

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GUARDANT HEALTH, INC.
Petitioner,

v.

COLD SPRING HARBOR LABORATORY
Patent Owner

IPR2025-XXXXX
U.S. PATENT NO. 12,234,510

DECLARATION OF GREGORY COOPER, PH.D IN SUPPORT OF
PETITIONER FOR *INTER PARTES* REVIEW

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

TABLE OF CONTENTS

	Page
I. INTRODUCTION	1
II. QUALIFICATIONS AND COMPENSATION.....	1
III. SUMMARY OF OPINIONS.....	6
IV. LEGAL STANDARDS	7
A. Claim Construction	7
B. Obviousness.....	8
V. THE '510 PATENT.....	9
VI. PROSECUTION HISTORY OF THE '510 PATENT.....	14
VII. RECENT BOARD UNPATENTABILITY DECISION IN IPR PROCEEDINGS BETWEEN PARSE BIOSCIENCES AND 10X GENOMICS	16
VIII. CLAIM CONSTRUCTION	18
IX. PERSON OF ORDINARY SKILL IN THE ART	19
X. TECHNICAL BACKGROUND AND STATE OF THE ART	20
A. DNA	20
B. Amplification.....	21
C. DNA Tagging	24
D. Amplification Bias and Molecular Tagging.....	28
E. Sequencing	31
F. Copy Number	33
XI. THE PRIOR ART REFERENCES OF THE GROUNDS	34
A. Lo (Ex. 1010)	34
1. Lo Is § 102 Prior Art.....	34
2. Lo Overview	34
B. McCloskey (Ex. 1008)	42
1. McCloskey Is § 102 Prior Art.....	43
2. McCloskey Overview	43

TABLE OF CONTENTS
(continued)

	Page
C. Brenner (Ex. 1014).....	46
1. Brenner Is § 102 Prior Art	46
2. Brenner Overview	46
D. Chee (Ex. 1015).....	50
1. Chee Is § 102(e) Prior Art.....	50
2. Chee Overview.....	51
E. Porreca (Ex. 1017).....	54
1. Porreca Is § 102(e) Prior Art.....	54
2. Porreca Overview.....	55
XII. PRIOR ART RENDERS OBVIOUS THE CLAIMS OF THE '510 PATENT.....	58
A. Ground 1: Claims 1, 2, 9-11, 22, 26, 28, and 33-34 Are Unpatentable as Obvious Over Lo in View of The Knowledge of a POSA.....	60
1. Claim 1:.....	60
a) [1A] A method comprising: a) obtaining segments of genomic nucleic acids from a sample containing the genomic nucleic acids;.....	60
b) [1B] randomly tagging the segments of genomic nucleic acids with different nucleic acid tags, thereby generating unique tagged nucleic acid molecules from the segments of genomic nucleic acids, such that each of the unique tagged nucleic acid molecules comprises a segment of the segments of the genomic nucleic acids from step (a) and a nucleic acid tag of the nucleic acid tags;	61
c) [1C] subjecting the unique tagged nucleic acid molecules to a polymerase chain reaction (PCR), thereby generating copies of the unique tagged nucleic acid molecules;.....	77

TABLE OF CONTENTS

(continued)

	Page
d) [1D] generating tag associated sequence reads by sequencing the copies of the unique tagged nucleic acid molecules of step (c);	84
e) [1E] assigning each of the unique tagged nucleic acid molecules to a location on a genome by mapping each of the tag associated sequence reads of step (d) to a location in the genome; and	87
f) [1F] in a plurality of locations in the genome, obtaining relative copy number information of the genomic nucleic acids from the sample based on the number of unique tagged nucleic acid molecules that have been assigned to each of the plurality of locations in the genome.	90
2. Claim 2: The method of claim 1, wherein the polymerase chain reaction is performed in the presence of region specific PCR primers.	95
3. Claim 9: The method of claim 1, wherein the unique tagged nucleic acid molecules are subjected to hybrid capture prior to the PCR or the amplified tagged nucleic acid molecules produced in step (c) are subjected to hybrid capture prior to step (d).	98
4. Claim 10: The method of claim 1, wherein the genomic nucleic acids are from a single species, two or more organisms, two or more species, or from a population of microbes.	99
5. Claim 11: The method of claim 1, wherein the sample is a blood sample or bone marrow sample.	100
6. Claim 22: The method of claim 1, wherein the nucleic acid tags comprise a sufficiently large number of distinct nucleic acid tags.	101
7. Claim 26: The method of claim 22, wherein the nucleic acid tags comprise nucleotide sequences with constant portions and variable portions.....	104

TABLE OF CONTENTS

(continued)

	Page
8. Claim 28: The method of claim 26, wherein the length of the constant portions and the length of the variable portions are different.	107
9. Claims 33 and 34: The method of claim 22, wherein the nucleic acid tags comprise [nucleotide sequences that are not known before the tagged nucleic acid molecules are generated/random nucleotide sequences].	109
B. Ground 2: The Combination of Lo and the Knowledge of a Skilled Artisan in Further View of Specific Disclosure in McCloskey, Brenner, or Chee Renders Obvious Claim 3.	110
1. Claim 3: The method of claim 1, wherein the unique tagged nucleic acid molecules differ at more than one nucleotide or wherein the sequences of the nucleic acid tags are known prior to said generating unique tagged nucleic acid molecules from the segments of genomic nucleic acids.	110
C. Ground 3: The Combination of Lo and the Knowledge of a Skilled Artisan in Further View of Specific Disclosure in Brenner or Porreca Renders Obvious Claim 4.	114
1. Claim 4: The method of claim 1, wherein the segments of the genomic nucleic acids are produced by mechanical shearing, heating, or sonicating the genomic nucleic acids from the sample.	114
D. Ground 4: The Combination of Lo and the Knowledge of a Skilled Artisan in Further View of Specific Disclosure in McCloskey, Brenner, or Porreca Renders Obvious Claims 5-8.	116
1. Claim 5: The method of claim 1, wherein the nucleic acid tags comprise a sample tag or a sample tag set.	116
2. Claim 6: The method of claim 5, wherein the unique tagged nucleic acid molecules are pooled with a plurality of unique tagged nucleic acid molecules having a different sample tag prior to the PCR or prior to said sequencing of the amplified tagged nucleic acid molecules product of step (c).	119

TABLE OF CONTENTS

(continued)

	Page
3. Claim 7: The method of claim 6, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads.....	120
4. Claim 8: The method of claim 1, wherein the nucleic acid tags are double-stranded.	121
E. Ground 5: The Combination of Lo and the Knowledge of a Skilled Artisan in Further View of Specific Disclosure in Brenner, Chee, or Porreca Renders Obvious Claims 12, 15, 17, 23, and 35-36.....	122
1. Claim 12: The method of claim 1, wherein the unique tagged nucleic acid molecules are unique based on a combination of sequence information in the nucleic acid tags and sequence information in the segments of the genomic nucleic acids.	122
2. Claim 15: The method of claim 12, wherein the nucleic acid tags comprise nucleotide sequences with constant portions and variable portions.....	126
3. Claim 17: The method of claim 15, wherein the length of the constant portions and the length of the variable portions are different.	127
4. Claim 23: The method of claim 22, wherein the unique tagged nucleic acid molecules are unique based on a combination of sequence information in the nucleic acid tags and sequence information in the segments of the genomic nucleic acids.	129
5. Claims 35-36: The method of [claim 22/claim 1] wherein the nucleic acid tags are substantially unique nucleic acid tags.	131
F. Ground 6: Claims 18-21 And 29-32 Are Obvious In Further View Of The Knowledge Of The Skilled Artisan.....	132

I. INTRODUCTION

1. My name is Greg Cooper. I have been retained by Guardant Health, Inc. (“Guardant” or “Petitioner”) in the above-captioned *inter partes* review (“IPR”) to provide my independent evaluation and opinions regarding U.S. Patent No. 12,234,510 (the “’510 patent”) and the related prior art. EX1001.

2. This Declaration sets forth my analyses and opinions based on my technical background, skills, knowledge, training, education and experience in the relevant art, and the materials I considered. As I explain below, it is my opinion that claims 1-12, 15, 17-23, 26, and 28-34 of the ’510 patent are rendered obvious by the combination of prior art references discussed herein. I note that I am simultaneously presenting opinions regarding U.S. Patent No. 10,947,589, which is in the same family as the ’510 patent. I reserve the right to rely on my opinions presented in connection with U.S. Patent No. 10,947,589 herein.

II. QUALIFICATIONS AND COMPENSATION

3. I am currently a Faculty Investigator at the HudsonAlpha Institute for Biotechnology in Huntsville, Alabama. I have held this position since 2010. HudsonAlpha, established in 2008, is an independent, non-profit institution dedicated to improving human health and welfare through biological research, genomic medicine, education, and economic development. Since coming to HudsonAlpha, I have set up an independent human genomics research lab. My lab

works on projects related to the genetics of human disease, with a particular emphasis on research and clinical applications of genome sequencing for diagnosing children with developmental delays and intellectual disabilities. We also study the genetics and biochemistry of gene expression regulation, develop new computational approaches for analysis of human genomes and related datasets, and pursue other applications of genomic technologies to understand human biology.

4. I am also concurrently an Adjunct Associate Professor in the Department of Genetics at the University of Alabama at Birmingham. I have held an adjunct faculty position since 2010.

5. I earned a Bachelor of Arts in Microbiology and Bachelor of Science in Mathematics and Statistics, summa cum laude, from Miami University (Oxford, OH) in 2001. As an undergraduate student, I performed molecular biology and genetic research in research laboratories at Miami University, the University of Illinois at Urbana-Champaign, and the University of Cincinnati.

6. I earned a Ph.D. in Genetics from Stanford University in 2006 under the mentorship of Professor Arend Sidow, Ph.D. My thesis work was mainly comprised of computational biology research, with a focus on evolutionary comparisons of mammalian genome assemblies to better annotate functional elements in human genomes and assess whether mutations in those elements might contribute to disease. While there, I worked and/or received training in areas

including: theoretical and applied human genetics, molecular evolution, functional genomics, developmental biology, bioinformatics, genomic technologies, and analysis of datasets produced by high-throughput experimental platforms, especially DNA sequencers.

7. From 2006 to 2009, I was a Senior Fellow in the Department of Genome Sciences at the University of Washington working with Professors Evan Eichler, Ph.D., and Deborah Nickerson, Ph.D. While there, I conducted my postdoctoral research. I was primarily interested in applying genomic technologies to better understand the genetic basis for human disease. One series of projects centered on analytical and algorithmic developments to identify large deletion and duplication events, also known as “copy-number variants” (“CNVs”), from microarray data generated to perform Single Nucleotide Polymorphism (“SNP”) genotyping. This process involved extensive processing, normalization, and analysis of Illumina SNP array data. A second series of projects was aimed at better understanding the genetic basis for a variety of complex traits. These included genetic analyses of transcript levels, for example, and also a genome-wide association study of patient response to the commonly prescribed anticoagulant warfarin. Finally, another series of projects focused on analysis of high-throughput DNA-sequencing data, in particular the locations of genetic variants, from individuals with Mendelian diseases. I completed my postdoctoral studies in 2009.

8. From 2009 to 2010, I was an Acting Assistant Professor in the Department of Genome Sciences at the University of Washington.

9. I am an author or contributor to more than 100 publications in the areas of human genetics and genomics, including functional and evolutionary annotations of mammalian genomes, DNA sequencing, and DNA analysis, as they relate to human diseases, especially, but not exclusively, intellectual disabilities and other developmental issues. As of June 23, 2025, these publications have been collectively cited more than 50,000 times, according to Google Scholar estimates.

10. Among the awards and honors I have earned are: Harrison Scholar, Miami University (Oxford, OH) (1997-2001), Goldwater Scholar (1999-2001), Howard Hughes Medical Institute Doctoral Fellowship (2001-2006), a University of Washington Genome Training Grant Postdoctoral Fellowship (2007), and a Merck, Jane Coffin Childs Memorial Fund Postdoctoral Fellowship (2007-2010).

11. I am a Principal Investigator (“PI”) on a current human genomic research project funded by the National Institutes of Health (“NIH”) and a PI or co-Investigator on several additional projects funded by non-NIH sources, including both governmental and philanthropic sources. I have also been a PI or co-Investigator on multiple additional grants related to human genetics and genomics, from NIH and elsewhere, that have been completed.

