

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

FORESIGHT DIAGNOSTICS INC.,
Petitioner,

v.

PERSONALIS, INC.,
Patent Owner.

Case Nos. IPR2024-00170
U.S. Patent No. 11,584,968

DECLARATION OF HENRY MORRICE FURNEAUX, PH.D.

TABLE OF CONTENTS

	Page No.
I. ENGAGEMENT.....	1
II. BACKGROUND AND QUALIFICATIONS.....	2
A. Patents Awarded.....	7
B. Research And Teaching Experience	8
C. Industry Experience.....	12
D. Professional Society Involvement.....	12
III. MATERIALS CONSIDERED	12
IV. LEGAL STANDARDS	13
A. Anticipation	13
B. Nonobviousness.....	14
V. THE PERSON OF ORDINARY SKILL IN THE ART	15
VI. CLAIM 1 OF THE '968 PATENT	16
VII. CLAIM CONSTRUCTION OF “WHOLE GENOME SEQUENCING”	17
VIII. CHAN’S MPS IS NOT WHOLE GENOME SEQUENCING.....	25
IX. CHAN DOES NOT NECESSARILY AND INEVITABLY TEACH EXTRACTING NUCLEIC ACID MOLECULES FROM LEUKOCYTES	36
X. THE CLAIMED INVENTION IS NONOBVIOUS OVER CHAN	40
A. The data burden of whole genome sequencing	40
B. Whole genome sequencing was more expensive than other approaches	
46	
XI. THE CLAIMED INVENTION IS NONOBVIOUS OVER LEY	48
A. Limitation 1(a)(i)-(ii).....	48

TABLE OF CONTENTS
(cont'd.)

	Page No.
1. Ley counsels against Foresight's proposed modification	49
2. Dr. Quackenbush's references do not provide the motivation to modify Ley	51
a. Meyerson	52
b. Chapman	52
c. Lee	55
d. Ding	56
B. Limitation 1(c).....	58
XII. LACK OF ANTICIPATION AND OBVIOUSNESS BASED ON OKOSUN-II	61
A. Limitation 1(a)(ii).....	61
B. Limitation 1(c).....	67
XIII. CERTIFICATION	69

1. I, Henry Morrice Furneaux, Ph.D., make this Declaration in connection with *inter partes* review proceedings at the U.S. Patent & Trademark Office concerning U.S. Patent No. 11,584,968 (“the ‘968 patent”).

2. I previously submitted declarations in connection with the related *inter partes* review proceeding nos. IPR2023-00224 concerning U.S. Patent No. 11,384,394 (“the ‘394 patent”), IPR2023-00317 concerning U.S. Patent No. 11,408,033 (“the ‘033 patent”), IPR2023-00545 concerning U.S. Patent No. 11,299,783 (“the ‘783 patent”), and IPR2023-00546 concerning U.S. Patent No. 10,450,611 (“the ‘611 patent”).

3. I am over the age of twenty-one (21) and am competent to make this Declaration. I reside at 146 Kenyon Street, Hartford, CT 06105.

I. ENGAGEMENT

4. I have been retained by counsel for Personalis, Inc. (“Personalis”) in the above-captioned IPR matter as an independent technical expert to review and evaluate claims of the ‘968 Patent. My opinions are based on my skills, knowledge, training, education, and experience in matters of this nature, and my examination of the materials used in preparing this report.

5. I am being compensated on an hourly basis for my work performed in connection with this matter. I have received no additional compensation for my

work in this matter, and my compensation does not depend upon the contents of this report, any testimony I may provide, or the ultimate outcome of these matters.

II. BACKGROUND AND QUALIFICATIONS

6. In general, I have knowledge and experience with both the general scientific contexts relevant to this proceeding and specific details of the technologies employed. Based upon my experience, qualifications and expertise, I am qualified as an expert in the fields of nucleic acid biochemistry, the hybridization of DNA to RNA, DNA recombination, the isolation of human genes, the de novo sequencing and assembly of human DNA sequence, the use of PCR and DNA sequencing to identify mutations in human tumor tissue, the manipulation of gene expression by siRNAs and microRNAs, the mechanism of action of enzymes such as DNA ligase and DNA polymerase that are used in molecular biology protocols and the identification of SNPs in noncoding regions of the genome that regulate gene expression via microRNA.

7. I have summarized in this section my educational background, career history, relevant publications, and other relevant qualifications. My full curriculum vitae has been submitted as Exhibit 2010 in this proceeding.

8. I earned a Bachelor of Science in Biochemistry, with honors, from the University of Aberdeen (Scotland, United Kingdom) in 1975. This department was headed by Professor Hamish Keir, a student of Professor J.N Davidson, who wrote

the classic textbook “The Nucleic Acids”. Thus, this experience provided a rigorous introduction to nucleic acids.

9. I earned a Ph.D. in Biochemistry, also from the University of Aberdeen, UK in 1978. My thesis work was on the structure and function of poly-ADP ribose, a novel nucleic acid that I discovered had novel effects upon RNA transcription. See Furneaux HM, Pearson CK: The effects of NAD⁺ on RNA synthesis in isolated nuclei from BHK cells. *Biochem Soc. Trans*, 6:753-755, 1978.

10. In 1978, in a seminar given by Professor Ed Southern, I heard about the exciting hypothesis that mature eukaryotic mRNA might be spliced from a precursor mRNA. This mechanism was unknown and so I accepted a position in Jerry Hurwitz’s lab in New York, to try to establish a system that could recapitulate mRNA splicing in vitro. I was very fortunate to be trained by Dr. Hurwitz who not only discovered many of the enzymes used in Molecular biology but also provided a very rigorous training environment. Indeed, several of Dr. Hurwitz’s trainees either won Nobel prizes, became members of the National Academy or attained comparable levels of achievement. We were amongst the first to recapitulate this reaction and the first to show that mRNA splicing required the action of a ribonucleoprotein complex. See Furneaux HM, Perkins KK, Freyer GA, Arenas J, Hurwitz J: Isolation and characterization of two fractions from HeLa cells required for mRNA splicing in vitro. *Proc Natl. Acad. Sci*, 82: 4351-4355, 1985.

11. From 1984 to 1987, I was a Research Fellow in the Program in Molecular Biology at Sloan Kettering Institute, the experimental research arm of Memorial Sloan Kettering Cancer Center in New York City. From 1987 to 1989, I was promoted to an Assistant Laboratory Member in the Program in Molecular Pharmacology and Therapeutics. In this position, I pioneered the use of molecular biology techniques to address problems in Oncology. For example, my lab was amongst the first to perform PCR analysis on DNA from human Cerebral Spinal Fluid (CSF) to identify a cancer-causing virus. See Henson J, Rosenblum M, Armstrong D, Furneaux H: Amplification of JC virus DNA from brain and cerebrospinal fluid of patients with progressive multifocal leucoencephalopathy. *Neurology* 41:1967-1971, 1991. My lab also pioneered the recovery of DNA from archival paraffin embedded human tumor slides. My lab showed that this DNA could be sequenced and we were amongst the first to identify recurrent mutations in the p53 gene, that could be linked to benzopyrene. See Schlegel U, Rosenfeld M, Voldanandt M, Rosenblum M, Furneaux HM: p53 Mutations in primary lung tumors are conserved in brain metastasis. *J Neuro-Oncology* 14: 93-100, 1992.

12. From 1989 to 1995, I was promoted to Assistant Member and Head of the Laboratory of Molecular Neuro-Oncology at Sloan Kettering. Concurrently, I was also an Assistant Professor of Neuroscience at the Cornell University Graduate School of Medical Sciences (now known as the Weill Cornell Graduate School of

Medical Sciences). In this phase of my career, my lab focused on the isolation of the genes that encoded human tumor antigens. This was an innovative project, since the conventional dogma at that time was that the immune system did not recognize tumors. My lab was the first to clone the human paraneoplastic tumor antigen genes, HuD, HuR, LEMS and CDR62. In addition, the cloned tumor antigens provided an ELISA assay for the early detection of lung cancer and neurological disease, which was patented and licensed to Quest Diagnostics. These studies involved cDNA library expression cloning, extensive DNA sequence analysis, gene expression by RT-PCR and an innovative immunohistochemical technique that we devised to directly visualize the expression of tumor antigens and antigen-reactive T cells in human tumor tissue. *See Szabo A, Dalmau J, Manley G, Rosenfeld M, Wong E, Henson J, Posner JB, Furneaux HM: HuD a paraneoplastic encephalomyelitis antigen contains RNA binding domains and is homologous to Elav and Sex-Lethal. Cell 67: 325-333, 1991; Fathallah-Shaykh H, Wolf S, Wong E, Posner JB, Furneaux HM: Cloning of a leucine zipper protein recognized by the sera of patients with antibody-associated paraneoplastic cerebellar degeneration. Proc Nat Acad. Sci, 88:3451-3454, 1991.*

13. From 1996 to 2000, I was promoted to Associate Member. Concurrently, I was also an Associate Professor of Neuroscience at the Cornell University Graduate School of Medical Sciences. In this phase, my lab continued

to focus on the structure and function of the HuD and HuR tumor antigens and demonstrated their role in regulation of gene expression via their interaction with the AU-rich elements in the 3'UTR of mRNA. *See* Chung S, Jiang L, Cheng S, and Furneaux HM*. Purification and properties of HuD, a neuronal RNA binding protein. J. Biol. Chem. 271:11518-11524, 1996. It is important to note that mutation of these noncoding elements is a driver of carcinogenesis.

14. In 2000, I accepted the position of Associate Professor of Molecular, Microbial and Structural Biology at the University of Connecticut School of Medicine. Beginning in 2003 and through 2009, I was also the Director of the Molecular Biology and Biochemistry Graduate Program at the University of Connecticut School of Medicine. In this phase, my lab continued studies on HuR, and HuD. My lab also pursued studies on the structure and function of microRNAs. This included innovative studies on the regulation of gene expression by SNPs present in microRNA target sites in non-coding regions. *See* Jensen KP, Covault J, Conner TS, Tennen H, Kranzler HR, Furneaux HM. A common polymorphism in serotonin receptor 1B mRNA moderates regulation by miR-96 and associates with aggressive human behaviors. Molecular Psychiatry 2009 14(4):381-9.

15. In 2009 I left academia to create a company (Clonegene LLC) that could generate income to fund my basic science research. Clonegene is a for-profit company that designs, generates and distributes advanced reagents used in

biochemistry research. The income generated has permitted me to pursue, without constraint, novel independent research.

16. I am also concurrently an adjunct professor in the Department of Biology at Capital Community College in Hartford, Connecticut. I have held this position since 2011. This outreach activity allowed me to teach advanced concepts in molecular biology to a community that had traditionally been underexposed to cutting-edge biology.

A. Patents Awarded

17. I am a named co-inventor on at least the following six (6) U.S. Patents:

- U.S. Patent No. 5,603,934 (Antigen recognized by patients with antibody associated paraneoplastic sensory neuronopathy, issued Feb. 1997);
- U.S. Patent No. 5,614,371 (Ri fusion antigen recognized by antibodies associated with paraneoplastic opsoclonus and methods of use thereof, issued Mar. 1997);
- U.S. Patent No. 5,668,013 (Antigen recognized by patients with antibody associated paraneoplastic cerebellar degeneration, DNA encoding same and uses thereof, issued Sept. 1997);

- U.S. Patent No. 5,807,705 (Antigen recognized by patients with antibody associated paraneoplastic sensory neuronopathy, DNA encoding same and uses thereof, issued Sept. 1998);
- U.S. Patent No. 5,925,526 (Antigen recognized by patients with antibody associated cerebellar degeneration, DNA encoding same and uses thereof, issued July 1999); and
- U.S. Patent No. 6,193,948 (Antigen recognized by patients with antibody associated paraneoplastic sensory neuropathy, DNA encoding same and uses thereof, issued Feb. 2001).

