). This is shown in Figure 2A, which shows identifiable

bands in the “T” column, indicating the presence of a PCR product created through the hybridization of the PCR primer capture probe in the Tumor samples, but no band in the “N” column, indicating that those capture probes failed to hybridize to Normal DNA lacking the polymorphism. I have annotated exemplary lanes in Figure 2A to show this.

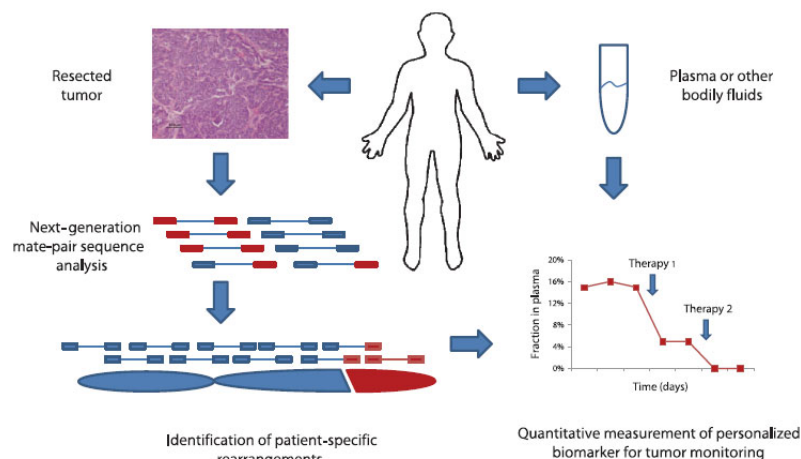


(iii) Step 1(c)(iii)

135. Leary discloses the claimed method “wherein said first sample or said second sample is obtained from said subject at a first time point.”

136. Leary explains that the PARE approach “requires the identification of patient specific rearrangements in tumor samples.” Ex. 1002 at 1. The tumor samples are, for example, obtained from the patient at a first time point, i.e., the time they are resected. *Id.*; *id.* at 2 (Fig. 1) (identifying “Resected tumor”).

Fig. 1. Schematic of PARE approach. The method is based on next-generation mate-paired analysis of resected tumor DNA to identify individualized tumor-specific rearrangements. Such alterations are used to develop PCR-based quantitative analyses for personalized tumor monitoring of plasma samples or other bodily fluids.



137. Leary also discloses that the second sample may be a plasma sample from the patient taken at time points after the initial tumor resection. Ex. 1002 at 4.

Leary explains that:

To determine whether the rearranged sequences could actually be detected in clinical samples, we evaluated circulating DNA from plasma samples of patients Hx402 and Hx403. *The sample from patient Hx403 was obtained before surgery, whereas the samples from patient Hx402 were obtained before and after surgery. . . .* PCR amplification of plasma DNA using primers spanning the breakpoints produced products of the expected sizes only in the plasma samples from patients with disease and not in plasma from healthy controls (Fig. 4A). Sequencing the PCR products from plasma DNA identified the identical breakpoints observed in the tumor DNA samples.

Ex. 1002 at 4.

(d) “conducting a second sequencing assay on said second subset of nucleic acid molecules to yield a second result comprising a second nucleic acid sequence”

138. Leary discloses step 1(d), “conducting a second sequencing assay on said second subset of nucleic acid molecules to yield a second result comprising a second nucleic acid sequence.”

139. Leary discloses that, after generating a second subset of nucleic acid molecules by contacting the second sample with PCR primer capture probe, as described above, “Sanger sequencing of seven PCR products confirmed the rearrangements in all cases tested.” Ex. 1002 at 3. The output of a Sanger sequencing assay is a nucleic acid sequence. *See* Section V.A.I (“Sequencing Techniques”).

140. When the second sample is a plasma sample from a patient, Leary discloses that the PCR products obtained by contacting such samples with its capture probes are subjected to a second sequencing assay, which results in a second nucleic acid sequence. Ex. 1002 at 4. Leary explains:

To determine whether the rearranged sequences could actually be detected in clinical samples, we evaluated circulating DNA from plasma samples of patients Hx402 and Hx403. The sample from patient Hx403 was obtained before surgery, whereas the samples from patient Hx402 were obtained before and after surgery. . . . PCR amplification of plasma DNA using primers spanning the breakpoints produced products of the expected sizes only in the plasma samples from patients with disease and not in plasma from healthy controls (Fig. 4A).

Sequencing the PCR products from plasma DNA identified the identical breakpoints observed in the tumor DNA samples.

Ex. 1002 at 4; *see also id.* (“Digital PCR analyses of plasma samples from patient Hx402 indicated that rearranged tumor DNA was detectable at all six time points analyzed throughout the course of treatment. (Fig. 4B).”).

- (e) **“using said plurality of capture probes to generate an additional subset of nucleic acid molecules from an additional nucleic acid sample isolated from an additional sample obtained from said subject at an additional time point different from said first time point”**

141. Leary discloses step 1(e) by disclosing the use of the PCR primer capture probes to generate additional subsets of nucleic acid molecules from a patient’s blood and/or tumor samples at multiple time points after the first time point, including to monitor the progression and/or response to treatment of the patient’s cancer, as discussed above.

142. Leary’s PCR primer capture probes are used as personalized biomarker probes that can detect rearrangements specific to a patient’s cancer at multiple timepoints. Leary explains:

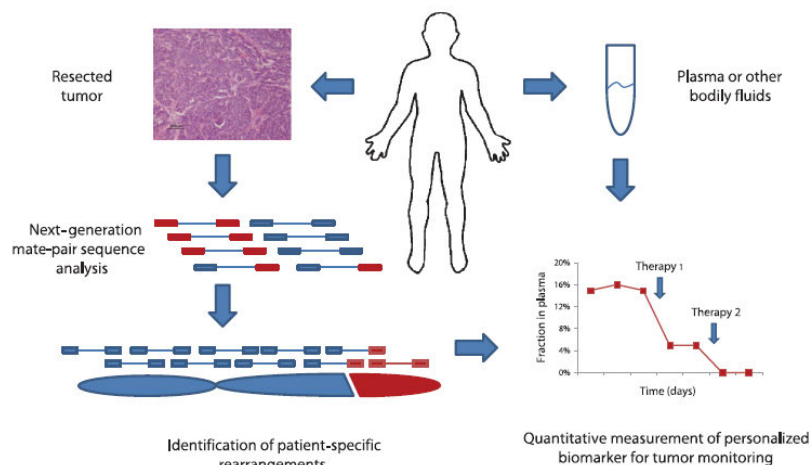
These results demonstrate that massively parallel sequencing can be used to develop personalized biomarkers based on somatic rearrangements. We were able to identify ***tumor-specific markers*** in each of the six breast and colorectal cancer cases analyzed. Moreover,

we demonstrated that the identified breakpoints can be used to detect tumor DNA in the presence of large quantities of normal DNA and in patient plasma. These results highlight the sensitivity and specificity of the approach and suggest broad clinical utility of the PARE method.

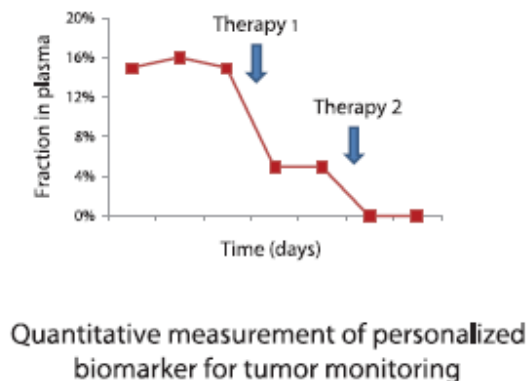
Ex. 1002 (Leary) at 5.

143. Leary explains that when the PCR primer capture probes developed following whole-genome sequencing in its methods are used to monitor a patient's cancer, such "PARE-derived clinical assays should have no false positives," because "PCR amplification of aberrant fusions of DNA sequences that are normally thousands of base pairs apart or on different chromosomes should not occur in nontumor DNA." *Id.* Leary demonstrates this by using the PCR primer capture probes as personalized biomarkers to monitor the patient's cancer over time, thereby generating additional subsets of nucleic acids. For example, Figure 1 shows the "monitoring of plasma samples or other bodily fluids" to monitor the patient's cancer over time using the PCR primer capture probes to facilitate "quantitative measurement of personalized biomarker for tumor monitoring." *Id.* at 2 (Fig. 1).

Fig. 1. Schematic of PARE approach. The method is based on next-generation mate-paired analysis of resected tumor DNA to identify individualized tumor-specific rearrangements. Such alterations are used to develop PCR-based quantitative analyses for personalized tumor monitoring of plasma samples or other bodily fluids.



144. The bottom right panel of Figure 1 (reproduced below) shows the creation of 7 subsets of nucleic acid molecules at different timepoints than the timepoint at which the first subset is generated. Specifically, the graph reports the amount of tumor biomarkers (the polymorphic rearrangements) in the plasma through the use of the PCR primer capture probes to detect such nucleic acids in each sample. The patient whose treatment is depicted in Figure 1 responded to therapy, indicated by the reduction of tumor DNA in the plasma at timepoints after treatment.