12. In general, I have knowledge and experience with both the general scientific contexts relevant to this matter and specific details of the technologies employed. For example, as described here, I have extensive knowledge about molecular techniques such as DNA sequencing, molecular tagging, PCR, gene expression analysis, alignment and counting of sequencing reads, and detection of genetic variation such as changes to copy number.

13. I have also previously served as an expert witness in three now concluded patent disputes and two ongoing patent disputes. The three completed cases, in which I worked on behalf of the plaintiffs, are *Illumina/Verinata v. Ariosa/Roche* (C.A. No. 3:12-5501-SI, N.D. Cal.), *Guardant Health v. FMI* (C.A. No. 17-1616-LPS-CJB, D. Del.), and *Guardant Health v. PGDx* (17-1623-LPS-CJB, D. Del.). The two ongoing cases, in which I am working on behalf of the defendants, are *Natera v. ArcherDX/Invitae* (C.A. No. 20-125-GBW, D. Del.) and *10X Genomics v. Parse Bioscience* (C.A. No. 22-1117-JHS, D. Del.). I have also recently agreed to serve as an expert on behalf of Guardant as the plaintiff in an infringement suit against Tempus AI, Inc. (1:24-cv-00687-GBW, D. Del.), although I have not yet begun work on that case as of July 8, 2025.

14. A copy of my Curriculum Vitae, attached as Exhibit EX1023, contains further details on my education, experience, publications, and other qualifications to render an expert opinion in this matter.

15. For this service, I am paid my standard hourly consulting fee of \$750/hour, and I am reimbursed for actual expenses. My compensation in no way depends on the outcome of this matter or the substance of this declaration.

III. SUMMARY OF OPINIONS

16. Based on my knowledge, experience, and the materials that I have reviewed, it is my opinion that claims 1, 2, 9-11, 22, 26, 28, and 33-34 of the '510 patent are obvious based on U.S. Patent Application No. 2009/0029377 to Lo et al. ("Lo") (EX1010) in view of the knowledge of the skilled artisan regarding molecular counting by tagging, amplification, and sequencing. This knowledge of the skilled artisan is reflected in at least four references: (1) M. McCloskey et al., *Encoding PCR Products with Batch-stamps and Barcodes*, 45 Biochemical Genetics 761 (Dec. 2007) ("McCloskey") (EX1008), (2) U.S. Patent No. 7,537,897 to Brenner et al. ("Brenner") (EX1014), (3) U.S. Patent No. 9,085,798 to Chee ("Chee") (EX1015), and (4) U.S. Patent No. 11,840,730 to Porreca et al. ("Porreca") (EX1017).

17. It is also my opinion that the following claims are obvious based on the following combinations:

- Claim 3 is obvious based on Lo and the knowledge of the skilled artisan in further view of specific disclosure in McCloskey, Brenner, or Chee.
- Claim 4 is obvious based on Lo and the knowledge of the skilled artisan in further view of specific disclosure in Brenner or Porreca.

- Claims 5-8 are obvious based on Lo and the knowledge of the skilled artisan in further view of specific disclosure in McCloskey, Brenner, or Porreca.
- Claims 12, 15, 17, 23, and 35-36 are obvious based on Lo and the knowledge of the skilled artisan in further view of specific disclosure in Brenner, Chee, or Porreca.
- Claims 18-21 and 29-32 are obvious based on further knowledge of the skilled artisan.

IV. LEGAL STANDARDS

18. I am not a lawyer and offer no legal opinion in this declaration. I have, however, been advised by counsel as to various legal standards that apply to the technical issues I address in this declaration. I have applied those standards in arriving at my conclusions expressed in this declaration. The recitation of the legal standards set forth here is virtually identical to those in the contemporaneously filed IPR proceedings for U.S. Patent No. 10,947,589 (the “’589 patent”) and those I relied upon in a recent set of IPR proceedings where I testified on behalf of Parse Biosciences.

A. Claim Construction

19. I have been informed that the words of the claims of a patent are generally given the ordinary and customary meaning as interpreted by a person of

ordinary skill in the art (POSA) at the time of the invention. I have been informed that patent claims are construed in light of the claim language, the patent specification (including the drawings), and the prosecution history (including the references cited during the prosecution of the patent). For this declaration, I have laid out my opinions on what terms should be construed in § VIII on Claim Construction below; I have given all other terms their plain and ordinary meaning as would have been understood by a POSA at the time of the invention.

B. Obviousness

20. I understand that even if a patent is not anticipated, a claimed invention is “obvious” if the difference between the claimed subject matter and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a POSA.

21. I understand that a POSA provides a reference point from which the prior art and claimed invention should be viewed. This reference point prevents the use of one’s insight or hindsight in deciding whether a claim is obvious.

22. Also, I understand that a claim can be obvious in light of a combination of multiple prior art references. I understand that such a combination must meet certain requirements. I understand that the prior art must teach or suggest all the requirements of the invention as recited in the claim. I understand that the prior art, coupled with knowledge generally available in the art at the time of invention, must

contain some suggestion or incentive that would have motivated a POSA to modify a reference or to combine the references. I also understand that this proposed modification must have had a reasonable expectation of success.

23. I understand that only analogous art may be considered in an obviousness analysis. Art is considered analogous if it is in the same field of endeavor as the patent or if it is reasonably pertinent to the particular problem with which the patent is involved.

24. I also understand that in evaluating whether patent claims are obvious, objective indicia of non-obviousness can be considered, such as but not limited to, commercial success, long-felt but unsolved needs, praise, unexpected results, industry acceptance, failure of others, and skepticism by experts.

V. THE '510 PATENT

25. The '510 patent is titled "Varietal Counting of Nucleic Acids for Obtaining Genomic Copy Number Information." *See generally* EX1001. The claims of the '510 patent generally recite methods for determining genomic copy number information by tagging nucleic acid segments, making copies of the segments, sequencing the segments, and counting the number of unique tagged molecules at genome locations.

26. Claim 1 is representative and is reproduced below:

1. A method comprising:

- a) obtaining segments of genomic nucleic acids from a sample containing the genomic nucleic acids;
- b) randomly tagging the segments of genomic nucleic acids with different nucleic acid tags, thereby generating unique tagged nucleic acid molecules from the segments of genomic nucleic acids, such that each of the unique tagged nucleic acid molecules comprises a segment of the segments of the genomic nucleic acids from step (a) and a nucleic acid tag of the nucleic acid tags;
- c) subjecting the unique tagged nucleic acid molecules to a polymerase chain reaction (PCR), thereby generating copies of the unique tagged nucleic acid molecules;
- d) generating tag associated sequence reads by sequencing the copies of the unique tagged nucleic acid molecules of step (c);
- e) assigning each of the unique tagged nucleic acid molecules to a location on a genome by mapping each of the tag associated sequence reads of step (d) to a location in the genome; and
- f) in a plurality of locations in the genome, obtaining relative copy number information of the genomic nucleic acids from the sample based on the number of unique tagged nucleic acid molecules that have been assigned to each of the plurality of locations in the genome.

27. The background of the '510 patent explains that genomic copy number information was commonly obtained using whole genome amplification. *Id.* at 1:37-40. Whole genome amplification, like the name suggests, relies on amplifying, or

making copies of, the entire genome or substantially all of the entire genome of a cell, as opposed to specific target regions. The background notes that the whole genome amplification method oversamples certain regions and introduces distortions. Ligation-mediated PCR solved some issues with whole genome analysis but was supposedly still susceptible to amplification bias. *Id.* at 1:37-40, 51-55, 60-66. The '510 patent purports to address amplification bias through the well-known use of tagging, amplification, and sequencing to perform molecular counting.

28. Consistent with this, the specification of the '510 patent reflects that the use of tagging before sequencing for molecular counting as recited in the claims is not inventive but rather a prior art technique. The specification explains, for instance, that “Miner et al. (2004) describe a method of molecular barcoding to label template DNA prior to PCR amplification, [which] allows for the identification of contaminant and redundant sequences by counting on distinctly tagged sequences.” *Id.* at 2:12-16. Similarly, the specification notes that McCloskey et al. (2007) describes a method of barcoding “template DNA molecules prior to PCR amplification” and that U.S. Patent No. 7,537,897 (Brenner) describes “methods for molecular counting by labeling molecules of an input sample with unique oligonucleotide tags and subsequently amplifying and counting the number of different tags.” *Id.* at 2:16-20, 23-26.

29. The specification, in my opinion, offers no satisfactory explanation for how these prior methods are materially different from the alleged invention. McCloskey, for instance, is criticized as using “template specific primers to barcode template DNA molecules prior to PCR amplification,” stating that this approach has an alleged drawback of requiring primers for each species of template molecule that is studied. *See id.* at 2:23-29. Yet, nothing in the claims excludes this approach. To the contrary the patent clearly states that the tags can be introduced by a range of methods including “priming.” For example, the specification states, “[t]ags may be associated with nucleic acid molecules by using, singly or in combination, ligation, nick-ligation, terminal transferase, hybridization, priming, and the like.” *Id.* at 21:38-40. Thus, it cannot be the case that McCloskey is distinguishable from the alleged invention on the basis of introducing tags via priming. The ’510 patent’s alleged distinction over Miner and Brenner is even more peculiar. Specifically, the patent purports to criticize these references because they use ligation for tagging, which is supposedly an “inefficient” reaction. *See id.* at 2:20-22. This is peculiar because the patent has an entire section on how to perform “Tagging by Ligation.” *See id.* at 22:43-24:29. Thus, it simply cannot be the case that Miner and Brenner are distinguishable from the alleged invention on the basis of their use of ligation for tagging. In general, the claims of the ’510 patent recite prior art techniques for

adding tags, and neither the specification nor the claims offer any explanation or basis for distinguishing over those prior art tagging methods.

30. The specification of the '510 patent further discloses that PCR was an amplification method known in the art. Specifically, the specification explains that “[g]enerally...the [amplification] steps are performed multiple times (e.g., polymerase chain reaction).” *Id.* at 13:50-54. The Abstract also references “subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction,” and claim 1 recites “subjecting the...tagged nucleic acid molecules to a polymerase chain reaction (PCR), thereby generating amplified tagged nucleic acid molecules.” *Id.* at Abstract, Claim 1.

31. The specification of the '510 patent discloses that high-throughput DNA sequencing was also well-known in the art. Specifically, the specification notes that any method directed towards determining an accurate copy number “should also be robust using existing methodologies for high volume, massively parallel sequencing.” *Id.* at 2:37-39. Nowhere does the specification suggest that massively parallel sequencing was something new or inventive.

32. I do not have an opinion on the priority date of the '510 patent. Based on the face of the patent, I understand that the earliest possible priority date of the '510 patent is October 22, 2010, based on the filing date of U.S. Provisional Patent Application No. 61/406,067. I have considered October 22, 2010 to be the relevant

date for assessing prior art invalidity, despite the differences in claimed subject matter between the provisional application and the '510 patent.

VI. PROSECUTION HISTORY OF THE '510 PATENT

33. On February 7, 2019, Patent Owner filed Nonprovisional Patent Application 16/269,818 titled "Varietal Counting of Nucleic Acids for Obtaining Genomic Copy Number Information." EX1009. Application 16/269,818 is a continuation of Patent Application No. 15/063,278 which issued as U.S. Patent No. 10,947,589 (the "'589 patent"). The '589 patent has been challenged in a co-pending IPR petition, and I have submitted a declaration providing my independent evaluation on the validity of the claims of the '589 patent.

34. I have been informed by counsel that during the prosecution of the '510 patent, the Examiner issued several non-final rejections on grounds including claim objections, indefiniteness, and enablement. EX1009.221-235, .310-17, .497-502. The Examiner also issued two final rejections, the last of which was on grounds including claim objections, indefiniteness, and obviousness. EX1009.271-281, .429-442.

35. On February 25, 2025, the '510 patent was issued. EX1009.540-541. In the Reasons for Allowance, the Examiner stated "The closest prior arts in the record are Casbon *et al.*, (US 8,715,967 B2, priority date: September 21, 2010) and Xie *et al.*, (BMC Bioinformatics, 10, 80, March 6, 2009). There [sic] prior arts do

not teach a combination of steps a) to f) of claim 68. These prior arts either alone or in combination with the other art in the record do not teach or reasonably suggest a method which comprises all of the limitations recited in claim 68.” *Id.* at .531-532.