B. Research And Teaching Experience

18. Over the course of my research career, I am the recipient of over \$3.75 million in public and outside funding, including from the American Chemical Society, Brain Tumor Society, Burke Prostate Cancer Fund, Connecticut Department of Public Health, National Institute of Health, and National Science Foundation. I am an author or contributor to at least eighty (80) publications in the areas of RNA biochemistry, mRNA stability, enzyme mechanisms, human genomics, RNA therapeutics, immunology, biochemistry, molecular biology, and cell biology as it relates to human diseases and developmental issues. A significant number of these publications are recognized as very innovative in their field. Overall, my publications have earned an h-index citation value of 55 (Google

Scholar indicates that an h-index of 20 is good, 40 is outstanding, and 60 is truly exceptional.)

19. Among my editorial work, I am: Associate Editor and member of the editorial board of BMC Biochemistry; and an ad hoc reviewer for the Journal of Biological Chemistry, Molecular and Cellular Biology, European Molecular Biology Organization Journal, Nucleic Acids Research, RNA Journal, and Journal of Clinical Investigation.

20. Among my recent federal study section service, I participated in the U.S. Department of Defense's 2007 Breast Cancer Research Program – Molecular Biology & Genetics Study Section, and 2009 Breast Cancer Research Program – Concept – Molecular Biology & Genetics Study Section; and the National Institute of Health's October 2007 KO8 – KO2 Awards – NHLBI Study Section, February 2008 Oncological Sciences (F09) Study Section, June 2008 Oncological Sciences (F09) Study Section, July 2008 KO8 KO2 Awards – NHLBI Study Section, and June 2009 Oncological Sciences (F09) Study Section.

21. Further, I am co-author or co-reviewer of several book chapters, including:

- Leslie Pick, Henry M. Furneaux & Jerard Hurwitz, Purification and characterization of wheat germ ligase and associated activities, 181 Methods Enzymology, 18-36 (1990).

- Jerome B. Posner & Henry M. Furneaux, Paraneoplastic neurological syndromes, 68 Proc. Ass'n. Rsch. Nervous Mental Disease, 187-219 (1990).
- Henry M. Furneaux, Neuronal nuclear auto antibodies, Type 1 (Hu), in Auto antibodies, Elsevier Science, 551-54 (J.B. Peter and Y. Shoenfeld eds., 1996).
- B. Joseph & Henry M. Furneaux, cDNA Libraries and Neurological Disease, 55(6) Archives Neurology, 785-88 (1998).
- Sung-Suk Chae, Ji-Hye Paik, Jonathan Schubert-Coleman, & Henry Furneaux, Timothy Hla: Application of the RNA Interference (RNAi) Technology to Angiogenesis Research, Methods in Endothelial Cell Biology - Springer Lab, (H. Augustin ed., 2004).

22. Among my educational and teaching experience, I was: Designer and Director of the Molecular Basis of Neurological Disease, a course for MD/PhD students, at the Cornell University Graduate School of Medical Sciences (now known as the Weill Cornell Graduate School of Medical Sciences) (1990-2000); Designer and Director of the Molecular Basis of Disease, a course for PhD and MD/PhD students, at the University of Connecticut Medical School (2001-2009); Course Director of the Molecular Biology and Biochemistry Program Journal Club at the University of Connecticut Medical School (2001-2006); Lecturer in

Biochemistry I, a first year biochemistry course for PhD students, at the University of Connecticut Medical School (2002-2006); Lecturer in Organ Systems II, a first year medical school course, at the University of Connecticut Medical School (2003-2006); Lecturer in outreach discovery lectures to encourage research careers for high school students (2003-2009); Co-designer and teacher of The Logic of Biology, a new discussion course for all incoming first year students to introduce them to experimental logic and contemporary problems in biology, at the University of Connecticut (2006-2009); and Lecturer in HB3, a first year medical school course, including design of a new introductory module that combines the fields of vascular biology, hematology and immunology, at the University of Connecticut Medical School (2007-2009).

23. Among my recent major scholastic administrative tasks, I was: Director of the Graduate Program in Biochemistry and Molecular Biology at the University of Connecticut Medical School (2003-2009); member of the executive committee of the Department of Molecular, Microbial and Structural Biology at the University of Connecticut Medical School (2003-2009); member of the Graduate School Planning Committee at the University of Connecticut Medical School (2004); first year graduate student advisor at the University of Connecticut Medical School (2004-2007); and member of the Scientific Advisory Committee – General

Clinical Research Center at the University of Connecticut Medical School (2005-2009).

C. Industry Experience

24. I am currently the Chief Executive Officer of Clonogene LLC (“Clonogene”) in Hartford, Connecticut. I have held this position since 2009. Clonogene is a for-profit company that designs, generates and distributes quality reagents used in Biochemistry research. These include monoclonal antibodies, RNA binding proteins, enzymes and synthetic mRNA. In particular, I use my skills to design novel research reagents based on my understanding of the needs of contemporary researchers.

D. Professional Society Involvement

25. Among the scientific societies and honors I have earned are: Member of the New York Academy of Science; Member of The Harvey Society; Member of The RNA Society; Holder of the Neil Levitan Chair in Neuro-Oncology at SKI, of the MSKCC (1998-2000); and cited in “Who’s Who in America.”

III. MATERIALS CONSIDERED

26. In forming my opinions, I have reviewed the following documents and materials:

- The '968 patent;

- The declaration submitted by Dr. John Quackenbush in this proceeding (Exhibit 1100) as well as exhibits cited in the Quackenbush declaration; and
- Materials cited in this declaration.

27. This pre-trial institution declaration does not specifically respond to each and every statement and opinion within the Quackenbush declaration, but that does not mean I necessarily agree with the statements and opinions within the Quackenbush declaration to which I have not specifically responded. If the Board institutes trial in this proceeding, I reserve the right to further respond to the Quackenbush declaration.

IV. LEGAL STANDARDS

28. I am not a patent attorney. I have been informed of the following legal standards that apply in this matter.

A. Anticipation

29. I have been informed that anticipation (lack of novelty) is established only when it is proven that a single prior art reference discloses each and every limitation of the claimed invention, either explicitly or inherently, as the limitations are arranged in the claim. I have also been informed that a single prior art reference may be deemed to “inherently anticipate” a claim if and only if a limitation that is not expressly stated in the reference is necessarily present in the product or process

that the reference describes, not merely probably or possibly present in that product or process.

B. Nonobviousness

30. I understand that a patent claim is invalid as obvious 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 to a person of ordinary skill in the art at the time of the invention. I understand that an obviousness analysis considers the following factors: (a) the scope and content of the prior art; (b) the differences between the prior art and the claims; (c) the level of ordinary skill in the art; and (d) any objective indicia of non-obviousness (such as unexpected results, skepticism in the field, industry praise, and commercial success). I understand that such objective indicia of nonobviousness must have a connection or “nexus” to the claims. I further understand that where the objective evidence is tied to a specific product that embodies the claimed invention and is coextensive with them, the objective evidence is entitled to a rebuttable presumption of nexus.

31. In connection with obviousness, I understand there should have been some reason that would have led a person of ordinary skill in the art to combine or modify the prior art to obtain the claimed invention. Furthermore, there must have been a reasonable expectation that a person of ordinary skill in the art would have succeeded in obtaining the claimed invention based on the prior art.

V. THE PERSON OF ORDINARY SKILL IN THE ART

32. Dr. Quackenbush has proposed that as of October 2014 a person of ordinary skill in the art (POSA):

[W]ould 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.

EX1100 (Quackenbush Declaration) ¶36.

33. I understand that the Board adopted Dr. Quackenbush's definition of the person of ordinary skill in the art in the related matters of *Foresight Diagnostics Inc. v. Personalis, Inc.*, IPR2023-00224, IPR2023-00317, IPR2023-00545, and IPR2023-00546. I have applied this definition of a person of ordinary skill in the art in forming the opinions described in this declaration. However, I understand

Personalis reserves the right to offer a competing definition in this proceeding if the Board institutes trial.

VI. CLAIM 1 OF THE '968 PATENT

34. Claim 1 is the only independent claim challenged by Foresight, and it recites four steps (a)-(d) as shown below:

Claim 1. A method of analyzing a biological sample obtained from a subject being screened for a cancer, wherein the biological sample is from a non-fetal source, the method comprising:

(a) obtaining a first set of sequence reads and a second set of sequence reads by sequencing nucleic acid molecules derived from a first set of nucleic acid molecules and a second set of nucleic acid molecules, wherein said first set of sequence reads corresponds to said first set of nucleic acid molecules and said second set of sequence reads corresponds to said second set of nucleic acid molecules, wherein said first set of sequence reads and said second set of sequence reads are generated using whole genome sequencing, and wherein (i) said first set of nucleic acid molecules is extracted from tissue sample of said subject, and (ii) said second set of nucleic acid molecules is extracted from leukocytes from a blood sample;

(b) using said first set of sequence reads and said second set of sequence reads to identify mosaic variants specific to said tissue sample of said subject in said first set of sequence reads that is not present in said second set of sequence reads;

(c) subsequent to (b), identifying said mosaic variants specific to said tissue sample of said subject in a first additional sample of said

subject, wherein said first additional sample is obtained from a source that is different than said tissue sample and at a time point that is different than said tissue sample; and

(d) providing a report based on an identification of said mosaic variants specific to said tissue sample of said subject in said first additional sample, wherein said report comprises predicting, diagnosing, and/or prognosing a status or outcome of a cancer in said subject.

VII. CLAIM CONSTRUCTION OF “WHOLE GENOME SEQUENCING”

35. Limitation 1(a) of the '968 patent recites, in part, “said first set of sequence reads and said second set of sequence reads are generated using *whole genome sequencing*.”¹

36. The term “whole genome sequencing” has taken on different meanings in different contexts in the literature. I have reviewed the claim language and the specification of the '968 patent to determine the meaning of “whole genome sequencing” in the context of this patent.

37. The claim language includes the term “whole genome sequencing.” The term contains the word “whole,” which modifies the word “genome” to indicate that the entirety—not a part or subset—of the genome must be sequenced.

¹ All emphases supplied unless otherwise noted.

38. The specification distinguishes methods that sequence the “whole genome” from “targeted genome sequencing” that only seek to sequence a portion of the genome, such as the “exome.” *See, e.g.*, EX1001 (the '968 patent) at 3:22-29 (“In some embodiments, the sequencing comprises whole genome sequencing. In some embodiments, the sequencing comprises targeted genome sequencing.”); *id.* at 4:31-38 (“In some embodiments, the sequencing comprises whole genome sequencing. In some embodiments, the sequencing comprises targeted genome sequencing. In some embodiments, the sequencing comprises untargeted genome sequencing. In some embodiments, the untargeted genome sequencing comprises whole genome sequencing. In some embodiments, the targeted sequencing includes at least one mutational hotspot.”); *id.* at 17:36-64 (comparing “whole genome sequencing” and sequencing “over an entire exome (e.g., whole exome sequencing)”); *id.* at 25:46-52 (“For example without limitation, in some cases, the combined whole exome sequencing and a whole genome sequencing reactions may increase the sensitivity or specificity in detecting one or more copy number variations or diagnosing a specific condition when compared to the sensitivity or specificity of whole exome sequencing alone”).

39. The “exome” is the part of the genome containing the exons. *See, e.g.*, EX1001 (the '968 patent) at 30:19-22 (“The exome is often the part of the genome formed by exons. The exome may be formed by untranslated regions (UTRs), splice

sites and/or intronic regions. The entire exome or portion thereof may comprise one or more exons of a protein coding gene.”).

40. The specification consistently refers to “whole genome sequencing” as a separate category of sequencing than “exome sequencing.” EX1001 (the '968 patent) at 15:34-37 (referring to “exome or whole genome sequencing”); *id.* at 17:67-18:1 (“In some cases, sequencing specific genomic regions may comprise using exome sequencing.”).

41. The specification also distinguishes “whole genome sequencing” from “other methods” that sequence only a portion of the genome, such as “deep sequencing over a fraction of the genome.” EX1001 (the '968 patent) at 17:36-49 (“The methods as disclosed herein may comprise using whole genome sequencing. In some cases, the whole genome sequencing is used to identify mosaic variants in a person. The methods may further comprise sequencing the mosaic variant using other methods. In some cases, sequencing the mosaic variant may comprise deep sequencing over a fraction of the genome. For example, the fraction of the genome may be at least about 50; 75; 100; 125; 150; 175; 200; 225; 250; 275; 300; 350; 400; 450; 500; 550; 600; 650; 700; 750; 800; 850; 900; 950; 1,000; 1100; 1200; 1300; 1400; 1500; 1600; 1700; 1800; 1900; 2,000; 3,000; 4,000; 5,000; 6,000; 7,000; 8,000; 9,000; 10,000; 15,000; 20,000; 30,000; 40,000; 50,000; 60,000; 70,000; 80,000; 90,000; 100,000 or more bases or base pairs.”).