145. Leary further describes using PCR primer capture probes to generate additional subsets of nucleic acid molecules from additional nucleic acid samples generated at multiple different time points. For example, Leary describes using its PCR primer capture probes to generate additional subsets of nucleic acid molecules in patient Hx402 at six time points throughout the course of treatment:

Digital PCR analyses of plasma samples from patient Hx402 indicated that rearranged tumor DNA was detectable at all six time points analyzed throughout the course of treatment (Fig. 4B). The fraction of mutant DNA present in plasma samples was 37% just before surgery but decreased to 14% 1 day after resection of the primary tumor. The mutant DNA fraction decreased further after chemotherapy and subsequent removal of metastatic lesions from the right lobe of the liver. However the fraction of mutant tumor DNA did not reach zero (remaining at 0.3% at day 137), consistent with the fact that this patient had residual metastatic lesions in the remaining left lobe of the liver.

Ex. 1002 (Leary) at 4.

146. In Figure 4, Leary shows the detection of tumor specific rearrangements in patient Hx402 “at different time points using digital PCR to determine the fraction of genomic equivalents of plasma DNA containing the chromosome 4:8 rearrangement.” *Id.* at 4 (Fig. 4). The fraction of rearranged DNA – detected using Leary’s PCR primer capture probes – is depicted at six different time points, showing the response of the cancer to multiple treatments during that time. *Id.* Accordingly,

Leary discloses “using said plurality of capture probes to generate an additional subset of nucleic acid molecules from an additional nucleic acid sample isolated from an additional sample obtained from said subject at an additional time point different from said first time point.”

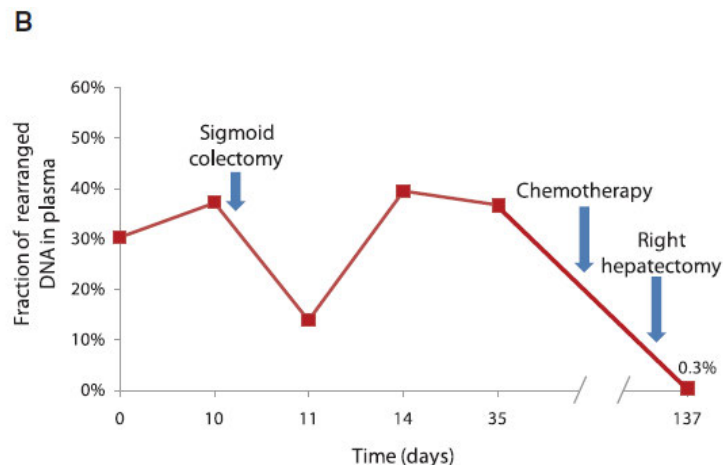


Fig. 4. Detection of tumor-specific rearrangements in plasma of cancer patients. (A) The identified chromosome 4:8 and 16 rearrangements were used to design PCR primers spanning breakpoints and to amplify rearranged DNA from tumor tissue and plasma from patients Hx402 and Hx403, respectively. A plasma sample from an unrelated healthy individual was used as a control for both rearrangements. (B) Plasma samples from patient Hx402 were analyzed at different time points using digital PCR to determine the fraction of genomic equivalents of plasma DNA containing the chromosome 4:8 rearrangement. The fraction of rearranged DNA at day 137 was 0.3%, consistent with residual metastatic lesions present in the remaining lobe of the liver.

(f) **“conducting a third sequencing assay on said additional subset of nucleic acid molecules from said additional nucleic acid sample to yield a third result comprising a third nucleic acid sequence; and”**

147. Leary renders obvious claim step 1(f). As explained above, Leary discloses generating a third subset of nucleic acid molecules when it describes using the PCR primer capture probes to generate additional PCR products from plasma

samples at multiple time points. Leary discloses or at least renders obvious conducting a sequencing assay on such PCR products.

148. Specifically, Leary describes using “PCR amplification of plasma DNA using primers spanning the breakpoints” of patient-specific rearrangements of interest in multiple plasma samples over time. Ex. 1002 at 4. The PCR primer capture probes are used to amplify DNA from plasma samples, and Leary discloses “[s]equencing of the PCR products from plasma DNA” to confirm whether the rearrangement observed in the plasma “identified the identical breakpoints observed in the tumor DNA samples.” *Id.* at 4. Accordingly, Leary discloses, and at a minimum renders obvious, conducting a sequencing assay on PCR products generated from each plasma sample, including the plasma sample from which the third subset of nucleic acid molecules was generated.

149. Moreover, a POSA would have found it obvious to conduct the claimed third sequencing assay on the PCR product generated from the third nucleic acid sample given Leary’s teachings and the knowledge of a POSA. For example, when designing a patient-specific PCR primer (or set of primers) for each patient’s tumor, a POSA would have found it obvious to perform Sanger sequencing (or employ other routine sequencing methods), to verify that the selected PCR primer capture probes selectively bind the tumor-associated nucleic acid sequences they were intended to target. As Leary explains, “[e]ach of the rearranged sequences identified through

PARE was unique, as no identical rearrangement was found in any of the other five tumor samples.” Ex. 1002 (Leary) at 4. An ultimate goal of using the unique identified PCR primer capture probes is “to serve as potential biomarkers,” and to use “PCR assays to detect them in the presence of increasing amounts of normal DNA.” *Id.* Thus, a POSA would have been motivated to sequence the unique rearrangement in each patient, at least to aid monitoring and the design of biomarkers, and for use in further clinical applications.

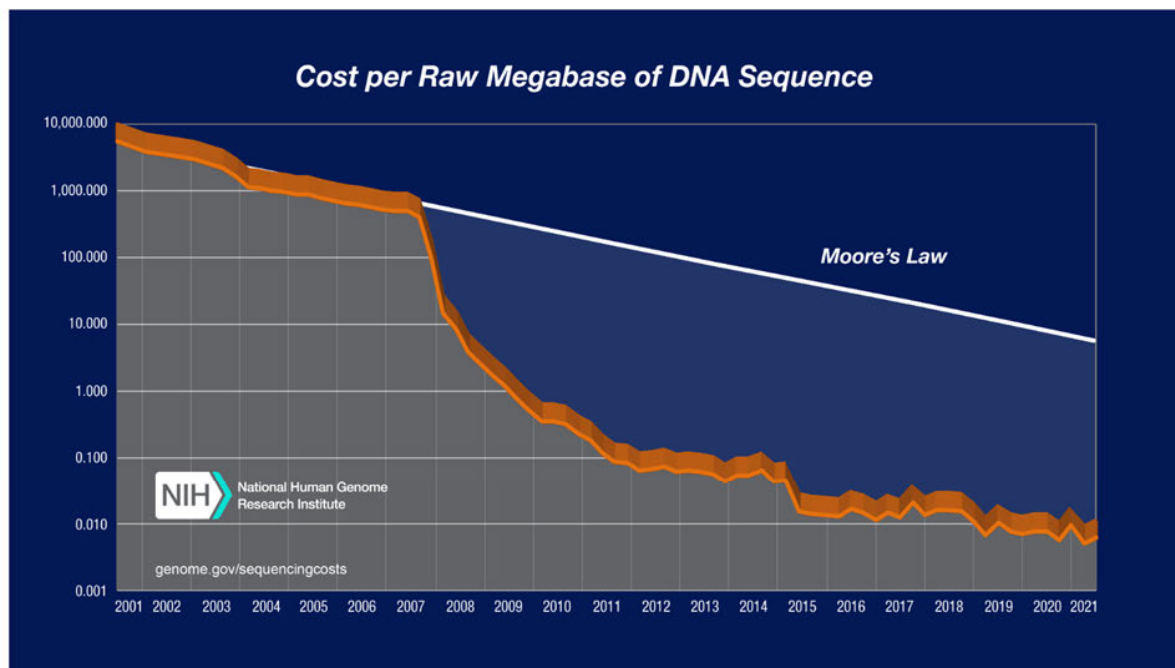
150. As explained above, *see* Section V.C (“Nucleic Acid Sequencing Technologies”), while PCR provides some information about the sequence of a nucleic acid, namely that there is sequence that is complementary to the primers used, it does not provide the same level of information as Sanger sequencing or NGS. This is because Sanger sequencing and NGS confirm, at the single nucleotide level, each nucleotide between the PCR primers used, and not just the sequences complementary to the capture probes used. Thus, a POSA would have been motivated to conduct a sequencing assay on the PCR products of Leary to confirm that the nucleic acids resulting from the PCR step contain the rearrangement’s sequence as identified in the initial whole-genome sequencing step. A POSA would have been particularly motivated to use a confirmatory sequencing step in the context of Leary’s methods, given that the purpose of the analysis is to predict the status of the patient’s cancer. As the impact statement published by *Science*

Translational Medicine accompanying Leary stated: “A diagnosis of cancer shatters the view of the world in an individual’s mind.” Ex. 1002 at 8 (“The individual starts asking questions and searching for possible remedies, and soon learns about the shockingly slow pace of successful cancer research.”). Providing the inquisitive patient and their physician with further information regarding the progress of the patient’s treatment would thus have further motivated a POSA to conduct sequencing of the nucleic acid subsets generated at additional timepoints. *Id.* (“PARE offers a leap forward in the clinical management and treatment of solid tumors . . .”).