36. In the last final rejection before issuance, the Examiner found the applications claims obvious as unpatentable over Casbon *et al.* (U.S. Patent No. 8,715,967) in view of Xie *et al.* (BMC Bioinformatics 10, 80, March 6, 2009). *Id.* at .434. In overcoming the Examiner’s obviousness objections, the Patent Owner emphasized that the claimed method (1) randomly tagged segments of genomic nucleic acids to generate the unique tagged nucleic acid molecules, (2) assigned each amplified unique tagged nucleic acid molecule to a location on the genome by mapping each tag associated sequence read to a location on the genome, and (3) obtained copy number information based on a count of the number of unique tagged nucleic acid molecules that were assigned the same genome location, rather than based on the number of different tags. *Id.* at .410; .531-532. Yet as explained below, the Examiner did not consider additional references that render these distinctions obvious.

37. During the prosecution of the ’510 patent, the Patent Owner identified Brenner and McCloskey in an IDS. *Id.* at .180; .183. The Examiner did not cite to or discuss Brenner or McCloskey in any of the non-final or final rejections. As discussed below, §§ XI.B.2, XI.C.2 *infra*, Brenner and McCloskey disclose the DNA

tagging, amplification, sequencing, and counting limitations of claim 1 of the '510 patent. Lo was not presented to the Examiner. Lo discloses the concept of obtaining DNA segments from a sample and “assigning each of the unique tagged nucleic acid molecules to a location on a genome by mapping each of the tag associated sequence reads...to a location in the genome” to determine copy number information as recited by claim 1 of the '510 patent.

VII. RECENT BOARD UNPATENTABILITY DECISION IN IPR PROCEEDINGS BETWEEN PARSE BIOSCIENCES AND 10X GENOMICS

38. I understand that this is not the first time that the Board has reviewed a patent directed towards nucleic acid tagging and counting. In fact, I provided testimony in IPR proceedings between Parse Biosciences, Inc. and 10x Genomics, Inc. that involved U.S. Patent No. 10,697,013 to Brenner et al. (the “’013 patent”) related to the use of tagging, amplification, and sequencing for molecular counting, which even relied on one of the same prior art references as the instant petition, namely the McCloskey reference. The specification of the ’013 patent explains, for instance, that “MIDs [tags] are employed to uniquely tag each of individual polynucleotide in a sample. Identification of the number of unique MIDs in a sample can provide a readout of how many individual polynucleotides are present in the sample.” EX1024 at 6:53-57.

39. Specifically, on February 5, 2024, the Board granted institution of *inter partes* review of the '013 patent in proceeding IPR2023-00958. EX1025 at 2. The '013 patent claimed priority to a provisional application filed on August 20, 2009, and this priority date was not contested during the IPR proceedings involving the '013 patent. The key point is that this priority date is well before the priority date of the patent at issue here. Therefore, to the extent that the '013 patent that was at issue in IPR2023-00958 is invalid, it is my opinion that the '510 patent would be similarly invalid, as I explain further below.

40. Indeed, claim 1 of the '013 patent contains many of the same elements as the '510 patent. Step (a) consists of “providing a sample comprising . . . a plurality of polynucleotides.” Step (b) consists of “performing combinatorial tagging to generate a plurality of tagged polynucleotides.” Step (c) consists of “amplifying said tagged polynucleotides.” This tagging process in step (b) involves introducing a “second tag sequence” that is used for “distinguishing said sample polynucleotide from other sample polynucleotides from said cell.” Step (d) consists of “sequencing said plurality of amplified polynucleotides.” Finally, step (e) consists of “using the sequences in step (d) to count sample polynucleotides.” Thus, claim 1 of the '013 is directed to the same general counting process as claimed in the '510 patent involving tagging, amplification, and sequencing. *See generally* EX1024.

41. After a full review of the merits, the Board found that all claims were unpatentable as obvious in light of prior art. In its final written decision issued on February 3, 2025, claims 1-12 and 19-28 were found unpatentable as obvious using Linnarsson (EX1026) in combination with McCloskey and further in view of the knowledge of a POSA. Similarly, claims 13-18 were found unpatentable as obvious using Linnarsson in combination with McCloskey and a second McCloskey reference (EX1027) and further in view of the knowledge of a POSA. EX1005 at 51.

42. As noted above, I submitted testimony in support of the petitioner in the IPR proceedings. *See* EX1028; *supra* ¶ 33. I reserve the right to rely upon my opinions in these prior proceedings here. Given the Board's decision to invalidate the '013 patent in the previous proceeding where I testified, it is difficult to see how the '510 patent claims could not be invalid, particularly when one considers that the '013 patent has an earlier priority date.

VIII. CLAIM CONSTRUCTION

43. I understand that for purposes of this Petition, the claims will be interpreted consistent with the Patent Owner's arguments and positions in the parallel district court proceeding, particularly in view of its contentions regarding how the '510 patent is allegedly infringed. EX1020.

44. I have been informed that the words of the claims of a patent are generally given the ordinary and customary meaning as interpreted by a POSA in the art at the time of invention. I understand that patent claims are construed in light of the claim language, the specification (including the drawings), and the prosecution history (including references cited during the prosecution of the patent).

IX. PERSON OF ORDINARY SKILL IN THE ART

45. I understand that whether a patent is obvious is analyzed from the perspective of a POSA at the time of the alleged invention. I am knowledgeable about the level of ordinary skill in the art of the time of the alleged invention based on my education and experience. *See supra* § II.

46. In my opinion, a POSA would have been someone with a Master's degree and/or Ph.D. in molecular biology, genetics, chemistry, engineering, or a related discipline and at least two years of post-Master's, postdoctoral, or industry experience. Alternatively, a POSA would have been someone with a Bachelor of Science degree in molecular biology, genetics, chemistry, engineering, or a related discipline and at least five years of industry or academic experience (including any experimental work toward a graduate degree) relating to DNA labeling techniques, DNA amplification methods, and DNA sequencing technologies.

47. Further, it is my opinion that a POSA would have been familiar with associated tools, methods, and techniques including: (1) DNA tagging, (§ X.C); (2)

nucleic acid amplification, (§ X.B); (3) use of tags to overcome amplification bias, (§ X.D); (4) sequencing and counting sequencing reads, (§ X.E); and (5) copy number (§ X.F).

X. TECHNICAL BACKGROUND AND STATE OF THE ART

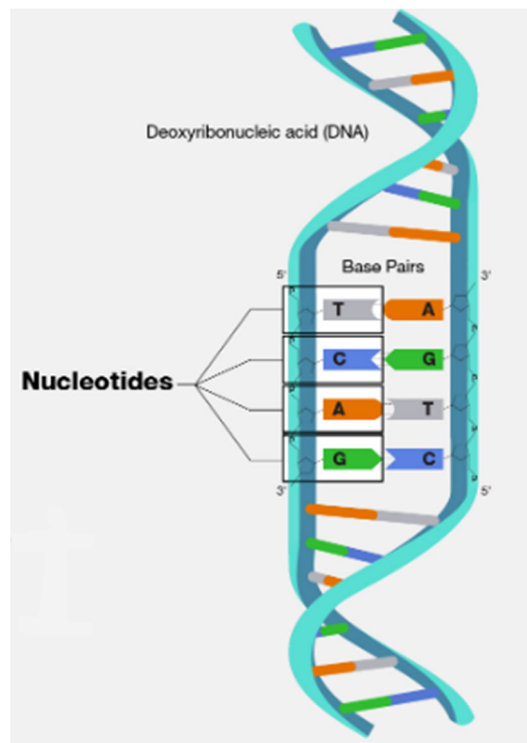
48. Below, I provide an overview of the relevant technology and the state of the art. Much of this background discussion is similar to that from my prior testimony on behalf of Parse Biosciences.

A. DNA

49. DNA, deoxyribonucleic acid, is a biological molecule that contains the information needed for cellular life. Individual cells use DNA for many aspects of cell development and maintenance. This information is stored in the sequence of adjacent bases in the DNA strand, or a “nucleotide sequence.” DNA is one type of “nucleic acid” and is thus commonly referred to as such.

50. DNA is a double-stranded molecule composed of monomeric units called nucleotides. A polynucleotide is a linear polymer comprised of multiple nucleotides. An oligonucleotide, or “oligo,” is a short polynucleotide. Oligonucleotides are commonly used, *inter alia*, as primers with various biological techniques such as DNA sequencing, polymerase chain reaction (“PCR”), and molecular cloning.

51. In DNA, each nucleotide comprises a deoxyribose sugar molecule, one of four nitrogenous bases, and one or more phosphate groups. The four nitrogenous bases of DNA are: adenine (“A”), thymine (“T”), cytosine (“C”), and guanine (“G”). The two strands of DNA are connected through complementary base pairing via chemical bonds, also known as “hybridization.” A pairs with T, and C pairs with G, as depicted below:



EX1029 at 2.

52. Each strand of the DNA molecule has two ends – one end is referred to as 5’ and the other end is referred to as 3’. DNA is synthesized in the 5’ to 3’ direction.

B. Amplification

53. Before it can be sequenced, a sample with a limited amount of DNA must sometimes be subject to amplification. Amplification refers to the process of creating multiple copies of one or more polynucleotides present in an input sample, and can be applied to the whole genome or to fragments depending on the specific technique employed. Various methods of amplification were known in the art before 2011, such as polymerase chain reaction (“PCR”), ligase chain reaction (“LCR”), nucleic acid based amplification (“NASBA”), and transcription-mediated amplification, etc. *See, e.g.*, EX1008 at Abstract (“Polymerase chain reaction (PCR) has become the mainstay of DNA sequence analysis”).

54. Brenner and McCloskey each describe amplification steps that were known and widely practiced in the art. *See* EX1014 at 33:6-27; EX1008 at 764-767. McCloskey, Chee, and Lo all recognize PCR reactions as a conventional form of amplification. EX1008 at Abstract (“Polymerase chain reaction (PCR) has become the mainstay of DNA sequence analysis.”); EX1015 at 3:2-3; EX1010 ¶ 36. Indeed, the ’510 patent also acknowledges the use of the conventional amplification technique of PCR, described in detail below, as part of its claims and description. EX1001 at Abstract, Claim 1.

55. PCR is a technique that uses repeated cycles of DNA synthesis to amplify target nucleic acid sequences. Each cycle roughly doubles the amount of

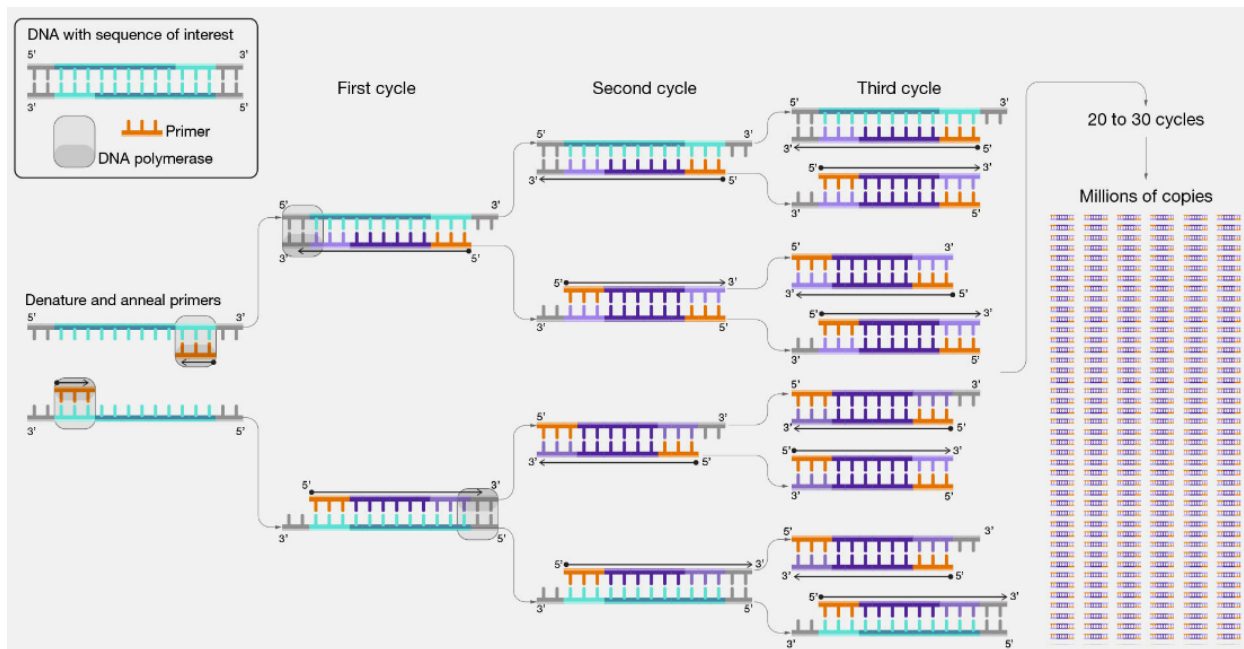
target DNA. PCR generally involves three steps: (1) denaturation of the double-stranded DNA, (2) annealing of primers, and (3) primer extension. EX1030 at 1154.