42. The specification further differentiates “whole genome sequencing” from methods that sequence sub-genomic regions. For example, the specification explains that “genomic regions” can be selected based on various traits, and “may comprise an entire genome or a portion thereof.” EX1001 (the '968 patent) at 28:40-41 (“The genomic region features may comprise an entire genome or a portion thereof.”). Alternatively, “genomic regions” may comprise “an entire exome or a portion thereof,” “one or more sets of genes,” or “one or more sets of regulatory elements.” *Id.* at 28:41-48 (“The genomic region features may comprise an entire exome or a portion thereof. The genomic region features may comprise one or more sets of genes. The genomic region features may comprise one or more genes. The genomic region features may comprise one or more sets of regulatory elements. The genomic region features may comprise one or more regulatory elements.”).

43. The specification further acknowledges that whole genome sequencing may encompass “low mean coverage regions” and “zero mean coverage regions.” EX1001 (the '968 patent) at 28:57-60 (“The genomic region features may comprise one or more low mean coverage regions from whole genome sequencing (WGS), zero mean coverage regions from WGS, validated compressions, or a combination thereof.”); *id.* at 32:37-46 (“The genomic region features may comprise one or more low mean coverage regions from whole genome sequencing (WGS), zero mean coverage regions from WGS, validated compressions, or a combination thereof. The

low mean coverage regions from WGS may comprise regions generated from Illumina v3 chemistry, regions below the first percentile of Poission distribution based on mean coverage, or a combination thereof. The Zero mean coverage regions from WGS may comprise regions generated from Illumina v3 chemistry.”).

44. The specification additionally refers to “low coverage whole genome sequencing,” which a person of ordinary skill in the art would understand constitutes its own category of sequencing rather than a subclass within whole genome sequencing. For example, the specification refers to “low coverage whole genome sequencing.” EX1001 (the '968 patent) at 26:6-13 (“In some cases, the methods as disclosed herein comprise untargeted sequencing data (e.g., a low coverage whole genome sequencing data) that may have a sensitivity, specificity, positive predictive value or negative predictive value that is less than 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 7 5%, 80%, 85% or 90% for one or more SNV.”); *see also id.* at 25:64-67 (“In some cases, the methods as described herein comprise combining an untargeted sequencing data (e.g., low coverage whole genome sequencing data) and one or more target-specific sequencing data.”). The same paragraph of the specification separately refers to “whole genome sequencing.” *See e.g., id.* at 26:18-25 (“In some cases, the untargeted sequencing (e.g., whole genome sequencing) in the methods as disclosed herein may have a sensitivity, specificity, positive predictive value or negative predictive value that is

less than 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85% or 90% for one or more SNV, one or more polymorphisms or one or more specific genes or genomic regions.”).

45. In addition, the literature recognizes the differences between “low coverage whole genome sequencing” (lcWGS) and “whole genome sequencing.” Before I discuss the literature, it is important to discuss the term “sequence coverage” since it can be used in different ways in different contexts. In the present context, I will use “sequence coverage” to refer to the extent to which sequence reads successfully align to the sequence of the genome under study. For example, in the present context, one would expect a “whole genome sequence” analysis to yield reads that align to all or nearly all of the segments of the human genome. This is typically established by the alignment of the sequence reads to a contemporary reference genome. This sequence coverage alignment analysis should not be confused with a planning event, sometimes called “coverage” that more accurately is described as sequence “depth,” in which an investigator plans a sequencing project by contemplating how many reads (which is a function of machine capacity and cost) would be needed to correctly identify each base throughout an entire genome. The depth of coverage is a simple arithmetical calculation which compares the number of bases that could be sequenced (the number of reads times the length of the read) to the size of the genome by, e.g., the Lander/Waterman equation $C=L*(N/G)$, where

C stands for depth of coverage, G is the haploid genome length, L is the read length and N is the number of reads. If the number of bases that could be sequenced is less than the number of bases in a genome, then one can very confidently predict that the sequencing project will not reveal all or nearly all of the sequence of the genome. For example, where the number of reads times the length of the reads was 30 billion bases, then one might be able to run a depth of approximately 10X on a human genome of 3.2 billion base pairs. If the sequencing technology were completely efficient, one might imagine that a sequence depth of 1X should be sufficient to correctly identify each base in the genome. However, it has been known, since the time of Sanger sequencing, that not all bases are equally identifiable. Thus, the current industry standard for many sequencing applications of normal human cells is to plan a potential 30X depth to efficiently identify the sequenceable bases. However, it is critical to point out that such an excess of sequencing power does not guarantee that all or nearly all of the segments of the genome will be covered. Indeed, if one continues to increase the sequence depth, one may simply be endlessly resequencing the same bases, whose identity has already been well established. The determination of sequence coverage achieved by a sequencing method includes the post-sequencing alignment analysis described above. Thus, sequence coverage is not predicted by the sequencing depth; the sequencing coverage is determined by a post-sequencing alignment analysis.

46. Literature recognizes lcWGS as an approach that “sacrifice[es] sequencing depth for greater genome coverage and larger sample size with the help of probabilistic statistical strategies.” See “Low-coverage Whole Genome Sequencing,” available at <https://www.cd-genomics.com/resource-lcwgs.html> (EX2191). “Low-coverage genome sequencing refers to the sequencing of a genome at a relatively low depth, typically resulting in incomplete coverage of the entire genome.” *Id.* “The LcWGS pre-processing process is similar to WGS, but an important difference is the need to use genotypic probabilities to explain the probability of genotypic uncertainty, such as downstream analysis using the site frequency spectrum (SFS).” *Id.* Although lcWGS could be used in an untargeted context, it can also be used in targeted approach. On the other hand, an untargeted approach is axiomatic to the success of the WGS objective. Further, in lcWGS there is no intention, nor indeed any expectation, to identify all or almost all of the bases throughout an entire genome.

47. I understand that the Board previously construed the term “whole genome sequencing” to mean “a non-targeted approach to sequencing all or almost all of a genome” in the related matter of IPR2023-00224. Based on the claim language and specification of the '968 patent, it is my opinion that the same construction should be applied in this case.

VIII. CHAN'S MPS IS NOT WHOLE GENOME SEQUENCING

48. Dr. Quackenbush equates Chan's MPS with whole genome sequencing. See EX1100 (Quackenbush Declaration) ¶¶113, 118-120. I disagree that Chan's MPS is equivalent to whole genome sequencing. MPS is simply a broad term that refers to hundreds of thousands to millions or even billions of DNA molecules being sequenced at the same time (in parallel). See Nguyen et al., "Automation of molecular-based analyses: a primer on massively parallel sequencing," Clin Biochem Rev 35:169-176 (2014) at page 169 (EX2192) ("In MPS, hundreds of thousands to millions of DNA sequencing reactions [being] performed simultaneously, or 'in parallel', with each sequencing step being coupled with detection"). "MPS is also referred to as next-generation sequencing (NGS)." *Id.* at page 169. As another example, the literature teaches that "Massively parallel sequencing, [is] otherwise known as next generation sequencing (NGS)...." Jessri et al., "Harnessing massively parallel sequencing in personalized head and neck oncology," J. Dent Res 93:437-444 (2013) at page 437 (EX2193). To be clear, Chan's MPS would only constitute WGS if the reads arising from MPS were found by alignment analysis to cover all or almost all of the genome.

49. I note that Foresight's exhibits use MPS as an alternative name for NGS. See EX1014 (Ross) at page 527 ("The recent introduction of the more thorough and in-depth sequencing technologies (*next-generation sequencing*

[NGS] or massively parallel sequencing) has heightened concerns that the traditional so-called Sanger-type sequencing method used to sequence the human genome 10 years ago will prove incapable of scaling sufficiently to support the breadth of information required to manage an ever-growing arsenal of targeted anticancer drugs.”); EX1013 (Shendure) at page 1135 (entitled “Next-generation DNA sequencing” and at Abstract discussing “massively parallel DNA sequencing platforms”). In NGS/MPS each “read” is from a single nucleic acid molecule or copies of the same nucleic acid molecule. In contrast, Sanger sequencing involves averaging a signal from all nucleic acid molecules in the reaction.

50. MPS can be used for many different types of applications that are not whole genome sequencing. MPS can be used for sequencing whole exomes, targeted genes, targeted exons, confirmatory sequencing, reduced representation sequencing, transcriptome sequencing (RNA sequencing including mRNA, micro RNA, lncRNA and tRNA), metagenomic sequencing, RNA sequencing, nuclease fragmentation, and chromatin immunoprecipitation-sequencing (ChIP-Seq). For example, Nguyen teaches that “MPS is being increasingly used in many areas of medicine” and explains that MPS can be used to sequence at least eight types of nucleic acid targets as follows:

Depending on the purpose, different nucleic acid targets may be sequenced:

- Whole genome
- Whole exome
- Targeted genes
- Transcriptome -RNA sequence including mRNA, tRNA
- Mitochondrial genome
- Metagenome-microbial DNA in biological samples
- Methylome – DNA sequence including cytosine methylation state
- Chromatin immunoprecipitation-Sequencing (ChIP-Seq): regions of DNA that bind to transcription factors, co-factors or chromatin proteins.

EX2192 (Nguyen) at page 173.

51. Foresight’s Ross reference teaches that MPS/NGS has several “major uses” that include “whole exome sequencing,” “whole genome sequencing,” “SNP detection,” “targeted exon sequencing,” “confirmatory sequencing,” and “targeted sequencing projects” as follows:

Table 3
Commercial Next-Generation Sequencing Platforms*

	Basic Technique	Major Uses	Time Needed	Instrument Cost (\$, thousands)
Illumina	Flow cell-based, reversible dye termination, and 4-color optical imaging	Whole exome sequencing; whole genome sequencing; SNP detection	Intermediate, 4 d fragment; 9 d paired end	500-900
454 Pyrosequencing	Emulsion PCR with bead-based pyrosequencing and CCD light imaging	Targeted exon sequencing; confirmatory sequencing; SNP detection	Fast, <1 d	500-700
Helicos	Oligo-dT captured PolyA-tailed DNA fragments; flow cell 4-color dNTP optical imaging	Single molecule sequencing; whole genome sequencing	Slowest, 8 d	999
SOLiD	Sequential dinucleotide ligation; flow cell-based 4-color optical imaging	Whole exome sequencing; whole genome sequencing; SNP detection	Slow, 7 d fragment run; 14 d paired end	600-700
Ion Torrent	Semiconductor-based nonoptical detection; standard dNTP sequencing chemistry	Targeted sequencing projects not demanding deep sequencing	Fast, <1 d	50

CCD, charge-coupled device; dNTP, deoxynucleoside triphosphate; PCR, polymerase chain reaction; SNP, single nucleotide polymorphism.
* Adapted from Metzker.⁵⁵

EX1014 (Ross) at page 530, Table 3.

52. Foresight's Shendure reference teaches that MPS/NGS has at least eleven applications:

Table 2 Applications of next-generation sequencing		
Category	Examples of applications	Refs
Complete genome resequencing	Comprehensive polymorphism and mutation discovery in individual human genomes	44
Reduced representation sequencing	Large-scale polymorphism discovery	45
Targeted genomic resequencing	Targeted polymorphism and mutation discovery	46–52
Paired end sequencing	Discovery of inherited and acquired structural variation	53,54
Metagenomic sequencing	Discovery of infectious and commensal flora	55
Transcriptome sequencing	Quantification of gene expression and alternative splicing; transcript annotation; discovery of transcribed SNPs or somatic mutations	56–63
Small RNA sequencing	microRNA profiling	64
Sequencing of bisulfite-treated DNA	Determining patterns of cytosine methylation in genomic DNA	60,65,66
Chromatin immunoprecipitation–sequencing (ChIP-Seq)	Genome-wide mapping of protein-DNA interactions	67–70
Nuclease fragmentation and sequencing	Nucleosome positioning	69
Molecular barcoding	Multiplex sequencing of samples from multiple individuals	61,71

EX1013 (Shendure) at page 1141, Table 2.