151. Moreover, between 2010 when Leary was published (and 2009 when the experiments were likely performed) and the January 2013 priority date of the ’394 patent, available platforms enabled faster sequencing assays, while the cost of sequencing had reduced drastically, which would have further motivated a POSA to conduct confirmatory sequencing steps following each instance of sample collection and creation of a subset of nucleic acids comprising the polymorphism(s) of interest, as in Leary.

152. For example, the National Institutes of Health have tracked the cost of sequencing for decades. In January 2010, right before Leary was published, the sequencing cost for 1 Mb of DNA was about \$0.52/Mb, and in 2009, when Leary’s sequencing was likely performed, the cost per Mb of sequencing fell from \$2.59/Mb

in January to \$0.78/Mb in October. By the January 2013 priority date of the '394 patent, this cost had reduced to \$0.06/Mb. *See* Ex. 1018 (NIH Spreadsheet)⁵; Ex. 1019 (<https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>).



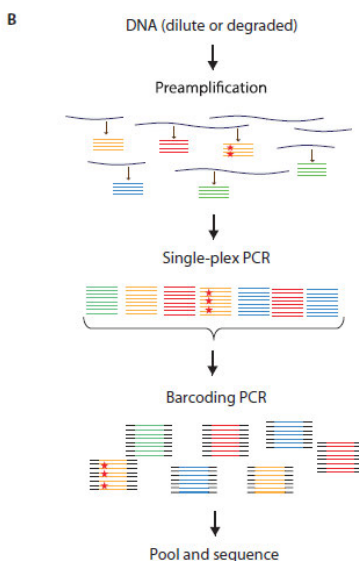
Sequencing cost per megabase - 2021

Thus, by the January 2013 (the '394 patent's priority date), the cost of the confirmatory resequencing step would have been 10-30 times cheaper than it was during the time Leary performed the tests reported in the article, which would have

⁵ This printout was created from the file Sequencing_Cost_Data_Table_Aug2021.xls, which was downloaded from https://web.archive.org/web/20221030082333/https://www.genome.gov/sites/default/files/media/files/2021-11/Sequencing_Cost_Data_Table_Aug2021.xls

further motivated the use of such a step in connection with the treatment of cancer patients.

153. Sequencing a PCR product to confirm the sequence of a rearrangement would have involved no more than the routine skill in the art. As Leary reported, such a step was already successfully employed in its methods. *See also* Ex. 1030 (Forsheew) at 1-2 (applying PCR-based assay “for detection of both abundant and rare mutations in circulating DNA from blood plasma of ... cancer patients. This sequencing approach allowed us to monitor changes in tumor burden by sampling only patient plasma over time.); Fig. 1B (reproduced below) (workflow showing the sequencing of DNA molecules carrying mutations of interest being sequenced following capture and amplification by complementary PCR probes).



154. Accordingly, a POSA would have reasonably expected success in performing such sequencing on additional subsets of nucleic acids collected at later time points (and further subsets at additional time points), as claimed.

(g) “generating a biomedical report that includes biomedical information of said subject, which biomedical information identifies the presence or absence of said one or more polymorphisms identified by said second result or said third result.”

155. Leary discloses step 1(g) because it discloses generating a biomedical report that identifies the presence or absence of a genetic polymorphism identified by the result of the second or third sequencing assay.

156. Leary discloses generating a biomedical report that identifies the presence or absence of the polymorphism identified by the second or third result (i.e., nucleic acid sequence). For example, in Figure 4B, Leary discloses a biomedical report that includes biomedical information about the subject that prognoses and/or diagnoses at least the disease state and efficacy of a drug therapy. Ex. 1002 at 4 (Fig. 4). Specifically, Leary generates a report depicting the “fraction of rearranged DNA in plasma” where the rearranged DNA represents the chromosomal rearrangements confirmed by sequencing in the prior steps. The report shows the patient’s disease status based on the level of circulating tumor DNA observed in the plasma, and also shows the effect of treatments. For example, Figure 4B shows that the treatment reduced this fraction of tumor DNA and concluded that

“[t]he fraction of rearranged DNA at day 137 was 0.3%, consistent with residual metastatic lesions present in the remaining lobe of the liver.” *Id.* Accordingly, Leary discloses step 1(g).

2. Claim 2

157. Claim 2 recites, “[t]he method of claim 1, wherein said biomedical information of said subject is predictive, prognostic, or diagnostic of one or more biomedical features selected from the group consisting of disease state, efficacy of a drug therapy, prediction of optimal drug dosage, recommendation of one or more therapies, and recommendation of a course of treatment of a disease.”

158. I incorporate my analysis of Claim 1 by reference here, including step 1(g). Leary’s biomedical report includes biomedical information that is predictive, prognostic, and/or diagnostic of the patient’s disease state. It is also predictive, prognostic, and/or diagnostic as to at least the efficacy of a drug therapy. *See* Ex. 1002, Fig. 4 (showing patient’s disease state and response to therapy at multiple timepoints).

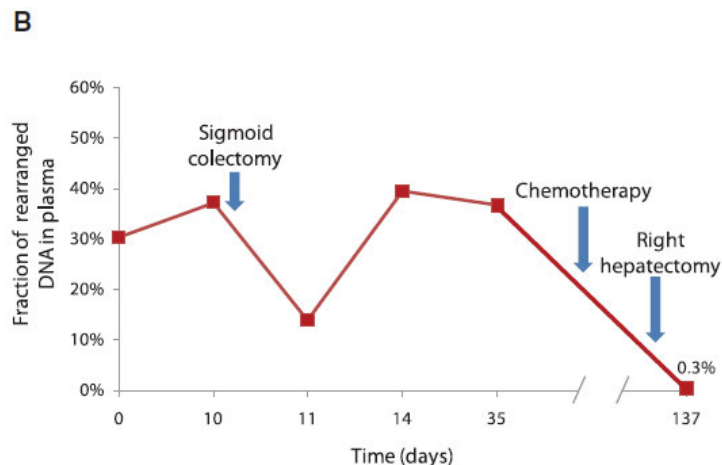


Fig. 4. Detection of tumor-specific rearrangements in plasma of cancer patients. (A) The identified chromosome 4:8 and 16 rearrangements were used to design PCR primers spanning breakpoints and to amplify rearranged DNA from tumor tissue and plasma from patients Hx402 and Hx403, respectively. A plasma sample from an unrelated healthy individual was used as a control for both rearrangements. (B) Plasma samples from patient Hx402 were analyzed at different time points using digital PCR to determine the fraction of genomic equivalents of plasma DNA containing the chromosome 4:8 rearrangement. The fraction of rearranged DNA at day 137 was 0.3%, consistent with residual metastatic lesions present in the remaining lobe of the liver.

159. Accordingly, Leary in view of a POSA's knowledge renders claim 2 obvious.

3. Claim 3

160. Claim 3 recites, “[t]he method of claim 1, wherein said one or more polymorphisms comprise one or more insertions, deletions, structural variant junctions, variable length tandem repeats, single nucleotide mutations, or a combination thereof.”

161. I incorporate my analysis of Claim 1 by reference here. The polymorphisms analyzed in Leary include chromosomal rearrangements, which constitute at least the claimed insertions (the rearranged DNA is inserted into its new

chromosomal location), deletions (the rearranged DNA is deleted from its prior chromosomal location), and structural variant junctions (the new junctions created as a result of the rearrangement).

162. Accordingly, Leary in view of a POSA's knowledge renders claim 3 obvious.

4. Claim 4

163. Claim 4 recites, "[t]he method of claim 1, wherein said one or more polymorphisms are in a set of genes with known biomedically interpretable variants."

164. I incorporate my analysis of Claim 1 by reference here. Given that genes comprise about 1% of the total genome, sequencing or assaying regions of the genome that contain genes would have been particularly useful for generating a meaningful biomedical report. Given that the genome encodes 25,000 genes, sequencing or assaying a subset of these genes or their known variants that have a recognized biological effect would also have increased the likelihood of providing information relevant to a patient's treatment, particularly if the goal is to produce a meaningful biomedical report.