56. During denaturation, double-stranded DNA molecules are separated into complementary single strands of DNA by using heat or other chemical or physical means of dissociation. *Id.* The two complementary strands of single-stranded DNA will be used as a template for replication. *Id.*

57. During annealing, primers or probes, which are oligonucleotides complementary to a portion of one of the original DNA strands, attach to the single-stranded DNA templates. When the reaction mixture comprising denatured DNA template strands is cooled to a certain temperature, the primers hybridize, i.e., bind, through base pairing interactions to either the left (5') or right (3') side of the sequence of interest. *Id.* The portions of the templates to which the primers bind are called priming sites.

58. During primer extension, each molecule of single-stranded DNA is converted to double-stranded DNA, doubling the number of starting strands. A thermostable enzyme, called "DNA polymerase," catalyzes the synthesis of new strands of DNA through primer extension. The enzyme adds nucleotides complementary to those in the single-stranded DNA template starting at the 3' end of the annealed primer. *Id.* Upon completion of one cycle, the DNA strands are subjected to additional cycles, i.e., the same steps of denaturation, primer annealing,

and primer extension. This results in an exponential amplification of target DNA sequences, as exemplified below.



EX1031 at 1.

59. Methods for applying PCR amplification to sample polynucleotides were well established and known in the prior art. *See* EX1014 at 5:37-42; EX1008 at Abstract.

C. DNA Tagging

60. Tagging is a known concept in genetics that involves adding a synthetic polynucleotide to a sequence of interest. The '510 patent notes in its background section that DNA tagging was used in a variety of ways in the art. For example, the '510 patent refers to “a method of molecular barcoding to label template DNA prior to PCR amplification,” which “allows for the identification of contaminant and

redundant sequences by counting only distinctly tagged sequences.” EX1001 at 2:12-16. The ’510 patent further acknowledges that various methods of tagging existed in the art. *See id.* at 2:20-22 (describing labeling by ligation); *id.* at 2:23-26 (describing labeling through specific primers); *id.* at 21:38-40 (“Tags may be associated with nucleic acid molecules by using, singly or in combination, ligation, nick-ligation, terminal transferase, hybridization, priming, and the like.”).

61. The field has been replete with routine methods of tagging for a variety of uses including acting as a unique “barcode” identifier (EX1017 at 31:5-8), providing primers to facilitate amplification (EX1008 at 762), counting and tracking molecules (EX1014 at 9:23-24), detecting errors in amplification and sequencing (EX1015 at 10:21-28), and/or removing amplification bias (EX1007 at 1).

62. It was a common practice in the art to introduce molecular tags to label nucleic acids before amplification. *See, e.g.*, EX1007 at 1; EX1015 at 3:1-3; EX1014 at 6:52-57; EX1008 at 762. Indeed, the ’510 patent itself cites McCloskey as well as Miner et al., which neatly explains how tagging “allows for quantification of the relative abundance of genomic” patterns and sequences “by correcting for skewing that can arise from PCR amplification or the cloning of the products.” EX1007 at 3.

63. Molecular tags have also been commonly known in the art to track the origin of nucleic acids. The concept of a “molecular barcode”—i.e., a barcode that

identifies the set of replication products derived from the same original molecule that was amplified or a barcode that identifies each sample in a pool of samples, for example—was known in the prior art. This includes the use of nucleic acid sequences to track which patient a sample came from. *See, e.g.*, EX1017 at 31:5-8. Porreca further discloses the use of a “differentiator tag sequence” which, when combined with target DNA templates, reduces the effects of bias by allowing for the identification of unique input molecules after amplification. *Id.* at 28:49-67.

64. The concept of using multiple tags or a combination of tags was also known. For example, McCloskey discloses a method of DNA analysis for single-cell samples, among others, by utilizing an encoding oligonucleotide with five information regions, two of which are tag sequences—one being a barcode that uniquely identifies and distinguishes one polynucleotide from another and another being a batch-stamp which associates a polynucleotide to “a specific sample on a specific date.” EX1008 at 762, 764. Chee also discloses tagging a DNA sample with a combination of tags. EX1015 at 20:12-15.

65. It was also known in the art that using a large number of tags is one way to ensure that each polynucleotide in a reaction mixture is tagged. For example, Chee discloses that the “large number of unique sequences that can be generated from the random assembly of nucleotides provides a high probability that each individual construct will be uniquely identified, even those constructs that contain

substantially the same” target DNA sequences. *Id.* at 11:26-30. Additionally, Porreca discloses using probes that “contain one of a large number of differentiator tags in their backbone such that the probability of any two [probes] in a reaction having the same differentiator tag sequence is low.” EX1017 at 60:49-52. Brenner also discloses that “a size of a tag population that ensures with high probability that substantially every target polynucleotide will have a unique tag.” EX1014 at 18:2-5. Using a large number of tags was a typical and routine method of increasing the chances that every polynucleotide from the original sample will be tagged and eventually sequenced. McCloskey, for instance, gives a general description of how one can tune the length of the barcode to distinguish among a specific number of molecules. As McCloskey explains, the “number of distinguishable barcodes in a population of oligonucleotides used in a reaction is determined by the number of random bases, n . This enables one to distinguish among 4^n allele copies per reaction.” EX1008 at 764.

66. Thus, tagging was a well-known step used as a prelude to sequencing polynucleotides. In fact, Dr. Michael Wigler, the named inventor of the '510 patent, admitted in a scientific publication that “[m]odifying molecules to facilitate counting is not a new idea.” EX1004 at 1. Dr. Wigler noted that the idea had been applied under a variety of names to identify PCR duplicates and improve counting of DNA templates. *Id.* He further cited several references that disclose the modification of

molecules to facilitate tagging, many of which are prior art to the '510 patent. *Id.* at 4.

D. Amplification Bias and Molecular Tagging

67. One of the known issues with nucleic acid amplification is amplification bias. Because efficiency of the PCR process can vary for both systematic and stochastic—i.e., randomized—reasons, the relative quantity of each sequence in the post-amplification sample often may not accurately reflect the distribution of such samples in the pre-amplification sample. *See* EX1032 ; EX1033. For instance, if an original sample had two DNA molecules having a first DNA sequence and four DNA molecules having a second DNA sequence, this 1:2 ratio may not be preserved by amplification. PCR amplification may, for instance, favor the first DNA sequence over the second sequence (e.g., due to random failure of primer binding to the second sequence in early PCR cycles), such that after amplification one ends up with roughly 10,000 molecules of the first sequence and 10,000 molecules of the second sequence. Whereas the ratio was 1:2 prior to amplification, it is roughly 1:1 after amplification. Such problems cloud the ability of researchers or clinicians to understand, for example, how much of a particular gene is being expressed in a particular cell of interest or the copy number of a particular gene in the DNA of an input sample.

68. This amplification bias problem can be particularly severe when the amplification starts from a small number of template nucleic acid molecules. A low number of template DNA requires more rounds of amplification, which exponentially increases the effect of bias. Additionally, there is an increased risk of amplifying any contaminant DNA in the starting mixture, which results in a skewed end product. EX1007 at 1. Thus, reducing amplification bias is of particular importance when working with a limited DNA sample.

69. To solve this problem, it was well-known and very common to add molecular tags to target sequences prior to the amplification step. *See, e.g.*, EX1007 at 1; EX1008 at 766; EX1015 at 2:65-3:3; EX1014 at 1:66-2:2; EX1017 at 5:12-17. For example, one can use a large number of tags that each have a random DNA sequence, and the tags are randomly applied to the sample molecules. If there are multiple DNA molecules in the input sample that are the same, after the tagging process they will almost certainly be distinguishable from one another because the molecules will likely have all received a different tag assuming one uses a large enough number of tags. One can then amplify the tagged molecules and, if there is bias in the amplification, the bias can be corrected using the tags. This is because even if the amplification process preferentially amplifies particular DNA sequences over others, the total number of molecules of any given species that were present in the original sample can nonetheless be determined by counting the number of unique

tags associated with that sequence. This results in addressing amplification bias by keeping track of the polynucleotides in the original sample. See EX1007 at 1; EX1008 at 766.

70. Tagging has been a known solution to address the problem of amplification bias. For example, McCloskey suggests the use of a multiple tag system wherein sample polynucleotides derived from a sample are tagged so that after amplification “nonredundant” barcodes can be used to count the number of original molecules. See EX1008 at 766 (“An additional benefit of our method for mosaic samples is that counting sequences with nonredundant barcodes provides a quantitative measure of template diversity.”)

71. The Miner reference teaches a similar process:

In response to the challenges of PCR redundancy and contamination associated with PCR amplification of limited amounts of DNA template, we have labeled genomic DNA fragments with molecular sequence barcodes and ‘batch-stamps’ prior to PCR amplification. This was accomplished by including these molecular labels in the hairpin linker sequence that we use in hairpin-bisulfite PCR (5). This encoded information enables us to track the genomic origin of each sequence obtained from PCR and subsequent bacterial cloning. *Each genomic fragment is marked prior to amplification, allowing us to identify contaminant and redundant sequences and to quantify accurately the proportion of cells carrying a particular sequence variant by counting only distinctly tagged sequences.* This highly sensitive method offers

confirmation of the independent genomic origin of all sequences in final data sets derived from PCR amplification.

EX1007 at 1. As I explain further below, there are yet additional references that describe this tagging concept for counting.

E. Sequencing

72. Nucleic acid sequencing is the process of determining the sequence of nucleotides (A, T, C, and G) in a nucleic acid molecule. The development of sequencing technologies can be traced to at least 1977 when Frederick Sanger developed DNA sequencing technology based on a chain termination method. EX1034 at 133, 140.

73. DNA sequencing technologies went through major improvements in subsequent years. In the early 2000s, next generation sequencing (“NGS”) technologies emerged that were different from the Sanger method in terms of throughput and cost. *Id.* at 133. By approximately 2008, techniques and platforms for NGS were commercially available and in widespread use including the Roche/454 FLX, Illumina/Solexa Genome Analyzer and Applied Biosystems SOLiD System. *Id.*

74. NGS is a type of high-throughput sequencing technology that uses “massively parallel sequencing” of many fragments of DNA. The platforms carrying out NGS are diverse in sequencing and detection biochemistry, but

massively parallel sequencing enables these sequencers to process and detect millions of sequence reads simultaneously. *Id.* DNA sequencing using NGS generally requires a minimum amount of DNA.

75. Once the sequences of such molecules have been obtained, they can be further analyzed using a so-called “reference genome.” *See id.* at 138. A reference genome is a genome sequence from the species that is being analyzed for which the sequence is already known. When sequence information from an unknown sample is determined, one can compare that unknown sequence information to the reference genome to determine where the unknown sequence came from in the species of the genome. The process of comparing an unknown sequence to a reference genome is known as “mapping” or “alignment.”

76. Molecular tags, whether used as barcodes, batch-stamps or for other purposes, are identified in the sequencing results. In McCloskey, for example, valid molecules were identified by detecting the valid batch-stamp sequence TCGAGCCT in the full molecular sequences. EX1008 at 763. Similarly, contaminant molecules were identified through the invalid batch-stamp, GTACCAT. *Id.*

77. Although a substantial improvement over previous sequencing technologies, NGS sequencing is not perfect and can produce both random and systematic errors. As just one example, some sequencing techniques face detection

issues when sequencing long sequences of the same nucleotides (e.g., distinguishing between TTTT and TTTTT). *See, e.g.*, EX1034 at 133; EX1017 at 57:10-14.

78. Sequencing errors were a known problem and one for which various solutions were developed. *See, e.g.*, EX1015 at 10: 54-57 (“In addition, error can be introduced as a result of sequencing error; this can be overcome by sequencing multiple instances of the same [target sequence]”); EX1014 at 21:14-18 (“Preferably, sequences of oligonucleotide tags are selected so that they can be distinguished even in the presence of sequencing errors. Thus, oligonucleotide tags are selected to be maximally different.”).