53. Chan cites to Forshew (EX2194) when describing MPS as follows:

With the advent of massively parallel sequencing (MPS),⁸ several groups have incorporated this approach for developing new plasma DNA-based cancer markers.

EX1008 (Chan) at 211. Citation 8 in Chan is the Forshew reference (EX2194). Forshew describes a “targeted deep sequencing” method of certain genomic regions, which does not constitute whole genome sequencing. EX2194 (Forshew) at page 1. For example, Forshew describes its targeted deep sequencing method as follows:

To amplify and sequence fragmented DNA, we designed primers to generate amplicons that tile regions of interest in short segments of about 150 to 200 bases (Fig. 1A and table S1), incorporating universal adaptors at 5' ends (fig. S1).

EX2194 (Forshew) at page 2-3. Forshew provides an overview of its targeted deep sequencing method in Figure 1:

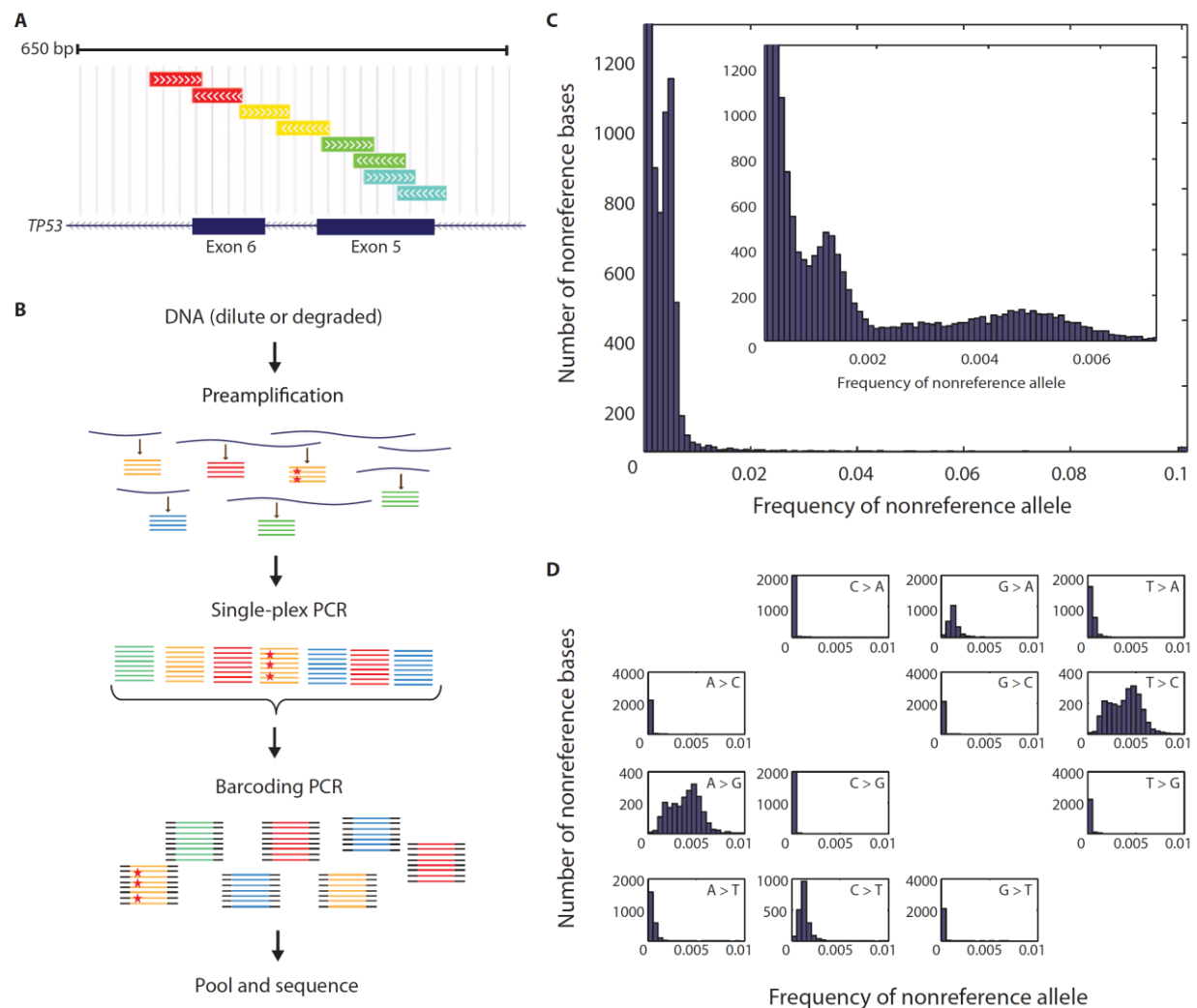


Fig. 1. Overview of tagged amplicon sequencing (TAm-Seq). **(A)** Illustration of amplicon design. Primers were designed to amplify regions of interest in overlapping short amplicons (table S1). Amplicon design is illustrated for a region covering exons 5 to 6 of *TP53*. Colored bars, segmented into forward and reverse reads, show regions covered by different amplicons (excluding primer regions). Sequencing adaptors are attached at either end, such that a single-end read generates separate sets of forward and reverse reads (fig. S1). Because amplicons are mostly shorter than 200 bp, the forward and reverse reads also partially overlap. Figure adapted from University of California, Santa Cruz, Genome Browser (<http://genome.ucsc.edu/>). **(B)** Workflow overview. Multiple regions were amplified in parallel. An initial preamplification step was

performed for 15 cycles using a pool of the target-specific primer pairs to preserve representation of all alleles in the template material. The schematic diagram shows DNA molecules that carry mutations (red stars) being amplified alongside wild-type molecules. Regions of interest in the preamplified material were then selectively amplified in individual (single-plex) PCR, thus excluding nonspecific products. Finally, sequencing adaptors and sample-specific barcodes were attached to the harvested amplicons in a further PCR. **(C)** Distribution of observed nonreference read frequencies, averaged over 47 FFPE samples, across all loci and all nonreference bases. Inset expands the low-frequency range. **(D)** Distribution of the observed background nonreference read frequencies averaged over 47 FFPE samples for the 12 different A/C/G/T base substitutions.

Id. at Figure 1.

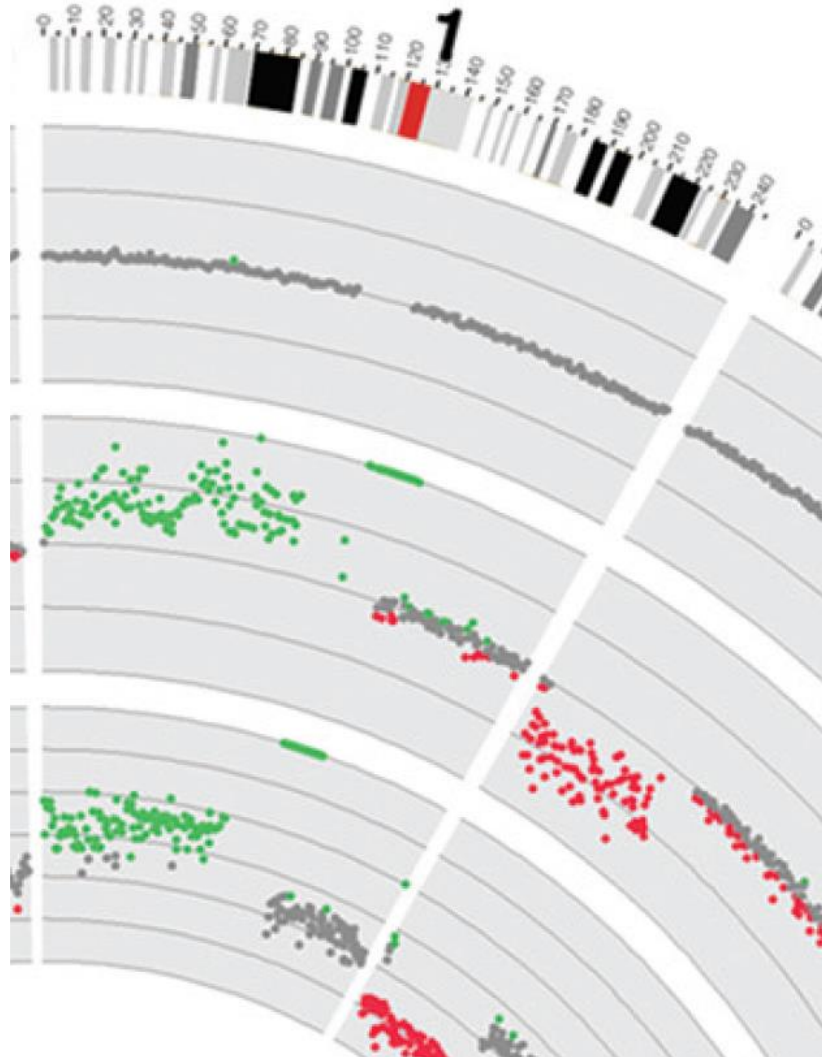
54. Dr. Quackenbush asserts that Chan describes its sequencing method as involving “shotgun MPS.” EX1100 (Quackenbush Declaration) ¶113. Chan’s mere use of a “shotgun” approach in the MPS sequencing protocol does not transform

Chan's MPS into whole genome sequencing. "Shotgun" MPS simply refers to MPS in which the first step is simply to break DNA into much smaller fragments. Subsequent steps may focus on specific targets within the genome. For example, Foresight's Gnirke reference performed "[t]argeting genomic loci by massively parallel sequencing," EX1005 at 182, using shotgun sequencing of the targeted loci, *id.* at 183-184. Although Gnirke fragmented the DNA in a presumably random manner, only some targeted fragments were sequenced. Thus, Gnirke's shotgun MPS is not whole genome sequencing. It sequenced targeted loci within the DNA fragment collection.

55. Chan does not report that its shotgun MPS sequenced all or almost all of the genome. Chan's assertion that it performed enough reads to provide an average **depth** of "29.5-fold" or "43-fold" does not mandate that its MPS sequencing provided sequence coverage across all or almost all of the genome. EX1008 (Chan) at page 215. Chan does not report the sequence coverage achieved by its sequence project. However, Chan did perform a low resolution chromosomal alignment analysis of its sequence reads, primarily to determine putative copy number changes. Chan's Figures 1 and 4 are "Circos" plots in which the sequencing reads resulting from the MPS sequencing data have been mapped to chromosomal locations. EX1008 (Chan) at page 215 ("Fig. 1 shows Circos plots").

56. First, it is evident that there was no analysis of reads aligning to chromosomes X and Y. Further, analysis of the alignment data reveals significant gaps in sequence coverage. The Circos plots show that large portions of sequencing data is missing from at least chromosomes 1, 9, 13, 14, 15, 16, 19, 21, and 22—and all data from chromosomes x and y are missing. *Id.* at Figs. 1, 4. Here, I have analyzed the outer ring which is the post-surgery plasma sample; it is difficult to clearly discern the gaps in the tumor sample, since the attendant copy number changes obscure the gaps. Importantly, Chan has no data regarding the buffy coat obtained from the blood of the HCC patients. Also, from the available information in Chan's Circos plots, it is evident that Chan did not sequence all or almost all of the genome.

57. For example, below is a zoomed-in excerpt from Chan Figure 1(A) showing that there are no sequencing reads mapped to chromosome 1 from approximately the 120,000,000 to 140,000,000 base positions for the HCC1 patient:



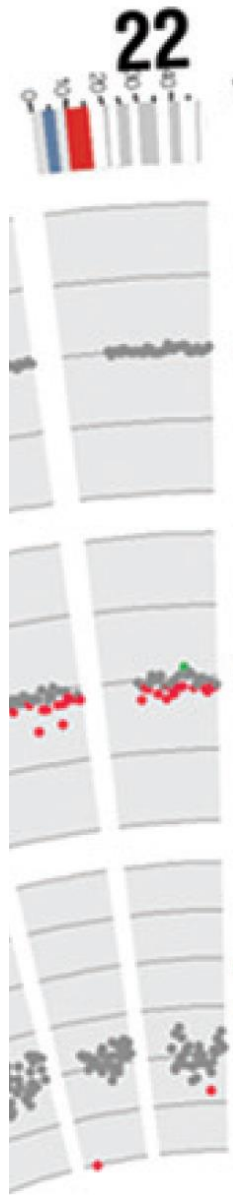
EX1008 (Chan) at 216, Fig. 1(A) (excerpted).