165. Thus, in addition to targeting polymorphisms that were identified in the sample, a POSA also would have found it obvious to target additional polymorphisms with capture probes using Leary's targeted resequencing technique,

including polymorphisms derived from a database of known variants. As explained in Section V.B (“Genetic Mutations”), a POSA would have been familiar with a number of publicly accessible databases that catalogued cancer-associated mutations. A POSA would have been motivated to include such mutations in a panel of tumor biomarkers, in addition to the patient-specific rearrangements identified in patient-derived tumor samples discussed above, in order to provide additional actionable information “to generate truly useful ‘clinically actionable’ reports.” Ex. 1014 (Ross 2011) at 527. *See* Section V.F (“Providing Actionable Information to Healthcare Providers”). For example, because certain cancer-associated-mutations were known to have companion diagnostics or therapies based on the presence or absence of a particular polymorphism or mutation, a POSA would have been motivated to include capture probes targeting such mutations in their panel to facilitate treatment. *See* Ex. 1014 (Ross 2011) at 527 (PLX-4032 drug candidate targeting cancers “that feature the V600E mutation of the *BRAF* gene,” and crizotinib having “impressive efficacy for the treatment of NSCLCs that feature an *EML4-ALK* translocation”).

166. Accordingly, Leary in view of a POSA’s knowledge renders claim 4 obvious.

5. Claim 5

167. Claim 5 recites, “[t]he method of claim 4, wherein said set of genes with known biomedically interpretable variants comprise one or more genes associated with cancer.”

168. As I explain in my analysis of claim 4, which I incorporate by reference here, a POSA would have found it obvious to monitor polymorphisms that occur in genes that are known to be associated with cancer, including to generate reports containing clinically actionable information and in order to determine whether a particular therapy is appropriate for a patient carrying a mutation known to be drug-able by a particular therapy.

169. Accordingly, Leary in view of a POSA’s knowledge renders claim 5 obvious.

6. Claim 6

170. Claim 6 recites, “[t]he method of claim 1, wherein said first time point and said additional time point are at least 1 week apart, or more.”

171. I incorporate my analysis of claim 1 by reference here. Leary’s Figure 4 discloses analyzing the patient’s nucleic acids at time 0, 10 days, 11 days, 14 days, 35 days, and 137 days. Accordingly, Leary discloses a first time point that is at least 1 week apart from an additional time point.

172. Accordingly, Leary in view of a POSA's knowledge renders claim 6 obvious.

7. Claim 7

173. Claim 7 recites, "[t]he method of claim 1, wherein said first time point and said additional time point are at least 1 month apart, or more."

174. I incorporate my analysis of claim 1 by reference here. Leary's Figure 4 discloses analyzing the patient's nucleic acids at time 0, 10 days, 11 days, 14 days, 35 days, and 137 days. Accordingly, Leary discloses a first time point that is at least 1 week apart from an additional time point. Moreover, a POSA would have found it obvious to use an additional time point more than a month apart from the first time point for several reasons. For example, certain cancers spread and proliferate at slower rates than others, which would motivate a POSA to use a longer interval between analyses. Moreover, certain treatments have different time courses and take different lengths of time to begin showing results; thus, it also would have been obvious to use a longer interval, such as a month long interval, before testing following a slower acting treatment. Such choices would involve only routine and predictable choices in experimental design.

175. Accordingly, Leary in view of a POSA's knowledge renders claim 7 obvious.

8. Claim 8

176. Claim 8 recites, “[t]he method of claim 1, wherein said additional time point is later than said first time point.”

177. I incorporate my analysis of claim 1 by reference here. Leary’s Figure 4 discloses analyzing the patient’s nucleic acids at time 0, 10 days, 11 days, 14 days, 35 days, and 137 days. Accordingly, Leary discloses a first time point that is at least 1 week apart from an additional time point, and thus is later than the first time point.

178. Accordingly, Leary in view of a POSA’s knowledge renders claim 8 obvious.

9. Claim 9

179. Claim 9 recites, “[t]he method of claim 1, further comprising comparing said first result and said second result.”

180. I incorporate my analysis of claim 1 by reference here. Leary conducts confirmatory sequencing of the second subset of nucleic acids to confirm that it has the same rearrangement that was observed in the first subset of nucleic acids. The act of “confirming” that the sequences are the same, i.e., that the same rearrangement has been observed at both time points, necessarily involves comparing the two sequencing results (i.e., the claimed first and second results).

181. Accordingly, Leary in view of a POSA’s knowledge renders claim 9 obvious.

10. Claim 10

182. Claim 10 recites, “[t]he method of claim 1, further comprising identifying the presence or absence of said one or more polymorphisms in said second nucleic acid sample.”

183. I incorporate my analysis of claim 1 by reference here. Leary discloses identifying the presence or absence of said one or more polymorphisms – the identified patient-specific rearrangements – in the second nucleic acid sample through confirmatory PCR and through the confirmatory sequencing that it renders obvious. *See also* Ex. 1002 (Leary) at Fig. 4B (confirming presence of rearranged DNA in plasma at multiple time points).

184. Accordingly, Leary in view of a POSA’s knowledge renders claim 10 obvious.

11. Claim 11

185. Claim 11 recites, “[t]he method of claim 10, wherein said identifying comprises observing an allele frequency of said one or more polymorphisms.”

186. I incorporate my analysis of claim 10 by reference here. Figure 4B of Leary depicts observing an allele frequency of the serially-monitored polymorphisms at multiple time points. Specifically Figure 4B shows on its Y-axis the “Fraction of rearranged DNA in plasma,” which is the polymorphic allele of interest that is being tracked over time. Leary observes this fraction (or frequency)

over time, and thus discloses observing an allele frequency to identify the presence or absence (and quantity of) a polymorphism in a sample at a particular time point.

187. Accordingly, Leary in view of a POSA's knowledge renders claim 11 obvious.

12. Claim 12

188. Claim 12 is not challenged in this ground.

13. Claim 13

189. Claim 13 recites, "[t]he method of claim 1, wherein said plurality of capture probes further comprise barcodes."

190. I incorporate my analysis of claim 1 by reference here. A POSA would have been motivated to further incorporate barcodes into Leary's PCR primer capture probes, especially in scenarios in which it was desired to track multiple polymorphisms identified through whole-genome sequencing serially in subsequent samples. A barcode is a short oligonucleotide sequence that can be used to identify a particular nucleic acid in a sequencing experiment. A POSA would have understood that barcodes were routinely employed in connection with PCR probes to allow the pooling of many samples in a single sequencing reaction. Because the barcode sequence would be incorporated into the PCR products, the barcodes could then have been used to deconvolute the data. *See* Ex. 1007 (Mamanova) at 113 (when using PCR amplification, "[f]or the sequencing to be economical, it is necessary to

barcode and pool many samples and to sequence these pools in a single lane. Several approaches to sample barcoding have been reported”). Accordingly, a POSA would have reasonably expected success in implementing this well-known technique in combination with the use of Leary’s PCR primer capture probes.

191. Accordingly, Leary in view of a POSA’s knowledge renders claim 13 obvious.

14. Claim 14

192. Claim 14 recites, “[t]he method of claim 1, wherein said samples isolated from said individual are isolated from a body fluid, cell, skin, tissue, organ, or combination thereof.”

193. I incorporate my analysis of claim 1 by reference here. As explained above, Leary utilizes samples isolated from bodily fluids like plasma, as well as tissue/organ biopsies.

194. Accordingly, Leary in view of a POSA’s knowledge renders claim 14 obvious.

15. Claim 15

195. Claim 15 recites, “[t]he method of claim 14, wherein said body fluid sample is a blood plasma sample.”

196. I incorporate my analysis of claim 14 by reference here. As explained above, Leary utilizes samples isolated from bodily fluids like plasma.

197. Accordingly, Leary in view of a POSA's knowledge renders claim 15 obvious.

16. Claim 16

198. Claim 16 recites, "[t]he method of claim 14, wherein said tissue sample is a tissue biopsy sample."

199. I incorporate my analysis of claim 14 by reference here. As explained above, Leary utilizes samples from tissue biopsies.

200. Accordingly, Leary in view of a POSA's knowledge renders claim 16 obvious.

17. Claim 17.

18. Claim 18

201. Claim 17 recites, "[t]he method of claim 1, wherein said whole genome sequencing produces at least 750,000,000 sequencing reads."

202. Claim 18 recites, "[t]he method of claim 1, wherein said whole genome sequencing produces at least 1,000,000,000 sequencing reads."

203. I incorporate my analysis of claim 1 by reference here. Leary discloses that its whole-genome sequencing was to "an 18-fold physical coverage of the human genome." Ex. 1002 (Leary) at 1. As explained in Section V.C.2 ("Whole Genome Sequencing"), above, sequencing to greater depths of coverage was routine by 2013, in particular using the Illumina sequencers that had become the industry

standard. Moreover, a POSA would have been motivated to sequence to greater depths of coverage in the context of cancer detection and monitoring to increase the accuracy and reduce the error rate of the assay.

204. A POSA would have reasonably expected success in sequencing to greater depths, given that it involves, at most, repeating experiments that were already performed successfully. *See* Ex. 1028 (“[I]t is very easy to increase the coverage or sequence depth, if you later decide you need more data. Provided you still have your original sample, you can just sequence more . . .”). And because this predictable step would increase the likelihood of success of the overall experiment by increasing the method’s accuracy, a POSA would have been further motivated to do so, especially given as sequencing times and costs decreased rapidly between 2010 and 2013.