F. Copy Number

79. Genes are sequences of DNA nucleotides and constitute functional units carrying hereditary information. Genes are located at specific positions on specific chromosomes. An individual typically inherits two copies of each gene, one from each parent.

80. In certain situations, however, copy number changes can occur such that genes are duplicated or deleted, resulting in a deviation from the standard two-copy state. Duplications of a gene could create more than two copies of a gene being present, while deletions could lead to one or zero copies of a gene being present. The concept of copy number simply pertains to the frequency with which a gene, or some other genetic region, is present in a person’s genome.

81. Copy number changes are important to the study of cancer, among other things. For example, a form of cancer might be caused by the deletion of a tumor suppressor gene in a cell. Various prior art techniques were directed towards identifying copy number changes, including the Brenner references that I rely upon here. EX1014 at 1:55-60.

XI. THE PRIOR ART REFERENCES OF THE GROUNDS

A. Lo (Ex. 1010)

1. Lo Is § 102 Prior Art

82. U.S. Patent Application Publication No. 2009/0029377 (“Lo”) entitled “Diagnosing Fetal Chromosomal Aneuploidy Using Massively Parallel Genomic Sequencing,” was published on January 29, 2009. EX1010. I understand that Lo is prior art under pre-AIA 35 U.S.C. § 102(a) and (b) because Lo was published at least one year before the earliest claimed priority date of the ’510 patent. Lo is analogous art to the ’510 patent because it is in the same field of genetic analysis and is directed to the problem of using molecular counting to determine copy number information.

2. Lo Overview

83. Lo teaches a method for determining whether a chromosome or chromosome region imbalance exists within a biological sample. EX1010 ¶ 14. That is, Lo teaches a method for determining copy number information from a sample. Specifically, Lo discloses diagnostic testing methods for chromosomal

defects by “determining imbalances between different nucleic acid sequences”. *Id.* ¶ 3.

84. Fetal chromosomal aneuploidy is a condition where a fetus has, for instance, an abnormal number of chromosomes. *Id.* ¶ 4. The number of chromosomes “can be abnormally high, e.g. the presence of an extra chromosome 21 or chromosomal region in trisomy 21; or abnormally low, e.g. the absence of a copy of chromosome X in Turner syndrome.” *Id.*

85. Lo’s method can be applied to determine copy number information due to mutated genetic sequences, genetic polymorphisms, or copy number changes. *See, e.g., id.* ¶ 37 (“Yet other examples include sequences which are mutated, deleted, or amplified in a malignant tumor, e.g. sequences in which loss of heterozygosity or gene duplication occur.”); *id.* ¶ 109 (“In another example the sequencing technique on plasma cell-free DNA may be used to detect the chromosomal aberrations in the plasma DNA for the detection of a specific cancer. Different cancers have a set of typical chromosomal aberrations. Changes (amplifications and deletions) in multiple chromosomal regions may be used. Thus, there would be an increased proportion of sequences aligned to the amplified regions and a decreased proportion of sequences aligned to decreased regions.”).

86. Lo’s method begins with obtaining a biological sample from, for instance, a pregnant female. *Id.* ¶ 54. The sample contains nucleic acid molecules

from the fetus and the mother. *Id.* These nucleic acid molecules are fragments from chromosomes. *Id.* Lo discloses that this sample may be enriched or sub-selected for specific sequences before being sequenced in the next step. *Id.* ¶¶ 47, 72, 77-79. Next, the nucleic acid molecules from the biological sample are sequenced. *Id.* ¶ 55. The sequences of the molecules are frequently referred to in Lo as “sequenced tags” or just “tags.” To avoid confusion, it is important to note that Lo’s use of the term “tags” in this manner does not connote the “tags” that I refer to above, which are the short DNA molecules that are appended to DNA molecules for identification purposes. Rather, in Lo, the terms “sequenced tags” or “tags” merely refer to a DNA sequence in the sample that comes from a particular chromosome or chromosome region.

87. Nevertheless, as Lo goes on to explain, the “resultant sequenced reads form a representative profile of the mix of nucleic acid species in the original specimen.” *Id.* ¶ 57. Lo teaches that an imbalance in the number of sequence reads coming from a chromosome or chromosome region is indicative of a copy number change and can be inferred from the percentage representation of the sequences from the region of interest relative to the other “sequenced tags” of the specimen. *Id.* ¶ 67.

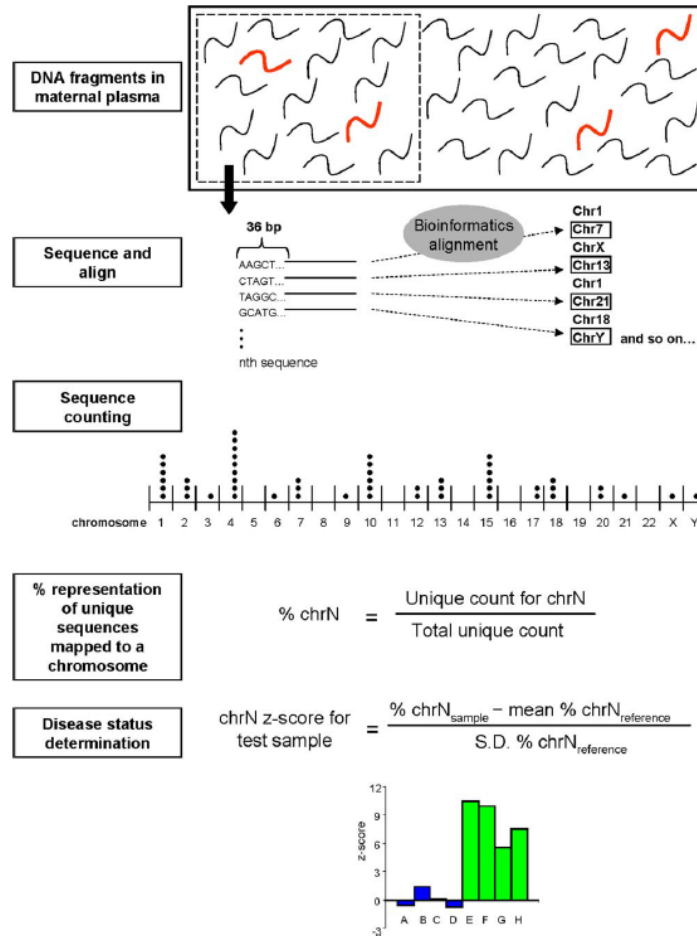
88. Lo discloses several ways to assess the percentage representation of a chromosome or chromosome region. *Id.* ¶ 60. These can include “counting the

number of sequenced tags, the number of sequenced nucleotides (basepairs) or the accumulated lengths of sequenced nucleotides (basepairs) originating from particular chromosome(s) or chromosomal regions.” *Id.* Lo teaches aligning the sequenced tags to a human reference genome sequence to determine chromosomal origin. *Id.* ¶ 70. This process of “aligning” is nothing more than the process of determining which particular chromosome or chromosome region a DNA sequence came from, and was known long before the priority date of the ’510 patent.

89. From there, the total number of individual “sequenced tags” aligned to each chromosome or chromosome region can be tabulated. *Id.* Lo then teaches determining a ratio of the number of sequence reads from the chromosome or chromosome region of interest to some other “reference” chromosome or group of chromosomes within the same sample that one expects will not have a copy number change. *Id.* ¶ 62. As Lo explains, this ratio may be a percentage of the total number of reads coming from the chromosome or chromosome region of interest relative to all of the other reads. *Id.* ¶ 63. Finally, this ratio or percentage is compared to one or more cutoff values, which in turn, allows for a determination of whether a copy number change exists for the first chromosome or chromosomal region. *Id.* ¶¶ 64, 66. Lo discloses that the cutoff values are empirically determined, using an external standard based on normal genomes. As Lo states, the “parameter (*e.g.*, the fractional representation of the clinically relevant chromosome) is then compared to a

reference range established in pregnancies involving normal (i.e. euploid) fetuses.” *Id.* ¶ 65; *see also id.* ¶ 70 (“The total number of individual sequenced tags aligned to each chromosome were tabulated and compared with the relative size of each chromosome as expected from the reference human genome or non-disease representative specimens. Chromosome gains or losses were then identified.”); *id.* ¶ 95 (“the chromosome 21 over-representation could be detected by comparing the percentage presentation of chromosome 21 obtained experimentally with the percentage representation chromosome 21 sequences expected for a euploid human genome.”).

90. FIG. 1 of a related 2008 publication by Lo (EX1011) neatly explains the method in Lo:



EX1011 at 20459. Fragments of fetal DNA are highlighted in red. From each fragment, 36 basepairs are sequenced. That short sequence is used to identify which chromosome or chromosome region the fragment comes from. The count of the number of fragments from each chromosome or chromosome region is compared with the total count of fragments to compute a percentage of sequences mapped to a particular chromosome or chromosome region. That calculated percentage is compared against the percentage for that chromosome or chromosome region in a reference standard genome. Samples containing several standard deviations more

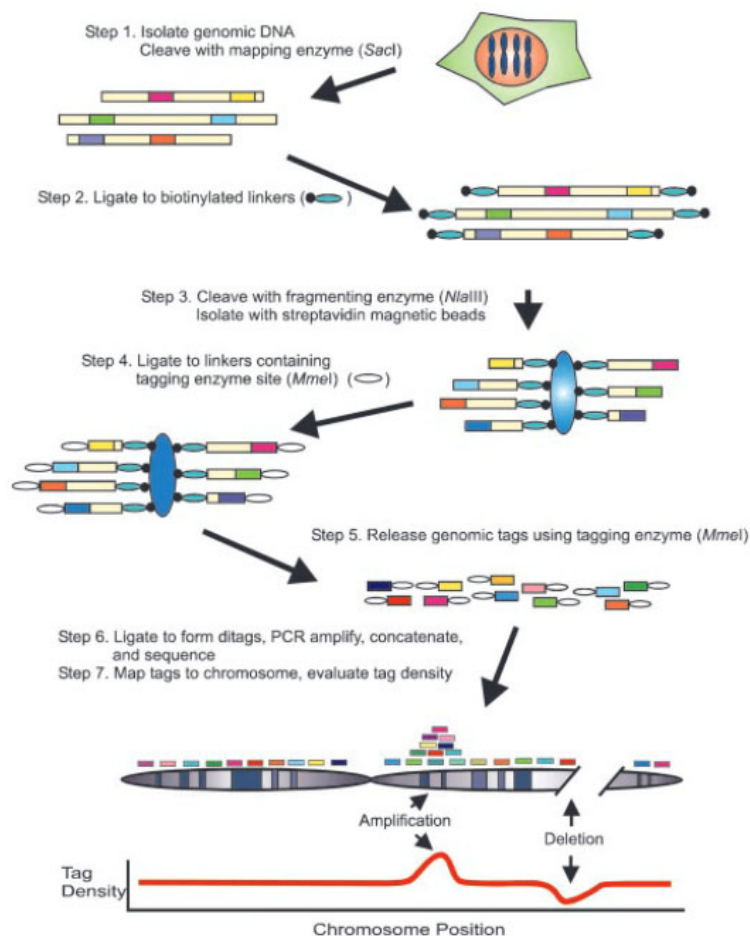
chromosomal content than the reference genome are likely aneuploid and are highlighted in green.

91. The method set forth in Lo based on looking for an imbalance in the number of sequence reads that map to a particular chromosome or chromosome region was not by any means a new method, but was rather an old idea that had been applied in many other contexts. Lo, in fact, acknowledges that the prior art discloses other such methods.

92. Lo, for instance, explains that the prior art Dhallan references teach similar methods: According to Lo, in Dhallan's method, the "proportion of chromosome 21 sequences contributed by the fetus in maternal plasma was determined by assessing the ratio of paternally-inherited fetal-specific alleles to non-fetal-specific alleles for single nucleotide polymorphisms (SNPs) on chromosome 21. SNP ratios were similarly computed for a reference chromosome. An imbalance of fetal chromosome 21 was then inferred by detecting a statistically significant difference between the SNP ratios for chromosome 21 and those of the reference chromosome. . . ." EX1010 ¶ 8.