58. Below is a zoomed-in excerpt from Chan Figure 1(A) showing that there are no sequencing reads mapped to chromosome 9 from approximately the 40,000,000 to 70,000,000 base positions for the HCC1 patient:



EX1008 (Chan) at 216, Fig. 1(A) (excerpted).

59. Below is a zoomed-in excerpt from Chan Figure 1(A) showing that there are no sequencing reads mapped to chromosome 22 from approximately the first to 20,000,000 base positions for the HCC1 patient:



EX1008 (Chan) at 216, Fig. 1(A) (excerpted). As mentioned above, there are similarly large gaps in sequencing reads for portions at least chromosomes 1, 9, 13, 14, 15, 16, 19, 21, and 22, and all data from chromosomes x and y are missing for each of the patients' data in Chan Figures 1 and 4.

60. Finally, it is important to reiterate that Chan’s Circos plots provide a very low resolution analysis, since it involves the mapping of a very short read (50 or 75 bp) to a 1 Mbase bin. See EX1008 (Chan) at page 215 (“Copy number aberrations were analyzed in 1-MB windows across the genome in the tumor tissues and compared with the plasma samples from a group of 16 healthy control individuals.”); *id.* at page 212 (“For each paired-end read, 50 bp or 75 bp from each end was aligned to the non-repeat-masked reference human genome (Hg18).”). Thus, Chan’s level of sequence coverage may be significantly lower than that estimated by this low-resolution analysis.

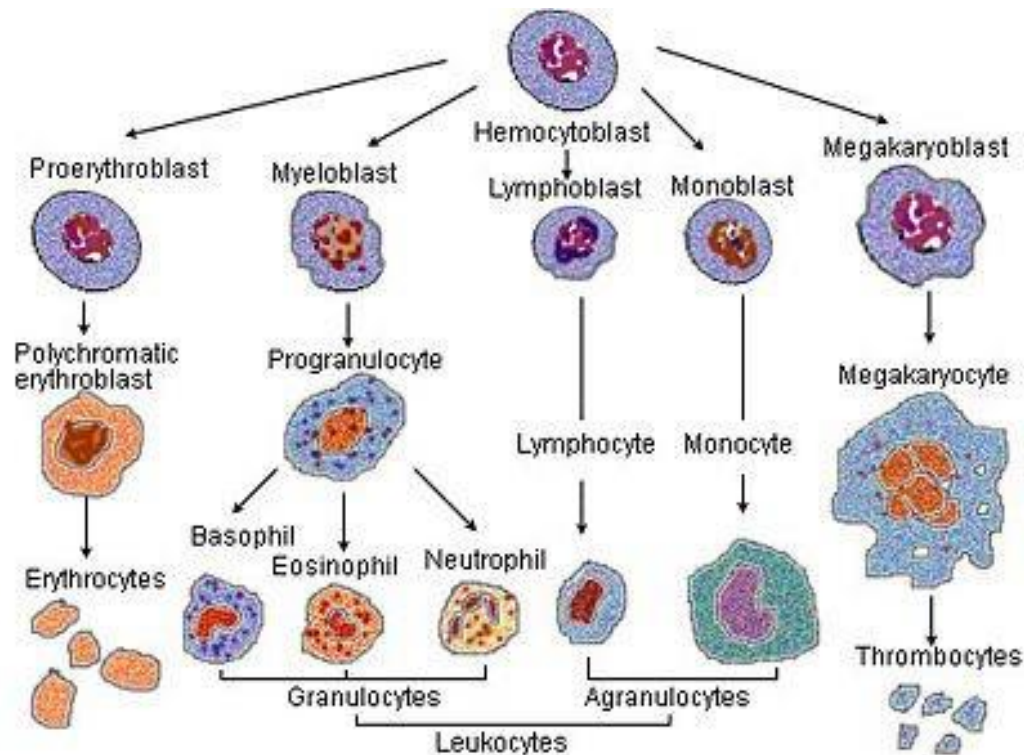
**IX. CHAN DOES NOT NECESSARILY AND INEVITABLY TEACH
EXTRACTING NUCLEIC ACID MOLECULES FROM
LEUKOCYTES**

61. Limitation 1(a)(ii) requires “said second set of nucleic acid molecules is extracted from *leukocytes* from a blood sample.”

62. Dr. Quackenbush assumes that the DNA extracted from Chan’s buffy coat sample must always, inevitably, and necessarily involve extraction of DNA from leukocytes because the buffy coat sample “contains leukocytes.” EX1100 (Quackenbush Declaration) ¶121. This assumption is incorrect.

63. Human blood contains a wide variety of cells in which the leukocyte family is only one component. Cells found in the blood arise from different lineages. For example, there is the hemocytoblast lineage which gives rise, via discrete

precursors, to erythrocytes, thrombocytes and the leukocyte family. The hemocytoblast lineage is depicted below:



64. There are also cells found in blood that arise from entirely different lineages such as the endothelial cell lineage and the mesenchymal lineage. There is an enormous catalogue of work that has been devoted to the purification of specific cell types from blood. The buffy coat technique is merely a starting point in these endeavors. In this technique, cells present in blood treated with anti-coagulant are resolved by centrifugation, which exploits the differing densities of blood components. Three layers result, a cell-free low-density layer called plasma, a high-density layer containing the red blood cells (erythrocytes), and a medium-density cell layer called the buffy coat which contains all the other blood cell types. This

technique was originally devised to purify red blood cells, and more selective techniques must be used to purify other cell types from the buffy coat layer.

65. Dr. Quackenbush apparently ignores the fact that a patient's buffy coat contains several types of nucleated non-leukocyte cells from which genomic DNA can be extracted. For example, a patient's buffy coat contains at least the following four types of nucleated non-leukocyte cells: (1) endothelial cells, (2) endothelial progenitor cells, (3) mesenchymal precursor cells, and (4) megakaryocytes and these are discussed below.

66. First, it is clear that **endothelial cells** (which do not derive from a hemocytoblast blast lineage) are present in the buffy coat layer of patients. Solovey "used immunohistochemical examination of buffy-coat smears to enumerate circulating endothelial cells." Solovey et al., "Circulating activated endothelial cells in sickle cell anemia," *The New England Journal of Medicine*, 1584-1590, 1997 at page 1584 (EX2196). Solovey determined that the "[m]ean ($\pm$ SD) numbers of circulating endothelial cells in normal blood donors, patients with sickle cell trait, and patients with hemolytic anemias not due to hemoglobin S were 2.6 ± 1.6 , 3.0 ± 2.6 , and 2.0 ± 0.8 per milliliter of whole blood, respectively." *Id.*

67. Second, this observation was validated and extended by other investigators. For example, Schmid provides a contemporary review of the field and also teaches that **endothelial progenitor cells** can be found in the buffy coat layer.

See Schmid et al., “Circulating Endothelial Progenitor Cells,” Chapter 8 in *Methods in Molecular Biology, Angiogenesis Protocols*, 2nd ed., 467:139-155 at pages 139, 146-148 (describing a protocol for isolating endothelial progenitor cells from the buffy coat) (EX2197).

68. Third, it is clear that **mesenchymal precursor cells** (which do not derive from a hemocytoblast blast lineage) are present in the buffy coat layer. Zvaifler teaches that “BMPCs [blood-derived mesenchymal precursor cells] were elutriated by diluting the buffy coats” in more than 100 normal persons. Zvaifler et al., “Mesenchymal precursor cells in the blood of normal individuals,” *Arthritis Res* 2:477-488 (2000) at page 478 (EX2195). Zvaifler reports that “[m]ore than 100 consecutive buffy coats from normal individuals were processed.... In every case, the appropriate elutriation fractions had cells with the BMPC morphology.” *Id.* at 479. Zvaifler explains that “it is likely that 500 ml of normal blood will have several thousand BMPCs.” *Id.* at 482.

69. Fourth, precursor cells to thrombocytes also can be found in the buffy coat. Efrati teaches that **megakaryocytes** (the precursor to thrombocytes) can be identified in the buffy coat layer from normal adults. Efrati “examined buffy coats from peripheral blood of 55 normal adults.” Efrati et al., “The morphology of buffy coat in normal human adults,” *Blood* 16:1012-1019 (1960) at page 1012 (EX2198). Efrati reported a mean value of 21.79 ± 17.27 megakaryocytes per ml of blood. *Id.* at

1017, Table 2. Efrati is particularly relevant to the current discussion, since megakaryocytes undergo endomitosis, a process in which they replicate their DNA without undergoing cell division. This results in an increased DNA content per cell, which can be up to 32 times the normal amount. Moreover, their unique multinucleate structure makes their presence in the buffy coat readily apparent with routine histochemical staining.

70. In short, the buffy coat layer contains a heterogeneous mixture of cell types, whose sole common attribute is that their densities are simply between the density of plasma and red blood cells.

X. THE CLAIMED INVENTION IS NONOBVIOUS OVER CHAN

71. Dr. Quackenbush does not provide any discussion that illuminates why he believes it would have been obvious for a person of ordinary skill in the art to modify Chan's MPS to achieve whole genome sequencing. See EX1100 (Quackenbush Declaration) ¶152. A person of ordinary skill in the art would not have been so motivated based on at least (A) the large data burden imposed by whole genome sequencing, and (B) the constraining expenses levied by whole genome sequencing.

A. The data burden of whole genome sequencing

72. The exhibits filed by Foresight demonstrate that whole genome sequencing was undesirable because it imposes a large data burden that is an order

of magnitude larger than targeted approaches. For example, Meyerson (EX1031) teaches that successful whole genome sequencing of normal tissue, typically requires the approximately 3 billion bases of the human genome to be sequenced at least 30 times on average, requiring the machine capacity to sequence 90 billion bases as follows:

Two important issues to consider when planning whole-genome sequencing experiments are depth of coverage and physical coverage. Sequence depth is measured by the amount of over-sampling: typically, to detect nucleotide alterations with high sensitivity, ***the 3 billion bases of the human genome are covered at least 30-fold on average, requiring the generation of 90 billion bases of sequence data per sample.*** For cancer samples, this number needs to be increased to account for the decreased purity and often increased ploidy of each sample.

EX1031 (Meyerson) at page 688. The data burden is even higher for Chan's cancer samples "to account for the decreased purity and often increased ploidy of each sample." *Id.* For example, if tumor cells were only 5% of the cells in a sample, then in planning to achieve a depth of 30-fold one would require the capacity to sequence 1,800 billion bases.

73. Meyerson further explains that planning to sequence the whole genome at 30-fold predicted depth requires the capacity to sequence 90 billion bases (90Gb) whereas 75-fold predicted depth of the exome can be achieved with only the capacity

to sequence 3 billion bases (3Gb), which is an order (and nearly two orders) of magnitude higher than the 3 Gb of sequence required for targeted approaches:

Because the exome represents only approximately 1% of the genome, or about 30 Mb, vastly higher sequence coverage can be readily achieved using second-generation sequencing platforms with considerably less raw sequence and cost than whole-genome sequencing. For example, whereas 90 Gb of sequence is required to obtain 30-fold average coverage of the genome, 75-fold average coverage is achieved for the exome with only 3 Gb of sequence using the current state-of-the-art platforms for targeting.

EX1031 (Meyerson) at page 690. Meyerson states that there are “computational issues” caused by “the huge quantity of experimental data” as follows:

Computational issues

Tools for computational analysis need to develop rapidly to keep up with ***the huge quantity of experimental data*** produced by second-generation sequencing. Many of the computational tools for second-generation sequencing were developed for analysing non-cancer samples for human population genetic studies.