205. From the coverage depth formula, one can calculate the number of reads N by the formula:

$$N=(C*G)/L$$

where C is the depth of coverage, G is the length of the genome (3 billion bases) and L is the read length. Thus, POSA would have conducted whole-genome sequencing that produces 1,000,000,000 sequencing reads when sequencing to a depth of coverage of 30x using an average read length of 90 bp or less. By way of example, using Leary’s 25 bp read length and sequencing to a depth of coverage of 30x would

have resulted in 3,600,000,000 sequencing reads. In short, determining the optimal number of sequencing reads for a given application would have required no more than routine optimization of predictable and well-known sequencing techniques.

206. Accordingly, Leary in view of a POSA's knowledge renders claims 17 and 18 obvious.

19. Claim 19

207. Claim 19 recites:

The method of claim 1, further comprising:

- (i) repeating said third sequencing assay one or more times on one or more subsequent samples to yield an additional result comprising a fourth nucleic acid sequence; and
- (ii) generating an additional biomedical report that includes biomedical information of said subject, which biomedical information identifies the presence or absence of said one or more polymorphisms identified by said additional result.

208. I incorporate my analysis of claim 1 by reference here. I particularly emphasize my analysis of steps 1(e)-1(g), which in addition to rendering the steps leading to a third sequencing assay and generation of a report at that time point obvious, renders obvious collecting samples, generating subsets of nucleic acid molecules using PCR primer capture probes, and sequencing the resultant PCR products at four, five, six or more additional timepoints as well. Leary discloses and/or renders obvious repeating the third sequencing assay one or more times, as well as generating additional biomedical reports that include biomedical information

identifying the presence or absence of polymorphisms of interest. Specifically, as explained above, Leary discloses reanalyzing the patient's biomedical status (e.g., cancer progression) at multiple time points following the first sequencing assay, including by monitoring tumor-associated nucleic acid levels in multiple samples collected at different subsequent time points and plotting them on a biomedical report to monitor that status. *See* Ex. 1002 at Fig. 4B.

209. Accordingly, Leary in view of a POSA's knowledge renders claim 19 obvious.

B. Ground 2: Leary (Ex. 1002) in view of Ley (Ex. 1012) and the knowledge of a POSA.

210. As I explain in Ground 1, above, Leary in view of the knowledge of a POSA renders claims 1-11 and 13-19 obvious. Leary in view of Ley renders claims 1, 4, and 5 obvious for the following, additional reasons.⁶

1. Claim 1

211. I incorporate by reference by analysis of claim 1 as set forth in Ground 1, above.

⁶ Accordingly, the combination of Leary, Ley, and the knowledge of a POSA renders claims 1-11 and 13-19 obvious. Claims 2-3, 6-11, and 13-19 are obvious in view of the combination of Leary and Ley based on the disclosures of Leary, as discussed in Ground 1 above.

212. In Ground 1, I explain that Leary satisfies claim step 1(c)(ii), which requires “said capture probes hybridize to one or more polymorphisms which are based on or extracted from one or more databases of polymorphisms, *observed in a population of one or more samples*, or a combination thereof,” by designing PCR primer capture probes that hybridize to polymorphisms observed in a population of at least one sample collected from a patient.

213. Leary in view of Ley further renders obvious alternative embodiments within the scope of the claim step 1(c), including selecting different types of polymorphisms, such as SNPs and indels, as well as polymorphisms “based on or extracted from one or more databases of polymorphisms” whether alone or in combination with the polymorphisms Leary identified in its samples.

214. Specifically, Ley performed targeted resequencing on mutations comprising SNPs and indels that were present in genes of a patient’s tumor sample. Ex. 1012 (Ley) at 69. Table 2 describes the mutations detected and analyzed in Ley, as well as their effect on the relevant genes:

Table 2 | Non-synonymous somatic mutations detected in the AML sample

Gene	Consequence	Type	Solexa tumour reads WT:variant	Solexa skin reads WT:variant	Conservation score of mutant base	Mutations in other AML cases*
CDH24	Y590X	Nonsense	9:9	16:0	0.998	0/187
SLC15A1	W77X	Nonsense	15:12	19:0	1.000	0/187
KNDC1	L799F	Missense	7:8	20:0	NA	0/187
PTPRT	P1235L	Missense	9:13	16:0	1.000	0/187
GRINL1B	R176H	Missense	15:10	14:0	NA	0/187
GPR123	T38I	Missense	11:11	13:0	NA	0/187
EBI2	A338V	Missense	7:12	18:2	1.000	0/187
PCLKC	P1004L	Missense	19:9	15:1	0.98	0/187
FLT3	ITD	Indel	18:12	8:0	NA	51/185
NPM1	CATG ins	Indel	36:6	33:0	NA	43/180

Ins, insertion; WT, wild type.
* Patient cohort defined in ref. 23.

215. Moreover, given that genes comprise about 1% of the total genome, sequencing or assaying regions of the genome that contain genes would have been particularly useful for generating a meaningful biomedical report. Thus, as in Ley, a POSA would have been motivated to focus on identified polymorphisms that occur within the coding regions of proteins. Ex. 1012 at 66 (“[W]e restricted our downstream analysis to the coding sequences of annotated genes We also detected previously described, recurrent somatic insertions in the *FLT3* and *NPM1* genes.”). Similarly, given that the genome encodes 25,000 genes, sequencing or assaying a subset of these genes or their known variants that have a recognized biological effect may increase the likelihood of providing information relevant to a patient’s treatment, such as information that would aid the generation of a biomedical report that provides meaningful information on prognosis, diagnosis, and the patient’s response to particular therapies.

216. In addition to targeting polymorphisms that were identified in the sample, a POSA also would have found it obvious to target additional polymorphisms with capture probes using the same technique, including polymorphisms derived from a database of known variants. As explained in Section V.B (“Genetic Mutations”), a POSA would have been familiar with multiple publicly accessible databases that catalogued cancer-associated mutations. A POSA would have been motivated to include such mutations in a panel of tumor

biomarkers, in addition to the patient-specific rearrangements discussed above, to provide additional actionable information “to generate truly useful ‘clinically actionable’ reports.” Ex. 1014 (Ross 2011) at 527. *See* Section V.F (“Providing Actionable Information to Healthcare Providers”). For example, by targeting mutations that were known to have a diagnostic or therapy that was based on the presence or absence of a particular polymorphism or mutation, a POSA would have been motivated to include capture probes for such mutations in their biomarker panel to facilitate treatment. *See* Ex. 1014 (Ross 2011) at 527 (PLX-4032 drug candidate targeting cancers “that feature the V600E mutation of the *BRAF* gene,” and crizotinib having “impressive efficacy for the treatment of NSCLCs that feature an *EML4-ALK* translocation”).

217. Ley further confirms that a POSA would have been motivated to include known mutations in a panel of polymorphisms for serial monitoring at multiple time points, and further confirms that a POSA would have had a reasonable expectation of success in successfully monitoring such polymorphisms, as Ley successfully did so. Specifically, Ley detected via whole-genome sequencing, and then monitored using probe-based capture techniques similar to those employed by Leary, polymorphisms that were found in “the coding sequences of annotated genes.” Ex. 1012 (Ley) at 66; *see also id.* at 66, 70 (discussing desirability of “directed re-sequencing”); 69 (comparing allele frequencies for re-sequenced

polymorphisms at time of primary tumor sampling and at relapse). The genes monitored in Ley “affect genes participating in several well-described pathways that are known to contribute to cancer pathogenesis,” and included “well-known AML-associated mutations, including an internal tandem duplication of the *FLT3* receptor tyrosine kinase gene, which constitutively activates kinase signaling, and portends a poor prognosis, and a four-base insertion in exon 12 of the *NPM1* gene (NMPc). Both of these mutations are common (25%-30%) in AML tumours, and are thought to contribute to progression of the disease rather than to cause it directly.” *Id.* at 70. A POSA would have been motivated to track both known and previously unknown mutations, because monitoring only based on known polymorphisms in a “hypothesis-driven (for example, candidate gene-based) examination of tumour genomes,” may not provide a full picture, and “will miss key mutations.” *Id.*

2. Claims 4-5

218. Claim 4 recites, “[t]he method of claim 1, wherein said one or more polymorphisms are in a set of genes with known biomedically interpretable variants.” Claim 5 depends from claim 4 and recites, “[t]he method of claim 4, wherein said set of genes with known biomedically interpretable variants comprise one or more genes associated with cancer.”

219. I incorporate my analysis of Claim 1 by reference here, which explains why, in view of Leary, Ley, and the knowledge of a POSA, a POSA would have

found it obvious to capture polymorphisms detected within cancer-associated genes with biomedically interpretable variants to conduct targeted re-sequencing assays to conduct personalized monitoring of a patient's cancer.