93. While the examples and embodiment in Lo focuses on the use of cell-free DNA from pregnant women, wherein the sample comprises DNA from multiple genomes (i.e. mother and fetus), the basic counting method described in Lo was a standard method that had been previously applied straightforwardly to other sources

of single-genome DNA to detect copy number information. For example, Wang et al. describes a tagging and sequencing method to detect copy number changes called “digital karyotyping.” EX1012. Wang and co-workers described characterizing “210,245 genomic tags from the lymphoblastoid cells of a normal individual (NLB) and 171,795 genomic tags from the colorectal cancer cell line (DiFi).” *Id* at 16158. In step 7 of FIG. 1, reproduced below, an abnormal number of sequences mapped to a chromosome region indicates amplification while an absence indicates deletion:



Id. at 16157. The bottom of the figure with the chart labeled “Tag Density” most clearly illustrates how Wang’s method identifies copy number information (either an amplification or a deletion) by looking for imbalances in the number of sequence reads mapping to a region of a chromosome.

94. As another example, a 2009 paper by Chiang et al. titled “High-resolution mapping of copy-number alterations with massively parallel sequencing” describes a similar segmentation algorithm to identify copy number changes using a variable window size. EX1013 at 100-101. This method was applied to sequencing data “from three tumor cell lines . . . and their matched normal cell lines.” *Id.* at 101. Sequencing counts in a given genomic window were compared between the tumor and reference samples to identify “194 copy-number alterations in the HCC1954 cell line, 126 alterations in the HCC1143 cell line and 15 alterations in the NCI-H2347 cell line.” *Id.*

95. Thus, while the Lo publication focuses primarily on applications where the sample was cell-free DNA, I wish to make clear that there is nothing about the method in Lo that would limit it to such applications. In fact, a POSA would immediately recognize and understand that Lo’s method is applicable to other sources of DNA and had, in fact, been applied to other DNA sample types that were not cell-free DNA well before Lo’s patent publication.

B. McCloskey (Ex. 1008)

1. McCloskey Is § 102 Prior Art

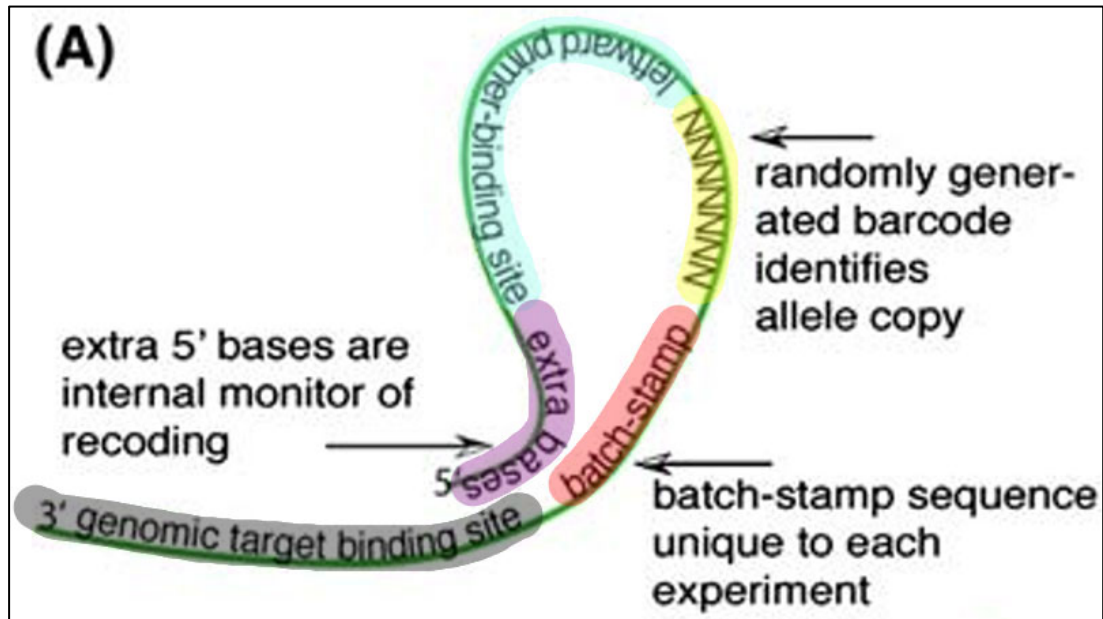
96. McCloskey *et. al.*, “Encoding PCR Products with Batch-Stamps and Barcodes,” is an article that was published in Biochemical Genetics on October 23, 2007. EX1008. I understand that McCloskey is thus prior art to the ’510 patent under pre-AIA 35 U.S.C. § 102(a) and 102(b).

2. McCloskey Overview

97. McCloskey is one of four references that I rely upon in this declaration that teaches the general approach of tagging, amplification, and sequencing for molecular counting.

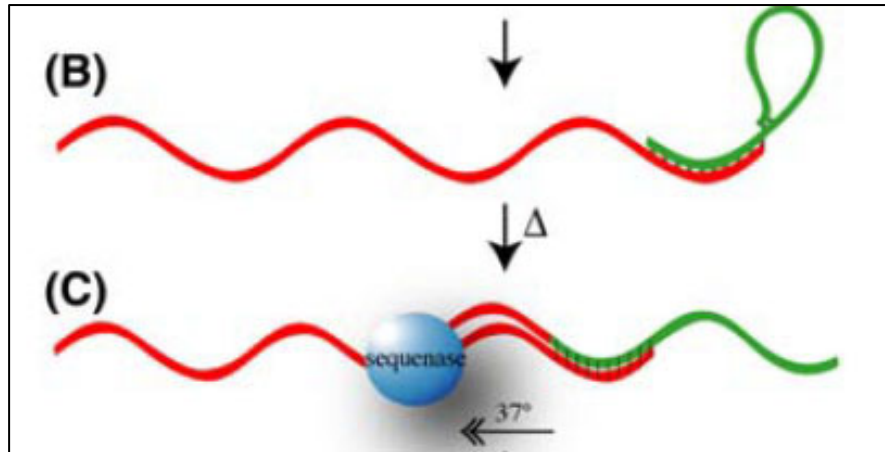
98. McCloskey describes a method of DNA analysis for single-cell samples (among other sample types) that utilizes an “encoding oligonucleotide” having “five distinct information regions,” two of which are tag sequences that distinguish, respectively, polynucleotides and samples. *Id.* at 762. As depicted below, the first tag is a “barcode” (highlighted in yellow), which is a randomly generated sequence intended to uniquely identify and distinguish one polynucleotide from another. *Id.* (“[T]he random barcode distinguishes among sequences arising from different cell or allele copies.”). McCloskey also teaches a second tag sequence, the “batch-stamp” (highlighted in red), which is unique to each experiment and is intended to associate a polynucleotide to “a specific sample run on a specific date.” *Id.* at 764. The other regions on the encoding oligonucleotide include: “a genomic target-

binding site” (highlighted in grey), “a primer-binding site for PCR” (highlighted in light blue), and “extra bases that serve as internal sentinels of recoding as well as tethering the oligonucleotide to itself” to minimize non-specific annealing (highlighted in purple, below). *Id.* at 763.



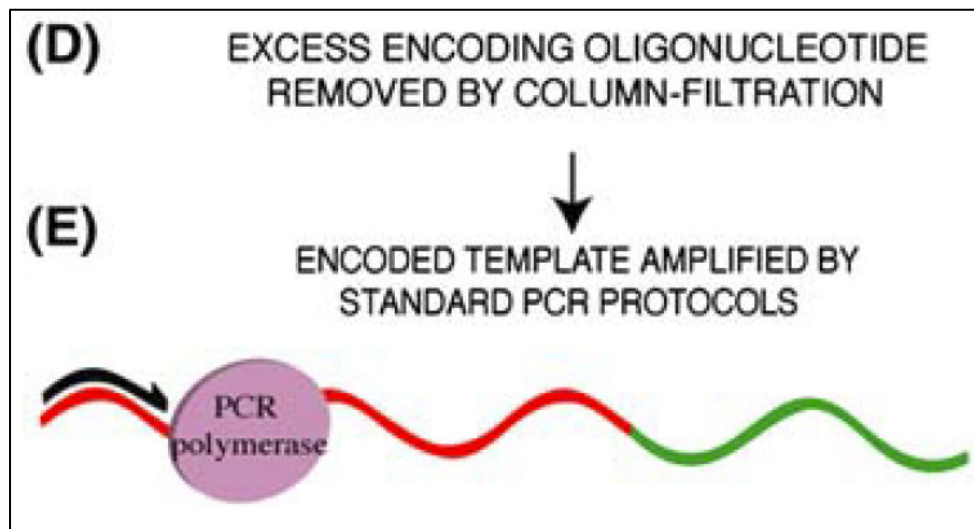
Id. at FIG. 1.

99. McCloskey’s method begins with annealing the 3’ end genomic target binding site (highlighted in grey, above) of the encoding oligonucleotide to denatured genomic DNA (see step (B) in figure, below). *Id.* at 763-766. The oligonucleotide-annealed DNA is then extended to build a complementary DNA sequence (see step (C) in figure, below):



Id. at FIG. 1.

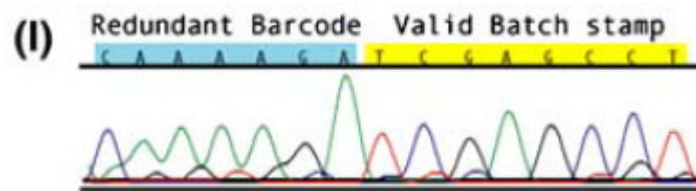
100. The resulting sample of encoded DNA strands are then amplified using PCR and sequenced for evaluation (see step(E) in figure, below). *Id.* at 765-766.



Id. at FIG. 1.

101. McCloskey's method addresses the element of uncertainty in sequence information caused by the risks of amplification bias, contamination, and redundant sequences. *Id.* at 762. McCloskey teaches that a benefit of its method is that it distinguishes redundant sequences found in the end product of PCR amplification

from redundant sequences existing in the sample before amplification. Researchers can readily identify the redundant sequences from the PCR reaction “because they are identical in both barcode and batch-stamp” as opposed to any encoded nucleotides with different barcodes, which suggests that the redundancy existed in the sample prior to amplification (see step (I) in figure, below). *See id.* at 766. Thus, by “counting sequences with non-redundant barcodes”, McCloskey’s method is useful in obtaining “a quantitative measure of template diversity.” *Id.*



Id. at FIG. 1. The McCloskey method based on the use of the random barcodes is thus little different from the claimed method in the '510 patent.

C. Brenner (Ex. 1014)

1. Brenner Is § 102 Prior Art

102. U.S. Patent No. 7,537,897 (“Brenner”) entitled “Molecular Counting” was published on May 26, 2009. EX1014. I understand that Brenner is thus prior art to the '510 patent under pre-AIA 35 U.S.C. § 102(a) and (b) because Brenner was published at least one year before the earliest claimed priority date of the '510 patent.

2. Brenner Overview

103. Brenner is the second of four references that I rely upon in this declaration that teaches the general approach of tagging, amplification, and sequencing for molecular counting.

104. Brenner teaches methods for counting molecules in a sample using unique oligonucleotide tags. *Id.* at Abstract. Brenner's purported approach is to convert the problem of counting molecules into the problem of counting tags. *Id.*; *see also id.* at 1:64-66 ("The invention provides a method of counting molecules in a sample by converting the problem of counting molecules into one of counting sequences of oligonucleotide tags."). Brenner further discloses various methods of counting these tags, including directly sequencing a sample of tags. *Id.* at Abstract; 2:16-20 ("In one aspect, the different oligonucleotide tags of the selected probes, and hence, the number of target molecules, is determined by sequencing a sample of the oligonucleotide tags amplified from the selected probes.").

105. Brenner's method purports to address the need for cost-effective, convenient, and accurate digital measurements of quantities of biomolecules. *Id.* at 52-55. Brenner discloses how this method has applications "in medical and research fields for determining a wide variety of quantities, including genetic copy number variation." *Id.* at 1:56-58; *see also id.* at 4:4-9 ("The present invention provides compositions and methods for making digital measurements of biomolecules, and has applications in the measurement of genetic copy number variation...").

Brenner's method employs conventional techniques of organic chemistry, polymer technology, molecular biology, cell biology, biochemistry, and immunology, known to those skilled in the art at the time of the invention. *Id.* at 4:58-63. These conventional techniques included polymer array synthesis, hybridization, ligation, and detection of hybridization using a label. *Id.* at 4:63-66.

106. Brenner teaches a method for estimating the number of target polynucleotides in a mixture. *Id.* at 3:15-24. First, target polynucleotides are labeled with unique oligonucleotide tags. *Id.* The target polynucleotides "are each associated or linked, e.g. by ligation, to an oligonucleotide tag from a population that has a much larger size than that of the target polynucleotide population." *Id.* at 17:61-64. "Generally, a size of tag population is selected that ensures with high probability that substantially every target polynucleotide will have a unique tag." *Id.* at 18:2-4.