EX1031 (Meyerson) at page 692. Meyerson states that it was a “major challenge” “to make biological sense of the mountains of genomic data” as follows:

It is likely that second-generation sequencing methods will continue to transform cancer genomics, leading to the comprehensive discovery of all the major alterations in the cancer genome, followed by the application of comprehensive sequencing approaches to cancer

diagnostics. Computational analysis will become a central part of these discovery and diagnostic efforts. ***The major challenge will be to make biological sense of the mountains of genomic data.***

EX1031 (Meyerson) at page 694.

74. Mertes (EX1006) teaches that targeted exome sequencing, as opposed to whole genome sequencing, provides “simpler data interpretation” that is a “crucial consideration” as follows:

The alternative of fully sequencing many complete genomes to high average coverage (~30x or higher) to enable things like genetic variation analysis, was ***simply not affordable***. Another reason for assaying, e.g. exome rather than whole-genome sequencing is the ***simpler data interpretation*** of the former. ***This is a crucial consideration as it is generally much more challenging to find the functional impact of variants in noncoding regions.***

EX1006 (Mertes) at page 382. In particular, the interpretation of data from noncoding regions was only in its infancy and, in fact, methods for interpreting data from much of the noncoding regions were not routine in the field as of 2014. Even today, interpretation of data from noncoding regions can be challenging.

75. This difficulty is evidenced by Ley’s decision to abandon the study of putative variants occurring in non-coding regions. In fact, Ley conceded that a targeted transcriptome analysis may have been more efficacious in the first place! See EX1012 (Ley) at page 71 (“Although we performed whole-genome sequencing

on this cancer sample, we restricted our initial validation studies to the 1-2% of the genome that encodes genes. This raises the issue of whether sequencing the complementary DNA transcriptome of this tumour would have been a faster, cheaper and more efficient way of finding the mutations.”).

76. Mamanova (EX1007) teaches that a person of ordinary skill in the art would pursue targeted methods because the demands “on laboratory time and funding” required by whole genome sequencing “would place a substantial burden on a research center’s informatics infrastructure” as follows:

Nevertheless, it is not yet feasible to sequence large numbers of complex genomes in their entirety because ***the cost and time taken are still too great***. To obtain 30-fold coverage of a human genome (90 Gb in total), would currently require several sequencing runs and would cost tens of thousands of dollars. ***In addition to the demands such a project would place on laboratory time and funding, the primary analysis during which the captured image files are processed, as well as storage of the sequences, would place a substantial burden on a research center’s informatics infrastructure.***

EX1007 (Mamanova) at 111.

77. Dr. Quackenbush does not explain why a person of ordinary skill in the art would have been motivated to modify Chan’s MPS approach to use whole genome sequencing. A person of ordinary skill in the art would not have been motivated to modify Chan’s MPS approach to use whole genome sequencing due to

the substantial burden of processing the very large amounts of data associated with whole genome sequencing. The above exhibits emphasize that a person of ordinary skill in the art would have avoided making such a modification to Chan due to the substantial burden of processing the very large amounts of data associated with whole genome sequencing.

78. In addition, a person of ordinary skill in the art would not have been motivated to modify Chan to use WGS in view of Ley. For example, a person of ordinary skill in the art would not be impressed by the meager collection of just 8 validated somatic SNVs reported by Ley's workflow that started with parallel whole genome sequencing of the tumor and normal genome. *See, e.g.*, EX1012 (Ley) at Figure 2 (reproduced below):

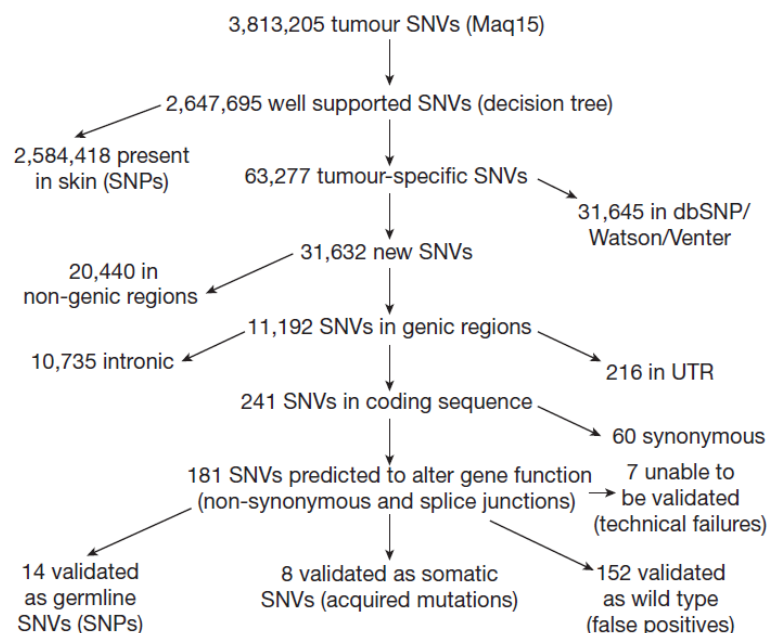


Figure 2 | Filters used to identify somatic point mutations in the tumour genome. See text for details. UTR, untranslated regions.

Ley's workflow involved "perform[ing] 98 full runs on the Illumina Genome Analyser to achieve a targeted level of 90% diploid coverage as determined by coverage of the high-quality SNP set" for the cancer genome and "34 Solexa runs (41.8 billion bases)... to attain 82.6% diploid and 84.2% haploid coverage" for the normal skin cells. EX1012 (Ley) at page 67. This was a heroic tour-de-force sequencing effort at the time. The extraordinary amount of work and cost that went into Ley's WGS to identify only 8 validated somatic SNVs would not motivate a person of ordinary skill in the art to take on the data burden (and cost burdens discussed below) of WGS in Chan's approach. Ley itself explains that a targeted transcriptome analysis may have been more efficacious than WGS. See EX1012 (Ley) at page 71 ("Although we performed whole-genome sequencing on this cancer sample, we restricted our initial validation studies to the 1-2% of the genome that encodes genes. This raises the issue of whether sequencing the complementary DNA transcriptome of this tumour would have been a faster, cheaper and more efficient way of finding the mutations.").

B. Whole genome sequencing was more expensive than other approaches

79. The exhibits filed by Foresight explain that whole genome sequencing was undesirable because it is significantly more expensive than other prior art approaches. The difference in expense is not trivial. Chan itself states "the shotgun sequencing approach we have described is relatively expensive." EX1008 (Chan) at

223. Meyerson (EX1031) explains that “whole-genome sequencing provides the most comprehensive characterization of the cancer genome but, as it requires the greatest amount of sequencing, it is the most costly.” EX1031 (Meyerson) at 687.

80. Mertes teaches that targeted approaches allow “costs and efforts [to] be reduced significantly compared with whole-genome sequencing” as follows:

As part of this, methods have been developed that aim to achieve ‘targeted enrichment’ of genome subregions, also sometimes referred to as ‘genome partitioning’. Strategies for direct selection of genomic regions were already developed in anticipation of the introduction of NGS. ***By selective recover[y] and subsequent sequencing of genomic loci of interest, costs and efforts can be reduced significantly compared with whole-genome sequencing***

EX1006 (Mertes) at 374.

81. Gnirke (EX1005) teaches that there is a “substantial economy in targeting” that “substantially lower[s] the sequencing costs of a project” as follows:

Sequencing entire human genomes will be an important application of next-generation sequencing. However, many research and diagnostic goals may be achieved by sequencing a specific subset of the genome in large numbers of individual samples. ***For example, there may be substantial economy in targeting the protein-coding fraction, the 'exome', which represents only ~ 1 % of the human genome.*** The economy is even greater for many key resequencing targets, such as genomic regions implicated by whole genome association scans and the exons of sets of protein-coding genes implicated in specific diseases.

Efficient and cost-effective targeting of a specific fraction of the genome could substantially lower the sequencing costs of a project, independent of the sequencing technology used.

EX1005 (Gnirke) at 182.

82. The exhibits filed by Foresight in this proceeding describe the real-world cost constraints imposed by whole genome sequencing as compared to other approaches. A person of ordinary skill in the art faced with the real-world cost constraints imposed by whole genome sequencing would not have been motivated to modify Chan's MPS to be whole genome sequencing. For example, as discussed above, Ley's WGS identified only 8 validated somatic SNVs. A person of ordinary skill in the art would not have been motivated to take on the large data and cost burdens of WGS in so modifying Chan's approach in view of the small amount of somatic SNVs identified by Ley.

XI. THE CLAIMED INVENTION IS NONOBVIOUS OVER LEY

83. Dr. Quackenbush asserts that Ley (EX1012) renders obvious claim 1. See EX1100 (Quackenbush Declaration) ¶¶167-198. Ley does not render obvious claim 1 for at least the reasons discussed below.

A. Limitation 1(a)(i)-(ii)

84. Limitation 1(a)(i)-(ii) recites "wherein (i) said first set of nucleic acid molecules is extracted from tissue sample of said subject, and (ii) said second set of nucleic acid molecules is extracted from leukocytes from a blood sample."

85. Dr. Quackenbush concedes that Ley fails to teach limitation 1(a)(i)-(ii) because Ley “differs from the claim insofar as the tumor sample is blood (rather than tissue) in a leukemia.” EX1100 (Quackenbush Declaration) ¶169. A person of ordinary skill in the art would not have been motivated to modify Ley’s approach to instead use a “tissue sample” as the tumor sample and leukocytes as the normal sample for several reasons.

1. Ley counsels against Foresight’s proposed modification

86. A typical tumor tissue sample contains a mixture of tumor and normal cells. For example, Meyerson (EX1031) teaches that a tumor tissue sample contains a mixture of malignant and non-malignant cells as follows:

[A] cancer specimen contains a mixture of malignant and non-malignant cells and, therefore, a mixture of cancer and normal genomes (and transcriptomes). Furthermore, the cancers themselves may be highly heterogeneous and composed of different clones that have different genomes. Cancer genome analytical models must take these two types of heterogeneity (cancer versus normal heterogeneity and within-cancer heterogeneity) into account in their prediction of genome alterations

EX1031 (Meyerson) at 686. Meyerson explains that such samples create “particular challenges for the detection and diagnosis of cancer genome alterations.” EX1031 (Meyerson) at 685. The normal (non-malignant) cells in a tumor tissue sample are very problematic because they cause a high background level of normal cell signal

that makes it very difficult to establish the identity of a presence or absence of genetic variants in the sample. Meyerson explains that this problem was not yet solved and explains that this problem might be solved in the future if there were large “data sets” of “sufficient statistical power to overcome experimental and biological noise.” *Id.* at 686.

87. Ley’s method gets around this problem by selecting a patient having “a typical acute myeloid leukemia [AML],” such that the patient’s blood sample provide a highly pure source of tumor DNA. EX1042 (Ley Supplementary Information) at 1 (providing the “[r]ationale for selection of an AML genome for sequencing”). Ley teaches that its method “**required** high sample purity” in the tumor sample to “eliminate the need to further purify malignant cells.” EX1042 at 1 (“2. High sample purity. The morphologic and flow cytometric criteria for M1 AML required high sample purity (e.g. a blast count of >80%) to eliminate the need to further purify malignant cells.”).

88. Ley teaches that contaminating normal cells would have undesirably complicated its analysis as follows:

From these 6 [cases of AML], we selected one case for study, UPN 933124, primarily because the patient’s initial bone marrow sample contained virtually 100% myeloblasts, eliminating concerns that contaminating normal cells would complicate our analyses. Although this case was carefully selected, the presentation of AML was typical,

and the case is representative of most patients with M1 AML, as described below.

EX1042 (Ley Supplementary Information at 1). Thus, Ley's method selected this leukemia patient precisely because the patient's blood sample provided a highly pure source of tumor DNA.

89. Ley's method stands in stark contrast to a typical tumor sample containing a mixture of tumor and normal cells that involves challenges for detection and diagnosis as conceded in Foresight's Meyerson exhibit. EX1031 (Meyerson) at 686. Dr. Quackenbush's obviousness theory requires a person of ordinary skill in the art to go against Ley's express teachings to instead use a tumor tissue sample contaminated with normal cells, which would defeat the purpose of Ley's rationale for selecting an AML patient to provide the highly pure source of tumor DNA required for Ley's method.