220. Accordingly, Leary in view of Ley and a POSA's knowledge renders claims 4 and 5 obvious.

C. Ground 3: Leary (Ex. 1002) in view of Chan (Ex. 1008), Liao (Ex. 1009) and the knowledge of a POSA.

221. Claims 1-19 would have been obvious based on the combination of Leary, Chan, Liao, and the knowledge of a POSA.

222. Leary in view of the knowledge of a POSA renders the challenged claims obvious. As explained above, Leary discloses performing an initial whole-genome sequencing step to identify mutations (polymorphisms) of interest in a tumor sample, and then interrogating those mutations serially by detecting them in the patient's blood at subsequent time points. To detect the mutations in the blood, Leary discloses generating subsets of nucleic acids from blood samples by contacting them with PCR primer capture probes that hybridize to the nucleic acid sequences harboring the mutations and renders obvious sequencing those subsets of nucleic acids in a sequencing assay. *See* Ground 1, claim 1.

223. Given, among other factors, the rapid pace of technological advancement in the application of NGS techniques between Leary's publication in

2010 and the filing date of the '394 patent in January 2013, one skilled in the art also would have found it obvious to modify Leary's protocol by, following an initial whole-genome sequencing step, generating subsequent subsets of nucleic acid molecules by contacting blood or plasma samples with solution-phase hybridization probes that target (capture) the additional types of polymorphisms taught in Chan, including SNVs/SNPs and CNVs, followed by next-generation sequencing of those subsets. A POSA would have found it obvious to use such techniques either as a replacement for Leary's PCR primer capture probes, or as an adjunct to them, given, for example, that solution-phase hybridization probes can be optimal for identifying and detecting mutations like SNPs. *See* Mertes (Ex. 1006), Section V.C.3, above.

224. Accordingly, and as explained in further detail below, a POSA would have been motivated to utilize solution-phase hybridization probes in Leary's workflow to generate subsets of nucleic acids, followed by NGS to sequence those subsets to detect the presence or absence of a polymorphism of interest, as claimed in steps 1(c)-1(f). Moreover, a POSA would have reasonably expected success in using such techniques, given that like PCR primer capture probes, solution-phase hybridization probes had been repeatedly demonstrated to successfully enrich target sequences of interest from plasma samples in clinical applications, were commercially available, and could readily and predictably facilitate the sequencing of target nucleic acids using any number of readily available NGS platforms.

225. The use of solution-phase hybridization followed by NGS to serially track tumor-associated polymorphisms in the blood following an initial whole-genome sequencing step is expressly contemplated by Chan. Ex. 1008.

226. Chan explains that “[t]he presence of tumor-derived DNA in the plasma of cancer patients offers exciting opportunities for the detection and monitoring of cancer.” *Id.* at 211. It explains that “[m]ost of the previously published work on plasma DNA as a cancer marker has focused on the detection of specific and predetermined molecular targets known to be associated with cancer by means of such methods such as the PCR.” *Id.* “With the advent of massively parallel sequencing (MPS),” Chan continues, “several groups have incorporated this approach for developing new plasma DNA-based cancer markers.” *Id.* Describing and citing Leary (Ex. 1002) and McBride (Ex. 1015)⁷, Chan states: “One approach is to use MPS on tumor samples to first identify specific genomic rearrangements that can subsequently be detected in plasma.” *Id.*

227. Chan explains that approaches like those of Leary and McBride have a “targeted nature,” and thus may “provide only a partial glimpse of the tumor genome in the plasma of cancer patients.” *Id.* at 212. Chan then proposes an approach in which a “genomewide view of the tumor genome in the circulation” is obtained by

⁷ Chan’s cited reference 13 is Leary and its cited reference 14 is McBride. *See* Ex. 1008 at 223.

using “a nontargeted random—or shotgun—sequencing approach.” *Id.* Thus, the article reports on “the use of shotgun MPS to obtain a noninvasive, genomewide view of cancer-associated copy number variations and mutations in DNA in plasma.” *Id.* Using DNA analyzed from tumor samples and from plasma taken at multiple timepoints, including before and after surgery to resect the tumor, Chan observed that its approach “would allow genomic aberrations harbored by tumor cells within a patient to be scanned at different levels of resolution, ranging from a global view of cancer-associated copy number aberrations to point mutations carried by different clones of tumor cells.” *Id.* Chan concluded that the methods would “allow qualitative and quantitative aspects of such aberrations to be followed serially, as evidenced by the. . . correlation between the fractional tumor DNA concentration measured and the size of the tumor.” *Id.* at 221. Thus, Chan provides another example of techniques utilizing whole-genome sequencing to identify mutations of interest, and to track them over time in order to monitor a patient’s cancer. *Id.* at 222 (“It was the aim of this study to develop a noninvasive approach to assess the tumor burden of the body.”); *see also id.* (“careful analysis of the sequencing data permits one to dissect out the presence of DNA in the plasma originating from individual tumor cell populations. Targeting genomic regions exhibiting copy number aberrations specific to the breast and ovarian cancers

allowed the elucidation of the relative contributions of these tumors to the circulating DNA pool.”).

228. Chan also expressly taught using solution-phase hybridization-based capture probes that target particular polymorphisms identified in an initial whole-genome sequencing assay to serially monitor and detect cancer in a patient’s plasma over time. As discussed further below, Chan’s disclosures would have motivated a POSA to employ solution-phase hybridization probes, including solution-phase hybridization probes targeting additional types of polymorphisms, followed by NGS in Leary’s methods.

1. Claim 1

229. As I have explained above, a POSA would have been motivated to utilize solution-phase hybridization capture probes as the claimed “capture probes” in steps 1(c) and 1(e), followed by NGS as the claimed “second sequencing assay” in step 1(d) and as the claimed “third sequencing assay” in element 1(e). Below I describe how a POSA would have implemented the obvious combination of Leary, Chan, and Liao.

(i) Step 1(a)

230. In implementing the combination, a POSA would have performed step 1(a) in the manner disclosed in Leary, as set forth in Section VIII.A(a) (Ground 1, Step 1(a)), above. Chan discloses similar methods of “generating a first subset of

nucleic acid molecules from a first nucleic acid sample,” specifically by obtaining nucleic acids from, for example, “a plasma sample upon presentation of a cancer patient...” Ex. 1008 (Chan) at 223; *see also id.* at 211 (“DNA was extracted from tumor tissues, and the preoperative and postoperative plasma samples of these patients were analyzed with shotgun MPS.”). As with Leary’s tumor sample, a POSA would have understood that extracting DNA from a plasma sample would have result in the collection of less than all of the DNA that was present in the sample, which satisfies step 1(a) under Patent Owner’s understanding of the term.

(ii) Step 1(b)

231. In implementing the combination, a POSA would have performed step 1(b) in the manner disclosed in Leary, as set forth in Section VIII.A(b) (Ground 1, Step 1(b)), above. Chan also discloses element 1(b) as it teaches “conducting a first sequencing assay on said first subset of nucleic acid molecules to yield a first result comprising a first nucleic acid sequence wherein said first sequencing assay comprises whole genome sequencing” by “using shotgun sequencing of a plasma sample upon presentation of a cancer patient.” Ex. 1008 (Chan) at 223; *see also id.* (“[W]e have demonstrated the performance of cancer genome scanning via shotgun sequencing of plasma DNA.”); *id.* at 211 (“In this study, we explored the use of shotgun massively parallel sequencing (MPS) of plasma DNA from cancer patients to scan a cancer genome noninvasively.”).

(iii) Step 1(c)

232. In implementing the combination, a POSA would have performed step 1(c) in the manner disclosed in Leary, as set forth in Section VIII.A(c) (Ground 1, Step 1(c)), above, with the following modifications. In view of Chan, a POSA would have performed step 1(c) by using solution-phase hybridization probes targeting polymorphisms such as SNPs and CNVs, in addition to the PCR primer capture probes targeting rearrangements taught by Leary, or as an alternative to such PCR primer capture probes, depending on, e.g., the needs of a particular patient and the types of polymorphisms the patient's cancer exhibits as identified in the initial whole genome sequencing step. Chan itself motivates this combination, which would provide an even greater "numerical advantage" in the cancer-related polymorphisms to be monitored, thereby increasing the sensitivity of the overall assay. Ex. 1008 (Chan) at 222 ("[T]he detection of mutations—passengers and drivers—in a genome-wide manner provides a numerical advantage that has the potential for more sensitive, precise, and representative detection of tumor DNA in plasma."). Chan highlighted the desirability of monitoring SNVs, which would have motivated a POSA to analyze them in a serial monitoring protocol: "[M]ost of the tumor-associated SNVs we detected were 'passenger' rather than 'driver' mutations. . . . We therefore believe that the detection of mutations—passengers and drivers—in a genome-wide manner provides a numerical advantage that has the potential for more

sensitive, precise, and representative detection of tumor DNA in plasma.” *Id.* at 222. Additionally, Chan noted that “selective analyses of driver mutations may allow more readily the identification of medically actionable targets,” and “[c]onsequently, the nonselective genome-wide approach can be combined in a synergistic manner with the more conventional detection strategies for selective targets to achieve better cancer management.” *Id.* at 222-23. Thus, a POSA would have been motivated to add to the fundamental methodology set forth in Leary the analysis of the additional types of polymorphisms (e.g., SNVs/SNPs and CNVs) that Chan found to be present in greater numerical quantities in patients’ cancer samples. *Id.* (Table 3) (monitoring group of 2,216 SNVs). Adding the monitoring of such polymorphisms to Leary’s methods would also have further motivated the use of solution phase hybridization probes in addition to, or as a replacement for, Leary’s PCR primer capture probes, depending on the patient.