107. To these ends, Brenner further discloses methods for constructing sets of tags and attaching tags to polynucleotides. *Id.* at 12:55-56 ("A general procedure for attaching oligonucleotide tags to polynucleotides is illustrated in FIGS. 2A-2B."); *id.* at 24:51-25:67 ("Nucleotide sequences of sets of such oligonucleotide tags are readily generated by conventional computer programs that generate all possible sequences of a preselected length followed by sorting all prospective tag sequences that fulfill the desire criterion on homopolymer size."); *id.* at 28:35-30:5 (Example

I: Construction of a Metric Tag Set); 30:6-31:40 (Example II: Construction of a Binary Tag Set).

108. After the tagged polynucleotides are generated, the tags are amplified using PCR. *Id.* at 3:20-21 (“amplifying the oligonucleotide tags of the labeled target polynucleotides”); *id.* at 17:52-55 (“In this latter approach, oligonucleotide tags may be amplified using emulsion PCR, e.g. as also disclosed in the cited references, so that clonal populations of each oligonucleotide tag of a sample are formed on beads, which are then sequenced.”); *id.* at 23:36-38 (“Once such an emulsion is formed, conditions are selected for implementing an amplification reaction, such as PCR....”).

109. Finally, the number of different oligonucleotide tags in the sample of amplified oligonucleotide tags are counted to determine the number of target polynucleotides in the starting sample. *Id.* at 2:2-4 (“The number of different oligonucleotide tags detected, or counted, is equal to the number of molecules in the sample.”); *id.* at 3:30-34 (“[T]he number of different oligonucleotide tags in a sample is determined by determining the nucleotide sequences thereof and then counting the number of oligonucleotide tags with different sequences.”); *id.* at 5:17-23 (“[S]ubstantially every molecule to be counted in a sample, e.g. the number of single stranded DNA molecules of a particular genetic locus in a sample of genomic DNA, is associated with a probe having a different tag, so that the process of

counting multiple copies of the same molecule is transformed into a process of counting the number of different kinds of associated tags.”).

110. In particular, Brenner teaches methods of counting by directly sequencing samples of oligonucleotide tags. *Id.* at Abstract (“[D]ifferent tags may be counted by directly sequencing a sample of tags using any one of several different sequencing methodologies.”); *id.* at 22:19-22 (“The oligonucleotide tags are then amplified (5252) out of conjugates (5246), after which a sample of such tags are sequenced and the number of different tags are determined (5254).”); *id.* at 24:30-35 (“The size of the sample of oligonucleotide tags from selected probes that is sequenced is a design choice in that the more oligonucleotide tags that are sequenced the greater the probability that every different oligonucleotide tag (and hence the correct number of target polynucleotides) will be determined.”).

D. Chee (Ex. 1015)

1. Chee Is § 102(e) Prior Art

111. U.S. Patent No. 9,085,798 (“Chee”) entitled “Nucleic Acid Constructs and Methods of Use” was published on July 21, 2015. EX1015. Chee issued from U.S. Patent Application No. 13/266,568, which is the national phase entry of International PCT Application No. US2010/033064 (“the Chee PCT”), filed on April 29, 2010. Chee and the Chee PCT share a specification. The Chee PCT designated the United States and was published in English. EX1016. As such, I understand

from counsel that Chee is prior art to the '510 patent under pre-AIA 35 U.S.C. § 102(e).

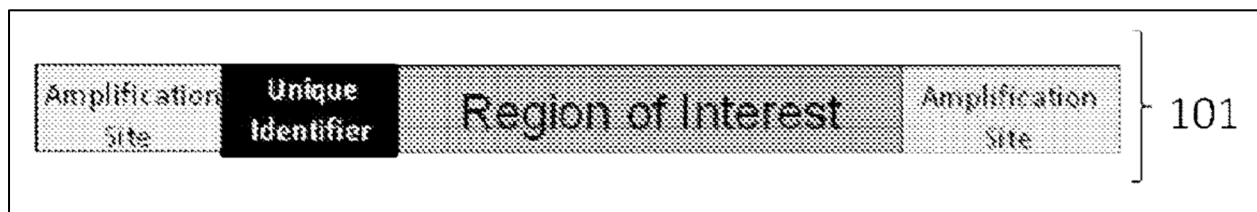
2. Chee Overview

112. Chee is the third of four references that I rely upon in this declaration that teaches the general approach of tagging, amplification, and sequencing for molecular counting.

113. Chee teaches counting uniquely tagged molecules for “identifying potential sample bias introduced by quantitative analysis techniques.” EX1015 at 5:18-20. Chee’s purported approach relies on so-called “oligonucleotide constructs.” *Id.* at Abstract (“The present invention provides oligonucleotide constructs, sets of such oligonucleotide constructs, and methods of using such oligonucleotide constructs . . .”). These nucleic acid constructs comprise a unique identifier (or “tag”), a region of interest (“ROI”), and at least one amplification site. *Id.* at 2:21-33. Chee teaches creating these constructs by tagging nucleic acids derived from a biological sample. *Id.* at 12:17-22 (“[W]hen the constructs are to be used in quantitative analysis of nucleic acids derived from a biological sample, the nucleic acids containing the ROI are isolated, and the additional elements of the constructs are separately created and added to the sample nucleic acids using molecular biology techniques.”); *see also id.* at 4:1-3 (“In another aspect the constructs are created using fragmented nucleic acids, and the unique identifier and

amplification regions are added to the individual fragments”); *id.* at 11:18-20 (“[A] diverse but pre-defined set of sequences can be synthesized and attached to the ROIs, e.g., via PCR primers or by ligation.”).

114. An example of a Chee construct can be seen in the below excerpt of Figure 1.



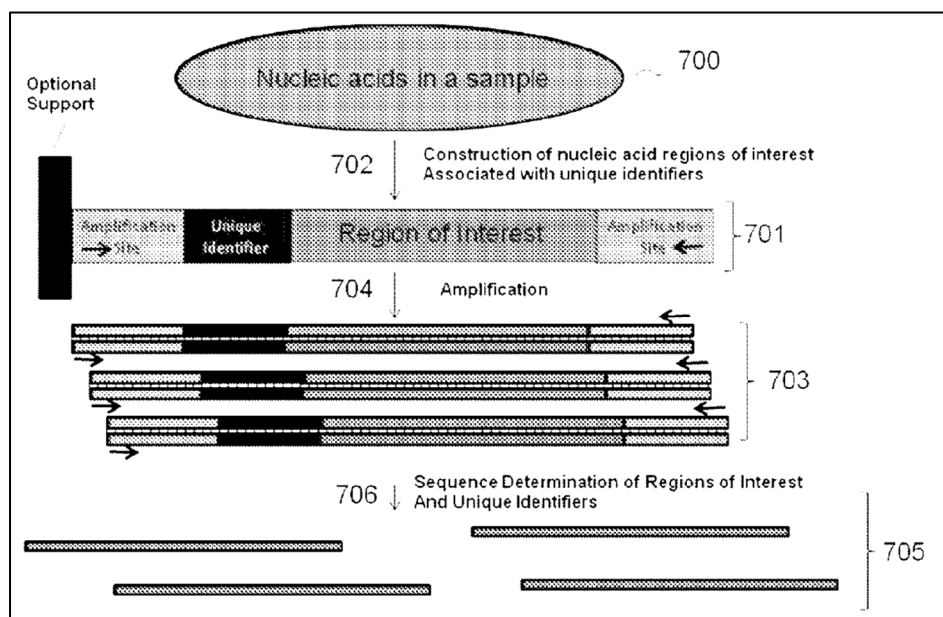
Id. at FIG 1.

115. Chee further discloses amplifying the set of constructs. *Id.* at 13:21-25 (“Following the selection of a set of constructs from a master set, the selected constructs undergo a limited amplification to provide additional material for determining accuracy of the ROI and the associated unique identifier that corresponds to an ROI.”); *id.* at 13:33-37 (“In one example, universal primers that are complementary to the amplification sites in the constructs are used to perform a limited number of amplification cycles, so that each unique molecule now has a small number of identical copies each containing an ROI and the unique identifier.”).

116. Chee further discloses that this amplification can be carried out using conventional PCR methods. *Id.* at 13:40-43 (“Amplification can be carried out in solution, using techniques such as PCR or emulsion PCR (with or without beads),

and/or using solid-phase amplification, e.g., on a support surface.”); *id.* at 14:33-36 (“In one particular aspect, the individual molecules are cloned onto beads and amplified using emulsion PCR to produce many copies of the original template on the surface of a bead.”); *see also id.* at 3:13-17 (“These constructs allow the unique identifier and the ROI to be selectively amplified using single primer-dependent amplification techniques such as asymmetric PCR.”).

117. Following amplification, Chee discloses using sequencing methods to confirm the sequence of the ROI and the unique identifier of a particular construct. *Id.* at 14:3-5 (“[C]omplete constructs are sequenced to confirm both the ROI and the unique identifier of a particular construct.”). The below excerpt from Figure 7 shows tagging, amplification, and sequencing as taught by Chee.



Id. at FIG. 7. Chee teaches that “molecules with identical identifiers arise via amplification of the same original construct.” *Id.* at 14:44-46.

118. Chee further discloses counting the number of constructs with unique identifiers. *Id.* at 9:13-21 (“The ability to uniquely identify specific ROI-containing constructs (as opposed to identifying the same ROI present in multiple constructs) . . . allows the identification of technical bias in, say, amplification procedures used in quantitative analysis, *as a skewed proportion of a specific construct in a sample can indicate a technical bias.*”); *id.* at 10:41-46 (“[T]he unique identifier will be amplified with the ROI of each construct and thus the number of starting molecules can be *inferred by association of each initial ROI present in a sample with its unique identifier.*”). As Chee explains, “[b]y creating constructs comprising the nucleic acids directly isolated from a sample, bias due to subsequent manipulation of the molecules,” e.g., bias in amplification of certain sequences, “can be identified by determining the amounts of constructs with unique identifiers.” *Id.* at 23:56-59. “Thus, if a nucleic acid is either over-represented or under-represented due to amplification bias, the number of molecules comprising unique identifiers will also be over-represented or under-represented, and the quantification of the molecules can be corrected to reflect this manner of bias.” *Id.* at 23:59-64.

E. Porreca (Ex. 1017)

1. Porreca Is § 102(e) Prior Art

119. U.S. Patent No. 11,840,730 (“Porreca”) entitled “Methods and Compositions for Evaluating Genetic Markers” was published on December 12, 2023. EX1017. I understand that Porreca claims priority to at least International PCT Application No. US2010/001293 (“the Porreca PCT”), filed on April 30, 2010. *Id.* at (63). Porreca and the Porreca PCT share a specification. The Porreca PCT designated the United States and was published in English. EX1018. I understand from counsel that Porreca is therefore prior art to the ’510 patent under pre-AIA 35 U.S.C. § 102(e).

2. Porreca Overview

120. Porreca is the fourth of four references that I rely upon in this declaration that teaches the general approach of tagging, amplification, and sequencing for molecular counting.

121. Porreca teaches counting uniquely tagged nucleic acid molecules for determining genotypes in patient samples, with a focus on reducing bias and increasing the reproducibility of multiplex genetic analysis. EX1017 at Abstract (“Aspects of the invention relates to methods and compositions that are useful to reduce bias and increase the reproducibility of multiplex analysis of genetic loci.”); *id.* at 1:23-24 (“The invention relates to methods and compositions for determining genotypes in patient samples.”).

122. Porreca's purported approach uses nucleic acid tags (referred to as differentiator tags) to tag target nucleic acid molecules. *Id.* at 5:1-5 ("In certain embodiments, the disclosure provides preparative methods for capturing target nucleic acids (e.g., genetic loci) that involve the use of differentiator tag sequences to uniquely tag individual nucleic acid molecules."). These differentiator tags allow for "the detection of bias based on the occurrence of combinations of differentiator tag and target sequences observed in a sequencing reactions." *Id.* at 5:5-8.

123. Porreca's purported method is useful for "detecting nucleic acid deletions or insertions in regions containing nucleic acid sequence repeats." *Id.* at 55:49-51; *see also id.* at 20:21-25 ("Methods of the invention may be used for any type of mutation, for example a single base change (e.g., insertion, deletion, transversion or transition, etc.), a multiple base insertion, deletion, duplication, inversion, and/or any other change or combination thereof.").