2. Dr. Quackenbush's references do not provide the motivation to modify Ley

90. Dr. Quackenbush cites four references that supposedly support his opinion that Ley obviates claim 1 of the '968 patent: Meyerson, Chapman, Lee, and Ding. EX1100 (Quackenbush Declaration) ¶¶171-176. First, none of these references teaches to use leukocytes as the normal sample. Second, each of these references show that a person of ordinary skill in the art would not have been motivated to use leukocytes as the normal sample.

a. Meyerson

91. Dr. Quackenbush states that Meyerson reported that “[c]omplete sequencing of the genome of cancer tissue to high redundancy, using germline DNA sequence from the same patient as a comparison, has the power to discover the full range of genomic alterations,” and also “the first of seven references Meyerson cites for this proposition is Ley.” EX1100 (Quackenbush Declaration) ¶171. The mere fact that Meyerson cites Ley as the first of seven references for this proposition does not establish that a person of ordinary skill in the art would have been motivated to modify Ley’s approach to instead use a tissue sample as the tumor sample and leukocytes as the normal sample to arrive at limitation 1(a)(i)-(ii). The quote from Meyerson has nothing to do with using leukocytes as the normal cell sample. It is a generic statement that merely acknowledges the long-held belief that a person of ordinary skill in the art would compare putative genomic alterations with the DNA sequence from normal cells.

b. Chapman

92. Dr. Quackenbush states that Chapman reported “the massively parallel sequencing of 38 tumor genomes and their comparison to matched normal DNAs” and he asserts that Chapman cites Ley as an “exemplary” “recent report.” EX1100 (Quackenbush Declaration) ¶172. In the introduction, Chapman states:

Recent reports have described the sequencing of whole genomes from a single patient⁵⁻⁹. Although informative, we hypothesized that a larger number of cases would permit the identification of biologically relevant patterns that would not otherwise be evident.

EX1103 (Chapman) at 467. Citation number 7 in Chapman is the Ley reference. I do not believe that this statement from Chapman about “recent reports” is consistent with the reports being “exemplary.” Indeed, a person of ordinary skill in the art may well believe that the Ley paper was underwhelming and not motivating. Chapman’s citation to Ley does not establish that a person of ordinary skill in the art would have been motivated to modify Ley’s approach to use a tissue sample as the tumor sample and leukocytes as the normal sample to arrive at the claimed invention.

93. First, Chapman does not teach extracting nucleic acids *from leukocytes* as required by limitation 1(a)(ii).

94. Second, a person of ordinary skill in the art would not have been motivated to modify Ley to use a tumor tissue sample instead of a blood sample as the tumor sample. Chapman teaches using *bone marrow aspirate* tumor samples from patients suffering from multiple myeloma, which an uncommon type of cancer that affects the bone marrow. EX1103 (Chapman) at 471. A person of ordinary skill in the art would not regard multiple myeloma as a representative model for general oncological investigations because it is an uncommon bone marrow cancer in which one would employ the uncommon strategy of sampling bone marrow aspirates. Not

only is multiple myeloma an uncommon type of cancer, but it is not a representative model additionally because bone marrow aspirate samples are a complex mixture of cells, including precursor stem cells, in various states of development. Chapman's bone marrow aspirate tumor sample contains a mixture of tumor and normal cells. Chapman teaches that the use of bone marrow aspirate tumor samples from patients suffering from multiple myeloma raised complexity issues that required sequencing **38 separate patients**. *Id.* ("Importantly, most of these discoveries could not have been made by sequencing only a single multiple myeloma genome—the complex patterns of pathway dysregulation required the analysis of multiple genomes."); *id.* (Chapman) at 467 (discussing sequencing of "38 tumour genomes and their comparison to matched normal DNAs"). Chapman's approach that required 38 separate patients is an entirely different scenario from Ley's carefully selected leukemia patient that provided a highly pure blood tumor sample and sequencing a single subject. EX1042 (Ley Supplementary Information) at 1.

95. Chapman's sequencing of 38 separate patients only allowed a "preliminary" analysis, which caused the authors to caution the person of ordinary skill in the art that "[a]dditional multiple myeloma genomes will be required to establish the definitive genomic landscape of the disease and determine accurate estimates of mutation frequency in the disease." EX1103 (Chapman) at 471. Dr. Quackenbush does not explain why a person of ordinary skill in the art would have

altered Ley's approach to use a tissue sample as the tumor sample and leukocytes as the normal sample in view of Chapman's failure to teach leukocytes and the complexity issues raised by Chapman. Chapman demonstrates that using a tissue sample as the tumor sample greatly increases complexity issues and requires performing the assay on at least 38 patients to provide just a preliminary analysis.

96. Third, a person of ordinary skill in the art would have been concerned about Chapman's use of blood as the normal control for the patients suffering from multiple myeloma (which is a B cell tumor). A person of ordinary skill in the art would understand that although the majority of the tumor cells would reside in the bone marrow, there is a significant possibility that tumor cells would be present in the patient's blood. This scenario would confound the ability to identify tumor mutations with appropriate sensitivity. In addition, further purification of cell types from peripheral blood would likely make this problem worse.

c. Lee

97. Dr. Quackenbush asserts that Lee "performed WGS on paired tumor and adjacent normal tissue of a lung cancer patient." EX1100 (Quackenbush Declaration) ¶173. The adjacent normal tissue in Lee are primarily composed of lung cells, not leukocytes. EX1104 (Lee) at 473. Lee fails to teach obtaining a normal sample by extracting nucleic acids from leukocytes, as required by limitation 1(a)(ii).

d. Ding

98. Dr. Quackenbush asserts that Ding used “peripheral blood” as the normal sample for whole genome sequencing. EX1100 (Quackenbush Declaration) ¶¶174-175. First, Ding does not use the term “whole genome sequencing.” Ding discloses MPS. EX1105 (Ding) at 999, Abstract (“Massively parallel DNA sequencing technologies provide an unprecedented ability to screen entire genomes for genetic changes associated with tumour progression.”).

99. Second, Ding’s use of “peripheral blood” as the normal sample does necessarily or inevitably teach the use of leukocytes as required by limitation 1(a)(ii). It is common knowledge that peripheral blood contains a complex mixture of cells, cell fragments, hormones, enzymes, and other proteins, electrolytes, amino acids, glucose and other carbohydrates, cholesterol, vitamins, fatty acids, glycerol, triglycerides, respiratory gases, water, and wastes (such as, lactic acid, creatinine, urea, bilirubin, ammonia). Peripheral blood contains many types of nucleated non-leukocyte cell types from which DNA can be extracted, including at least mesenchymal precursor cells, endothelial cells, endothelial progenitor cells, and megakaryocytes. *Supra* Section IX. In addition, Dr. Quackenbush states that peripheral blood contains circulating extracellular DNA (“cell-free DNA” or “cfDNA”) from non-leukocyte sources of healthy and diseased cells. EX1100

(Quackenbush Declaration) ¶89 (“there is always cfDNA being released into and circulating through the bloodstream”).

100. Further, a person of ordinary skill in the art would not have been motivated to modify Ley’s approach to use a “tissue sample” as the tumor sample and leukocytes as the normal sample based on Ding. Ding teaches using **breast cancer** and **brain metastasis** tumor samples. EX1105 (Ding) at Abstract (“Here we describe the genomic analyses of four DNA samples from an African-American patient with basal-like breast cancer: peripheral blood, the primary tumour, a brain metastasis and a xenograft derived from the primary tumour.”). Ding teaches that there was considerable heterogeneity in the primary tumor. *See id.* at page 1004 (“The wide range of mutant allele frequencies suggests considerable genetic heterogeneity in the cellular population at the primary site. The mutation frequency range narrowed in brain metastasis and xenograft, indicating that the metastatic and transplantation processes selected for cells carrying a distinct subset of the primary tumour mutation repertoire. The overlap between the mutation frequency changes seen in the metastatic and xenograft samples argues that cellular selection during xenograft formation is similar to that during metastasis.”). Ding’s analysis of these samples involved heterogeneity complexities that required the design of an unpublished “cnvHMM algorithm” to “align[] sequence reads” and “calculate[] the read depth contributed from the tumor cells alone.” EX1105 (Ding) at 1001 (“The

cnvHMM algorithm (K.C., X.S., E.R.M., L.D. and R.K.W., unpublished) was applied to the aligned sequence reads to detect regions of copy number alterations in all three tumours.”). Dr. Quackenbush does not explain whether a person of ordinary skill in the art would have been able to develop or otherwise access such an algorithm to use such tumor samples to arrive at the claimed invention with a reasonable expectation of success. In contrast to Ding’s approach, Ley selected a leukemia patient that provided a highly pure blood tumor sample to “eliminate the need to further purify malignant cells” because “contaminating normal cells would complicate our analysis.” EX1042 (Ley Supplementary Information) at 1. Thus, Ding emphasizes that using a tissue sample as the tumor sample would increase the complexity of analyzing the samples, which a person of ordinary skill in the art would find undesirable.

B. Limitation 1(c)

101. Limitation 1(c) recites “subsequent to (b), identifying *said mosaic variants specific to said tissue sample* of said subject in a first additional sample of said subject, wherein said first additional sample is obtained from a source that is different than said tissue sample and at a time point that is different than said tissue sample.

102. The “said” mosaic variants specific to said tissue sample of limitation 1(c) refers back to the mosaic variants identified using the first and second set of

sequence reads of limitation 1(b). Limitation 1(b) requires “using said first set of sequence reads and said second set of sequence reads to identify mosaic variants specific to said tissue sample of said subject in said first set of sequence reads ***that is not present in said second set of sequence reads.***”

103. Reading these two limitations together, the “said” mosaic variants specific to said tissue sample in the first additional sample of limitation 1(c) must be “specific to said tissue sample of said subject in said first set of sequence reads ***that is not present in said second set of sequence reads***” as required by limitation 1(b).

104. For limitation 1(c), Dr. Quackenbush relies on the Ley’s “relapse tumor” sample data in Ley Figure 3 and Table 3. EX1100 (Quackenbush Declaration) ¶¶195-196. He asserts that “Figure 3 summarizes [Ley’s] data, and shows the frequency of mosaic variants (‘ten somatic mutations’) in each of the primary tumor, relapse tumor, and skin samples.” *Id.* ¶196. He further asserts “[t]his shows that Ley identified the presence of mosaic variants (green bars) specific to the tumor sample in the subsequently obtained relapse tumor sample.” *Id.* ¶196.