233. A POSA would have been motivated to implement such techniques for multiple reasons, including to aid the detection and monitoring of a larger number of polymorphisms in parallel, and to adapt the capture technique to the types of mutations identified in the whole-genome sequencing step, as taught by Chan. A POSA would have had a reasonable expectation of success in using solution phase hybridization probes, including given the reports of success reported in Liao and

Gnirke, among many other references successfully employing solution phase hybridization capture probes before 2013.

234. Specifically, Chan recognized that there were limitations with the main approach disclosed in its article including that “the shotgun sequencing approach we have described is relatively expensive compared with many targeted approaches.” *Id.* at 223. Accordingly, Chan recognized that there would be benefits associated with combining “the nonselective genome-wide approach . . . in a synergistic manner with the more conventional detection strategies for selective targets to achieve better cancer management.” *Id.* at 222-23; *see also id.* at 222 (“[T]he detection of mutations—passengers and drivers—in a genome-wide manner provides a numerical advantage that has the potential for more sensitive, precise, and representative detection of tumor DNA in plasma.”).

235. Accordingly, Chan expressly motivates first conducting whole-genome sequencing – as in its experimental protocols, where such sequencing is performed on an initial plasma sample, as well as in prior art protocols like Ley (Ex. 1012), Leary (Ex. 1002), and McBride (Ex. 1015), each of which performed an initial whole-genome sequencing step on a tumor sample – and then using that information to target specific mutations identified through whole-genome sequencing at subsequent timepoints with capture probes, followed by sequencing of the captured nucleic acids. As Chan observed, such an approach allows the identification of many

relevant mutations, increasing the chances of identifying actionable information: “a numerical advantage that has the potential for more sensitive, precise, and representative detection of tumor DNA in plasma.” *Id.* at 222.

236. In light of this clear motivation to serially monitor specific target mutations following whole-genome sequencing, Chan concluded with the following suggested methodology:

Given that plasma DNA in cancer patients consists of a mixture of tumor and nontumor DNA, with the former being present as a minority population in many patients, one needs to perform relatively deep sequencing to obtain the necessary analytical power. Thus, as currently implemented, the shotgun sequencing approach we have described is relatively expensive compared with many targeted approaches (8). With the continuing and, thus far, rapid reduction in sequencing costs, cost is not expected to be as much of an obstacle in the near future. *One can also envision the possibility of using shotgun sequencing of a plasma sample upon presentation of a cancer patient, mining the shotgun sequencing data of plasma DNA for copy number aberrations and SNVs, and then specifically targeting such sequences for serial monitoring purposes. In this regard, solution-phase hybridization-based capture approaches have already been successfully used to enrich plasma DNA for noninvasive prenatal diagnosis (34, 35). Hence, we see that the targeted and shotgun approaches can be used in combination for cancer detection and monitoring.*

Id. at 223 (emphasis added). Thus, Chan expressly taught using solution-phase hybridization-based capture probes that target particular polymorphisms identified in an initial whole-genome sequencing assay to serially monitor and detect cancer in a patient's plasma over time. A POSA would have been motivated to utilize such capture probes following the initial step of performing a whole-genome sequencing assay in Leary's protocol as well, given the significant overlap in methodology, including to improve detection and monitoring of CNVs and SNVs identified via whole-genome sequencing, as well as to enable the serial monitoring of a larger number of targets by using multiplexed sequencing of the hybrid capture products using NGS.

237. By 2010, those skilled in the art were familiar with high throughput sequencing platforms in which capture probes were utilized to selectively amplify or enrich a nucleic acid sequence of interest. For example, a review of the state of Next-Generation sequencing applications published in 2010 observed that “[a]lthough whole-genome sequencing analysis may be soon feasible for the clinical laboratory from a technical perspective, targeted analysis of specific genomic regions may be preferable in order to answer a specific clinical question.” Ex. 1004 (Anderson 2010) at 48

238. As explained in Section V.C(b) (“Solution-Based Hybrid Capture”), above, by the time Chan was published, a POSA would have understood how to

implement and use such techniques to enrich targets before analyzing them with NGS. *See* Ex. 1005 (Gnirke 2009). To summarize the steps of this process, first, a pool of capture probes (or “baits”) consisting of nucleic acid fragments that target a particular nucleic acid sequence (i.e., the complementary nucleic acid sequence that the bait’s sequence pairs with by hybridization) is prepared. These capture probes/baits are typically prepared to include a ligand, such as biotin, that allows them to be bound by beads during subsequent steps to facilitate extraction. Second, the prepared baits are contacted with the nucleic acid sample (the “pool”), and the baits bind to the sequences of interest. This results in a “catch” of nucleic acids bearing the sequences of interest (i.e., sequences complementary to the baits), which are “caught” through the hybridization of strand-to-strand base-pairing. The catch is then pulled down or fished out using beads that bind the biotin that was incorporated into the baits, and the catch can then be PCR amplified before analysis with Next-Generation Sequencing. Ex. 1005 (Gnirke) at 183; Fig. 1 (reproduced below). *See also* 1006 (Mertes 2011) at 374 (“By selective recover[y] and subsequent sequencing of genomic loci of interest, costs and efforts can be reduced significantly compared with whole-genome sequencing.”); 375 (“Targeted enrichment can be useful in a number of situations where particular portions of a whole genome need to be analyzed.”).

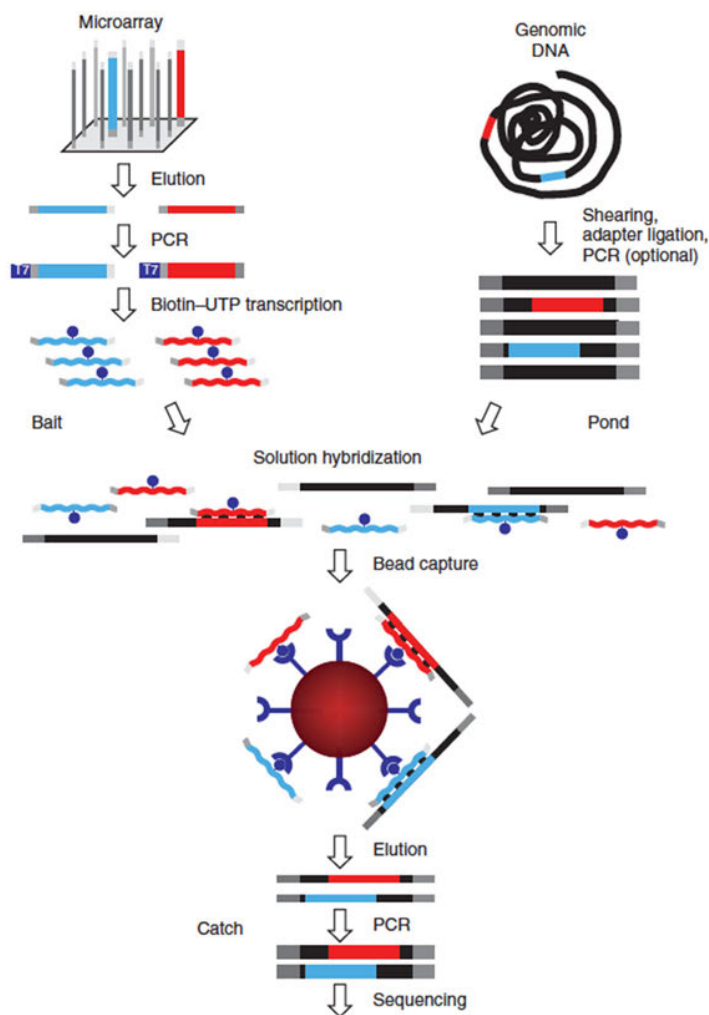


Figure 1 Overview of hybrid selection method. Illustrated are steps involved in the preparation of a complex pool of biotinylated RNA capture probes (bait; top left), whole-genome fragment input library (pond; top right) and hybrid-selected enriched output library (catch; bottom). Two sequencing targets and their respective baits are shown in red and blue. Universal adaptor sequences are gray. The excess of single-stranded nonself-complementary RNA (wavy lines) drives the hybridization.

239. As of 2011 and thus well before the priority date, “an increasing number of protocols and vendors [had] begun offering out of the box solutions for hybrid capture, meaning, the researcher need not do development work but merely choose between a preset targeted enrichment regions (e.g. whole exome) or specify their

own custom enrichment region.” Ex. 1006 (Mertes) at 379. Thus, using the sequence information derived from Leary’s whole genome sequencing step, a POSA would have been readily able to generate a custom enrichment region for the donor-derived polymorphisms of interest, including by seeking such a kit from a vendor, such as SureSelect from Agilent.