105. Dr. Quackenbush fails to account for Ley’s teaching that the normal skin samples used to generate Figure 3 and Table 3 were contaminated with tumor cells. Specifically, Ley teaches that “the normal skin sample was contaminated with leukaemic cells containing the somatic mutations.” EX1012 (Ley) at 69. Thus, the mosaic variants in Ley’s relapse tumor sample were also present in the normal skin

sample. The yellow bars in Ley’s Figure 3 shows the presence of each of the mosaic variants in the normal samples:

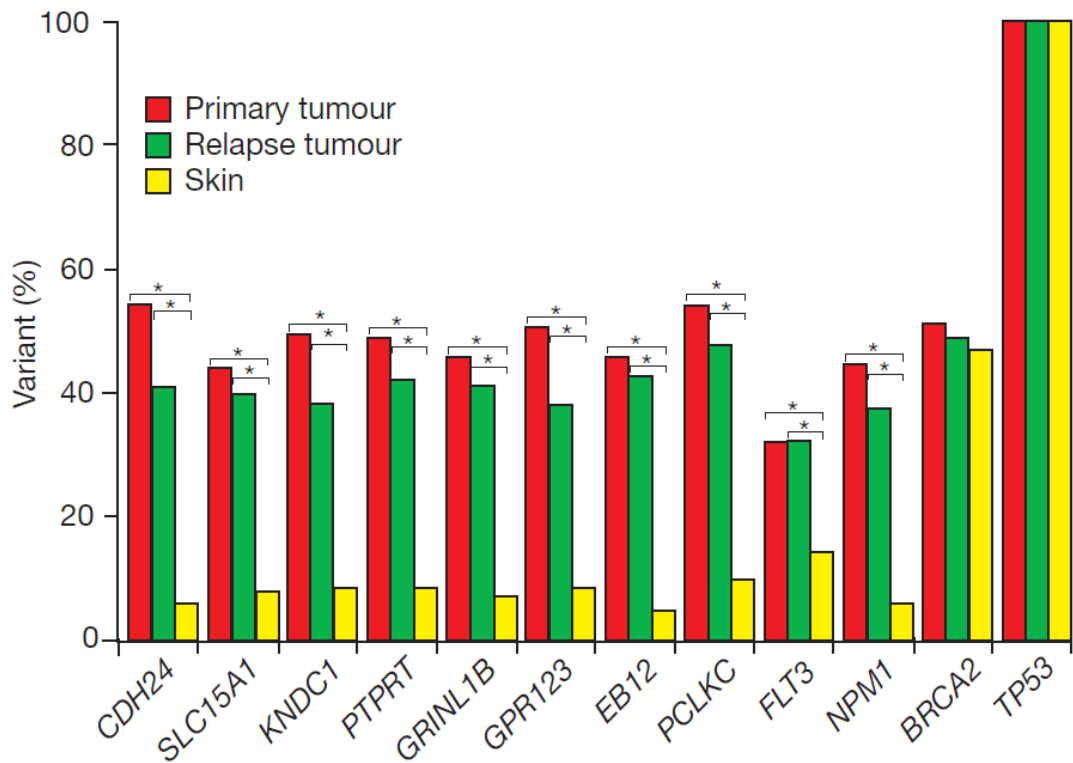


Figure 3 | Summary of Roche/454 FLX readcount data obtained for ten somatic mutations and two validated SNPs in the primary tumour, relapse tumour and skin specimens. The readcount data for the variant alleles in the primary tumour sample and relapse tumour sample are statistically different from that 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.

Because the data in Ley’s Figure 3 and Table 3 shows that the mosaic variants *were* present in the normal skin sample, Ley fails to teach that the sequence reads for mosaic variants in the relapse tumor sample were not present in the sequence reads

for the normal skin sample. Dr. Quackenbush therefore fails to prove that Ley teaches or suggests limitation 1(c).

**XII. LACK OF ANTICIPATION AND OBVIOUSNESS BASED ON
OKOSUN-II**

106. Dr. Quackenbush asserts that Okosun-II anticipates and renders obvious claim 1 of the '968 patent. In this section of the declaration I address why Okosun-II does not teach or suggest limitation 1(a)(ii) and limitation 1(c).

A. Limitation 1(a)(ii)

107. Limitation 1(a)(ii) recites that “said second set of nucleic acid molecules is extracted from leukocytes from a blood sample.”

108. Dr. Quackenbush asserts that Okosun-II’s disclosure of “peripheral blood specimens” meets limitation 1(a)(ii) because “peripheral blood specimens contain leukocytes and thus satisfy the claim.” EX1100 (Quackenbush Declaration) ¶245. Okosun-II’s disclosure of “peripheral blood specimens” does not necessarily or inevitably teach the use of leukocytes as required by limitation 1(a)(ii). It is common knowledge that peripheral blood contains a complex mixture of cells, cell fragments, hormones, enzymes, and other proteins, electrolytes, amino acids, glucose and other carbohydrates, cholesterol, vitamins, fatty acids, glycerol, triglycerides, respiratory gases, water, and wastes (such as, lactic acid, creatinine, urea, bilirubin, ammonia). Peripheral blood contains many types of nucleated non-leukocyte cell types from which DNA can be extracted, including at least

mesenchymal precursor cells, endothelial cells, endothelial progenitor cells, and megakaryocytes. *Supra* Section IX. In addition, Dr. Quackenbush states that peripheral blood contains circulating extracellular DNA (“cell-free DNA” or “cfDNA”) from non-leukocyte sources of healthy and diseased cells. EX1100 (Quackenbush Declaration) ¶89 (“there is always cfDNA being released into and circulating through the bloodstream”).

109. Dr. Quackenbush alleges that it would have been obvious to use leukocytes in place of Okosun-II’s peripheral blood specimens based on Meyerson’s teaching that “samples generally used for germline genome analysis [are] peripheral blood mononuclear cells [PBMCs].” EX1100 (Quackenbush Declaration) ¶272. Meyerson states the following regarding PBMCs:

Finally, cancer nucleic acid specimens are less pure than specimens used to analyse the inherited genome, especially in terms of genomic DNA purity. The samples generally used for germline genome analysis—peripheral blood mononuclear cells—are known to be heterogeneous only at the rearranged immunoglobulin and T cell receptor loci in a subset of cells. By contrast, a cancer specimen contains a mixture of malignant and nonmalignant cells and, therefore, a mixture of cancer and normal genomes (and transcriptomes).

EX1031 (Meyerson) at 686. In this paragraph, Meyerson is not teaching the suitable source of normal germline cells to compare and thereby identify somatic mutation in a tumor tissue. A person of ordinary skill in the art would be well aware that such

teaching is to select a tissue source that is ostensibly normal, available, and free from tumor cells. In the present context of Okosun-II, a person of ordinary skill in the art, would be very skeptical of using blood cells to serve as a normal control for a patient who suffers from follicular lymphoma (which is a B cell lymphoma). A person of ordinary skill in the art would understand that although the majority of the tumor cells would reside in a lymph node, there is a significant possibility that tumor cells would be present in the peripheral blood. A person of ordinary skill in the art would anticipate that an appropriate normal control would likely be a histologically negative lymph node.

110. In a different scenario where blood in fact is a suitable control, Meyerson, in the cited paragraph, is simply teaching its advantages. The principal “advantage” being that, in contrast to tumor cells, blood cells genomes are typically homogenous, the only exception being that the receptor genes of T and B cells are clonally rearranged. EX1031 (Meyerson) at page 686.

111. A person of ordinary skill in the art would not have been motivated to modify Okosun-II’s method of sequencing based on Meyerson for at least the following three reasons.

112. First, Meyerson does not exclusively teach using PBMCs for germline genome analysis. Meyerson also teaches that peripheral blood samples are appropriate. EX1031 (Meyerson) at 686 (“Cancer samples differ in their quantity,

quality, and purity from the peripheral blood samples that are used for germline genome analysis”); *id.* at 686 (“Nucleic acids from cancer are also often of lower quality than those purified from peripheral blood.”). Myerson’s peripheral blood samples are the same type of sample taught by Okosun-II to use in Okosun-II’s protocol.

113. Second, and more importantly, Meyerson’s statement about PBMCs “generally” being used for germline genome analysis provides no reason why a person of ordinary skill in the art would modify Okosun-II’s follicular lymphoma study to use PBMCs. Follicular lymphoma is a hematologic cancer. A significant percentage of follicular lymphoma patients evolve to (or in some cases present with) a leukemic phase in which the patients’ peripheral blood contains lymphocytes that are cancerous. *See* Spiro et al., “Follicular lymphoma: a survey of 75 cases with special reference to the syndrome resembling chronic lymphocytic leukaemia,” *British Journal of Cancer* 31, Suppl. II, 60-72 (1975) (EX2200) at page 60 (“In a series of 75 cases of follicular lymphoma collected between 1950 and 1971 the following features were singled out as characteristic of, or specially frequent in, follicular lymphoma: ... (4) the syndrome resembling and often confused with chronic lymphocytic leukaemia, in which the notched nucleus cell, when present, is diagnostic....”). Spiro reported “[a] characteristic form of lymphocyte, the ‘notched nucleus cell’ of Anday and Schmitz (1952), or the ‘haematogone’ of Rosenthal et al.

(1952) was found in the peripheral blood in 8 of 17 cases with lymphocytosis and accounted for the majority of lymphocytes in the cases in which the lymphocyte counts were considerably elevated, but the proportion was variable in those with only moderate lymphocytosis.” EX2200 (Spiro) at page 63. PBMCs include peripheral blood lymphocytes. A person of ordinary skill in the art would not have been motivated to use PBMCs as the normal sample for Okosun-II’s follicular lymphoma study because the PBMCs would be expected to be contaminated with cancerous cells for patients that enter or present with leukemic phase follicular lymphoma.

114. If a person of ordinary skill in the art elected to use blood as a control and were aware of the potential problem of circulating tumor B cells in a follicular lymphoma patient, they would not consider purifying PBMCs. Such a purification, would increase the percentage of tumor genomes in the sample (by removing the genomes of all the non-PBMC cell types we have discussed above) and thus diminish the prospect of detecting tumor variants in the follicular lymph node sample.

115. Third, the standard method for isolation of PBMCs is density gradient centrifugation. See Zvaifler et al., “Mesenchymal precursor cells in the blood of normal individuals,” *Arthritis Res* 2:477-488 (2000) at page 478 (EX2195). It is a relatively labor-intensive, tedious, and time-consuming process. It involves precise

layering of peripheral blood over a density medium, with careful pipetting of the floating cell layer after centrifugation. The user must take meticulous care throughout the procedure to keep from mixing layers before and after centrifugation, which invites user error. The carefulness required throughout this multistep procedure can lead to inconsistency between users and sample variability. In contrast, Okosun-II's peripheral blood samples are obtained by a simple blood draw from the patient. Even if a person of ordinary skill in the art could have modified Okosun-II to use PBMCs instead of peripheral blood, Dr. Quackenbush provides no supposed advantages that might arise from such a modification within the specific context of Okosun-II's protocol that would have motivated a person of ordinary skill in the art to further complicate Okosun-II's already-working protocol with unnecessary processing steps required to isolate PBMCs from the peripheral blood samples.

116. Fourth, Meyerson was published in 2010, which is five years before Okosun-II was published. Thus, Okosun-II was likely aware of the supposed general understanding described by Meyerson that PBMCs were “generally” used for germline genome analysis. If Okosun-II had thought that complicating its protocol with the additional processing steps of peripheral blood required to isolate PBMCs would have been a beneficial improvement, then Okosun-II would have done so.

B. Limitation 1(c)

117. Limitation 1(c) recites “subsequent to (b), identifying said mosaic variants specific to said tissue sample of said subject in a first additional sample of said subject, ***wherein said first additional sample is obtained from a source that is different than said tissue sample*** and at a time point that is different than said tissue sample.”

118. Dr. Quackenbush relies on Okosun-II’s follicular lymphoma (FL) and transformed follicular lymphoma (tFL) biopsies to allegedly meet the first additional sample required by limitation 1(c). EX1100 (Quackenbush Declaration) ¶¶250-252. Dr. Quackenbush also relies on Okosun-II’s FL and tFL biopsies to allegedly meet the tissue sample required by limitation 1(a). *Id.* ¶¶241-244. A person of ordinary skill in the art would be aware that “follicular lymphoma” (FL) refers to a relatively slow-growing (indolent) follicular lymphoma, whereas “transformed follicular lymphoma refers to a faster-growing (aggressive) follicular lymphoma. Both types of follicular lymphomas are from the same source (follicle tissue). Thus, Dr. Quackenbush relies on tissue samples from the ***same source*** (i.e., follicle tissue biopsies) to supposedly meet limitations 1(a) and 1(c). However, limitation 1(c) requires the first additional sample to be obtained from a ***source that is different*** than the tissue sample of limitation 1(a). Because Dr. Quackenbush relies on using a follicle biopsy tissue sample for limitation 1(a) and using a follicle biopsy tissue

sample for limitation 1(c), it necessarily fails to teach the requirement of limitation 1(c) that the first additional sample must be obtained from a source that is different than said tissue sample of limitation 1(a).

119. A person of ordinary skill in the art would not have been motivated to use a non-FL sample in Okosun-II's protocol to arrive at limitation 1(c) because the purpose of Okosun-II was to identify and monitor the "progression of follicular lymphoma" (FL). See EX1102 (Okosun-II) at title ("Integrated genomic analysis identified recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma."). Thus, modifications of Okosun-II to use non-FL samples would undermine the purpose of Okosun-II's approach to test FL-samples to identify and monitor the progression of follicular lymphoma.

XIII. CERTIFICATION

120. I 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: February 29, 2024

By: 
Henry Morrice Furneaux, Ph.D.