240. In the context of the tumor-surveillance assays disclosed by Leary and Chan, PCR followed by sequencing (as in Leary) and solution-phase hybridization capture followed by NGS (as in Chan) would have provided similar information about the patient’s cancer. Specifically, each approach would predictably result in information regarding the presence or absence of a particular nucleic acid sequence in a sample at a given timepoint, as well as its relative abundance.

241. A POSA would have reasonably expected success in using solution-phase hybridization capture probes to identify and detect polymorphisms of interest in plasma samples from a cancer patient, given that the technique was well-known and could be implemented from readily available commercial sources. *See* Section V.C(b) (“Solution-Based Hybrid Capture”), above. Moreover, Chan’s citation of Liao (Ex. 1009, Chan reference No. 34) would have further supported a POSA’s reasonable expectation of success. Specifically, Chan observed that the ability to detect and monitor cancer noninvasively using a patient’s plasma samples was based on the presence of cfDNA in the blood: “plasma DNA in cancer patients consists of

a mixture of tumor and nontumor DNA.” Ex. 1008 (Chan) at 223. Thus, Chan noted that Liao had already established that “solution-phase hybridization-based capture approaches have already been successfully used to enrich plasma DNA for noninvasive prenatal diagnosis.” *Id.* The success of using the technique does not depend on whether the target sequence is from cancer or from a fetus; accordingly, Liao’s success in enriching cell-free-fetal DNA from plasma would have furthered a POSA’s reasonable expectation that using solution phase hybrid capture probes followed by NGS to identify and detect a nucleic acid carrying a polymorphism of interest would be successful. *See* Ex. 1009 (Liao) at 92 (Background) (“In this study, we investigated [the] applicability in enriching selected genomic regions from plasma DNA and the quantitative performance of this approach.”); *id.* (Methods) (“DNA was extracted from plasma samples collected from 12 pregnant women carrying female fetuses. The SureSelect Target Enrichment System (Agilent Technologies) was used to enrich for exons on chromosome X. Plasma DNA libraries with and without target enrichment were analyzed by massively parallel sequencing.”); *Id.* (Conclusions) (“Targeted sequencing of maternal plasma DNA permits efficient and unbiased detection of fetal alleles at genomic regions of interest and is a powerful method for measuring the proportion of fetal DNA in a maternal plasma sample”).

(iv) Step 1(d)

242. In implementing the combination, a POSA would have performed step 1(d) in the manner disclosed in Leary, as set forth in Section VIII.A(d) (Ground 1, Step 1(d)), above, with the following modifications. In view of Chan, a POSA would have performed step 1(d) by using next-generation sequencing to conduct the second sequencing assay on second subset of nucleic acid molecules generated through the use of solution-phase hybridization probes. Liao reported success using its similar targeting strategy, observing that this approach “increase[d] the representation of the targeted region from 0.1% to 21.3% of the total sequenced reads.” Ex. 1009 (Liao) at 98. *See also id.* at 92 (“The SureSelect Target Enrichment System (Agilent Technologies) was used to enrich for exons on chromosome X. Plasma DNA libraries with and without target enrichment were analyzed by massively parallel sequencing.”); *id.* at 100 (“Because of the potential of targeted sequencing to save costs and increase throughput, this approach may play an important role in the translation of plasma DNA sequencing to molecular diagnostics, with impacts on such fields as ... cancer diagnostics ...”). For the reasons set forth in the discussion of element 1(c) above, a POSA would have been motivated to implement such techniques for multiple reasons, including to aid the detection and monitoring of a larger number of polymorphisms in parallel, and to take advantage of the faster speeds and lower sequencing costs of next-generation sequencing as compared to

Sanger sequencing. A POSA would have had a reasonable expectation of success in using next-generation sequencing to conduct sequencing assays on enriched populations of nucleic acids generated with solution phase hybrid capture probes for multiple reasons, including given the reports of success reported in Liao and Gnirke, among other references.

(v) Step 1(e)

243. In implementing the combination, a POSA would have performed step 1(e) in the manner disclosed in Leary, as set forth in Section VIII.A(e) (Ground 1, Step 1(e)), above, with the following modifications. In view of Chan, a POSA would have performed step 1(e) by using solution-phase hybridization probes in addition to the PCR primer capture probes taught by Leary, or as an alternative to such PCR primer capture probes, depending on the particular experiment. For the reasons set forth in the discussion of element 1(c), above, a POSA would have been motivated to implement such techniques for multiple reasons, including to aid the detection and monitoring of a larger number of polymorphisms in parallel, and to tune the capture technique to the types of mutations identified in the whole-genome sequencing step, as taught by Chan. A POSA would have had a reasonable expectation of success in using solution phase hybrid capture probes, including given the reports of success reported in Liao and Gnirke, among other references.

(vi) Step 1(f)

244. In implementing the combination, a POSA would have performed step 1(f) in the manner disclosed in Leary, as set forth in Section VIII.A(f)(Ground 1, Step 1(f), above, with the following modifications. In view of Chan, a POSA would have performed step 1(f) by using next-generation sequencing to conduct the third sequencing assay on the third subset of nucleic acid molecules generated through the use of solution-phase hybridization probes. For the reasons set forth in the discussion of element 1(c), above, a POSA would have been motivated to implement such techniques for multiple reasons, including to aid the detection and monitoring of a larger number of polymorphisms in parallel, and to take advantage of the faster speeds and lower sequencing costs of next-generation sequencing as compared to Sanger sequencing. A POSA would have had a reasonable expectation of success in using next-generation sequencing to conduct sequencing assays on enriched populations of nucleic acids generated with solution phase hybrid capture probes for multiple reasons, including given the reports of success reported in Liao and Gnirke, among other references.

(vii) Step 1(g)

245. In implementing the combination, a POSA would have performed step 1(g) in the manner disclosed in Leary, as set forth in Section VIII.A(a) (Ground 1, Step 1(g)), above.

2. Claims 2-11, 13-19

246. In Ground 1, I explain why Leary in view of the knowledge of a POSA renders challenged claims 1-11 and 13-19 obvious. In this ground, I explain that a POSA would have further found it obvious to specifically monitor polymorphisms such as SNPs and CNVs through the use of solution-phase hybridization probes followed by next generation sequencing in steps 1(c) (use of capture probes to generate a second subset of nucleic acids), 1(d) (use of NGS to sequence second subset), 1(e) (use of capture probes to generate an additional subset), and 1(f) (use of NGS to sequence additional subset), in addition to or instead of Leary's PCR capture probes and Sanger sequencing in those steps. Thus, Leary in view of Chan and Liao renders obvious claims 2-11 and 13-19, for the same reasons that Leary in view of the knowledge of a POSA, as I set forth in Ground 1.

3. Claim 3

247. Claim 3 recites, "[t]he method of claim 1, wherein said one or more polymorphisms comprise one or more insertions, deletions, structural variant junctions, variable length tandem repeats, single nucleotide mutations, or a combination thereof."

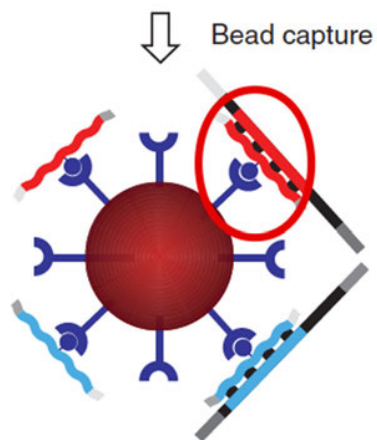
248. I incorporate my analysis of Claim 1 by reference here. The additional polymorphisms analyzed in the combination include SNPs ("single nucleotide mutations.")

249. Accordingly, Leary in view of Chan, Liao and a POSA's knowledge further renders claim 3 obvious.

4. Claim 12

250. Claim 12 recites, "[t]he method of claim 1, wherein said plurality of capture probes are conjugated to beads."

251. I incorporate my analysis of claim 1 here. Leary in view of Chan and Liao discloses conjugating the plurality of capture probes to beads when the hybrid capture probes are bound to magnetic beads to select the catch, as routine in the use of hybrid capture. Specifically, Liao explains that "[a]fter hybridization, we selected captured targets by pulling down the biotinylated probe/target hybrids with streptavidin-coated magnetic beads (Dynabeads M-280 Streptavidin; Invitrogen) and purified the targets with the MinElute PCR Purification Kit (Qiagen)." Ex. 1009 (Liao) at 94. A POSA thus would have understood that solution-phase hybridization probes will be conjugated to beads as part of the well-known enrichment process. *See* Ex. 1005 (Gnirke) at 183 (Fig. 1) ("Bead capture" step annotated below with red circle marking streptavidin on surface of a magnetic bead conjugated to biotin on capture probe).



252. Accordingly, Leary in view of Chan and Liao renders claim 12 obvious.

IX. Conclusions

253. For the reasons set forth above, a POSA would have found claims 1-19 of the '394 patent obvious.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Dated: November 30, 2022

By:  Digitally signed by
John Quackenbush
Date: 2022.11.30
17:50:49 -05'00'

John Quackenbush, Ph.D.