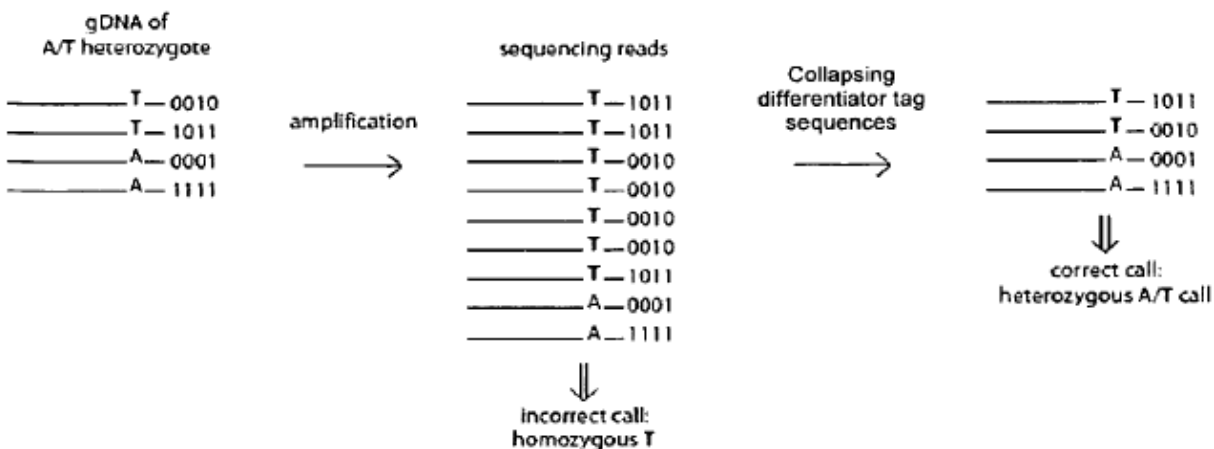
124. Porreca discloses tagging target nucleic acid sequences of interest with differentiator tags. *Id.* at 29:4-6 ("FIG. 4A-C depict various non-limiting examples of methods for combining differentiator tag sequence and target sequences."); *id.* at 29:28-33 ("During a library preparation process, each nucleic acid molecule may be tagged with a unique differentiator tag sequence in a configuration that permits the differentiator tag sequence to be sequenced along with the target nucleic acid sequence of interest (the nucleic acid sequence for which the library is being

prepared, e.g., a polymorphic sequence.”); *id.* at 31:10-13 (“The actual incorporation of the random differentiator tag sequences can be performed through various methods known in the art. For example, nucleic acids comprising differentiator tag sequences may be incorporated by ligation.”).

125. Porreca further discloses the use of amplification and sequencing techniques on the tagged molecules. *See, e.g., id.* at 72:11-12 (“Amplify Circular Material by PCR Using Primers Specific to the ‘Constant Region’ of the Probes”); *id.* at 72:45-47 (“Perform Sequencing of Library According to Manufacturer's Directions (e.g. Illumina, ABI, Etc), Reading Both the Target Sequence and the Differentiator Tag Sequence.”); *see generally* 71:24-73:5 (Example 2).

126. Porreca further discloses counting differentiator tags by sequencing. *Id.* at 60:65-67 (“The number of each differentiator tag sequence is counted, for example, by next-generation sequencing.”); *id.* at 72:50-51 (“Construct a table of target: differentiator tag abundances from the read data....”). The differentiator tags are counted “to distinguish independent target molecules from each other....” *Id.* at 28:29-30. “[I]f each of the plurality of target molecule sequences obtained in the assay is associated with the same differentiator tag, then they probably all originated from the same target molecule due to over-representation (e.g., due to biased amplification) of this target molecule in the assay.” *Id.* at 28:37-41.

127. Figure 5, reproduced below, summarizes the foregoing method of counting by the using of tagging, amplification, and sequencing:



Id. at FIG. 5; *see also id.* at 72:63-73:3 (“All ‘count’ entries should be ‘1’, since any particular target:differentiator tag mapping will not occur more than once by chance, and therefore will only be observed if bias was present somewhere in the sample preparation process. For any target:differentiator tag combination observed more than once, all such reads are ‘collapsed’ into a single read before consensus basecalls are determined.”).

XII. PRIOR ART RENDERS OBVIOUS THE CLAIMS OF THE ’510 PATENT.

128. I now discuss the specifics of how the prior art renders obvious the claims of the ’510 patent.

129. The basic structure of my theory is as follows. Lo teaches the same general approach set forth in claim 1 of the ’510 patent, wherein one determines copy number information by sequencing genomic segments, mapping the sequence

reads to a reference genome, and then finally determining copy number information by counting sequence reads for the genomic region(s) of interest.

130. To the extent any aspect of the independent claims is missing in Lo, it is the process of tagging the DNA sequences to generate unique tagged nucleic acid molecules followed by amplification of these tagged nucleic acid molecules. Yet, as I explained above and further below, this process was well-known to the skilled artisan and was a basic technique in their toolkit. Skilled artisans would have been motivated to use this basic tagging and amplification technique to improve the accuracy of molecular counting, which is what Lo's method is based upon. Indeed, using tagging and amplification with Lo would have been the simple application of a known method to improve an existing process, and a POSA would have had a reasonable expectation of successfully doing so given the wide and routine use of such methods. Four prior art references expressly instruct to apply the tagging and amplification method to determine copy number information by molecular counting. As such, claim 1 of the '510 patent is obvious in view of Lo and other prior art references.

131. As to the dependent claims in the '510 patent, they recite features that are already part of Lo's method or that are otherwise similarly well-known obvious techniques that a skilled artisan would have been motivated to use with a reasonable expectation of success.

A. Ground 1: Claims 1, 2, 9-11, 22, 26, 28, and 33-34 Are Unpatentable as Obvious Over Lo in View of The Knowledge of a POSA

1. Claim 1:

- a) [1A] A method comprising: a) obtaining segments of genomic nucleic acids from a sample containing the genomic nucleic acids;**

132. It is my opinion that Lo discloses a method comprising the step of “obtaining segments of genomic nucleic acids from a sample containing the genomic nucleic acids.”

133. First, Lo discloses receiving a biological sample, such as from a pregnant female, and states the “biological sample may be plasma, urine, serum, or any other suitable sample. The sample contains nucleic acid molecules from the fetus and the pregnant female.” EX1010 at ¶ 54. Lo further clarifies that “the nucleic acid molecules may be fragments from chromosomes,” i.e., segments of genomic nucleic acids. *Id.*

134. Further, Lo states that “[t]he term ‘biological sample’ as used herein refers to any sample that is taken from a subject (e.g., a human, such as a pregnant woman) and contains one or more nucleic acid molecule(s) of interest.” *Id.* ¶ 33. Lo additionally teaches obtaining maternal plasma samples that include DNA that is known to be naturally fragmented. For example, “[i]n one embodiment, random sequencing is performed on DNA fragments that are present in the plasma of a

pregnant woman, and one obtains genomic sequences which would originally have come from either the fetus or the mother.” *Id.* ¶ 83; *see also id.* ¶ 55 (“In step 120, at least a portion of a plurality of the nucleic acid molecules contained in the biological sample are sequenced....In one embodiment, the nucleic acid molecules are fragments of respective chromosomes. One end...both ends, or the entire fragment may be sequenced.”).

135. Therefore, in my opinion Lo teaches that the first step in the method is of obtaining a sample containing segments of genomic nucleic acids

- b) **[1B] randomly tagging the segments of genomic nucleic acids with different nucleic acid tags, thereby generating unique tagged nucleic acid molecules from the segments of genomic nucleic acids, such that each of the unique tagged nucleic acid molecules comprises a segment of the segments of the genomic nucleic acids from step (a) and a nucleic acid tag of the nucleic acid tags;**

136. It is my opinion that it would have been obvious to perform this claim step in combination with the method set forth in Lo. This would have been an obvious combination in view of the knowledge of the skilled artisan, as reflected in McCloskey, Porreca, Chee, or Brenner.

137. As of the earliest priority date of the '510 patent, it was well-known and routine in the art to attach unique nucleic acid tag sequences to genomic DNA fragments prior to amplification to obtain accurate genomic copy number

information. Indeed, the specification of the '510 patent admits as much in its discussion of Miner, McCloskey, and Brenner:

Parameswaran et al. (2007) and U.S. Pat. No. 7,622,281 describe methods of labeling nucleic acid molecules with barcodes for the purpose of identifying the source of the nucleic acid molecules, thereby allowing for high-throughput sequencing of multiple samples (4,5). Eid et al. (2009) describe a single molecule sequencing method wherein single-molecule real time sequencing data is obtained from a DNA polymerase performing uninterrupted template-directed synthesis using four distinguishable fluorescently labeled dNTPs (6). However, these methods do not provide genomic information unaffected by amplification distortion. Miner et al. (2004) describe a method of molecular barcoding to label template DNA prior to PCR amplification, and report that the method allows for the identification of contaminant and redundant sequences by counting only distinctly tagged sequences (22). U.S. Pat. No. 7,537,897 describes methods for molecular counting by labeling molecules of an input sample with unique oligonucleotide tags and subsequently amplifying and counting the number of different tags (23). Miner et al. and U.S. Pat. No. 7,537,897 both describe labeling of input nucleic acid molecules by ligation, which has been found to be an inefficient reaction. McCloskey et al. (2007) describe a method of molecular encoding which does not use ligation but instead uses template specific primers to barcode template DNA molecules prior to PCR amplification (24). However, such a method requires that template specific primers be made for each species of template DNA molecule studied.

EX1001 at 2:1-28.

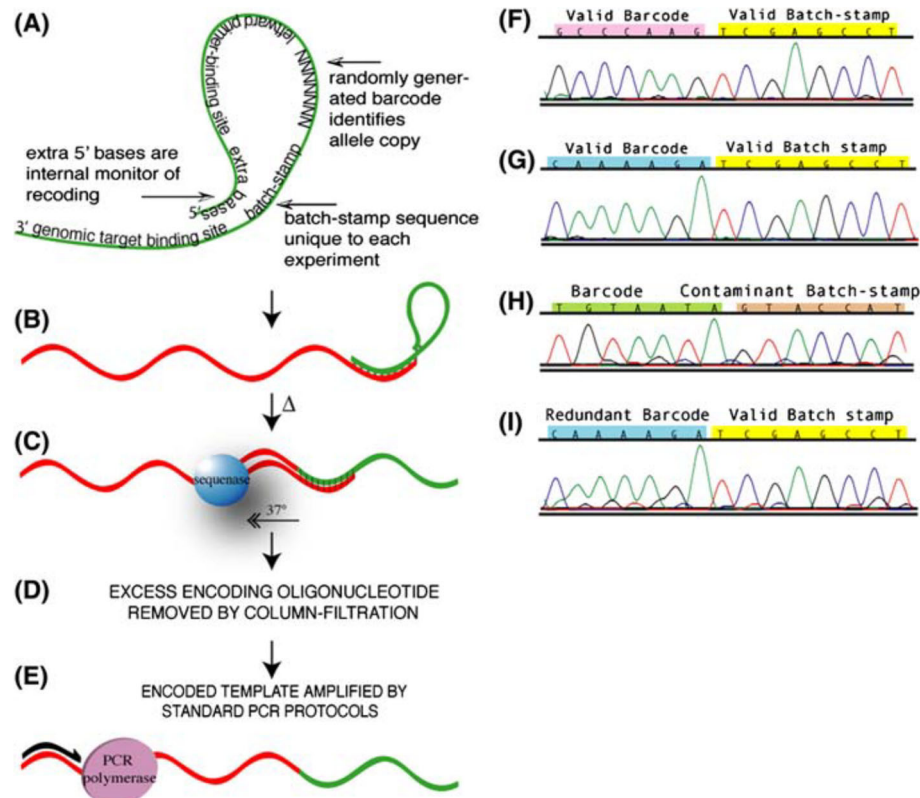
138. Below, I discuss some of these references and other references that describe the use of tagging for molecular counting in conjunction with amplification. Specifically, I discuss the Brenner, Chee, McCloskey, and Porreca references, which I summarized above and which reflect the general knowledge of the skilled artisan with respect to the use of tagging for molecular counting.

139. As I explain below, each of these references discloses “randomly tagging the segments of genomic nucleic acids with different nucleic acid tags, thereby generating unique tagged nucleic acid molecules from the segments of genomic nucleic acids, such that each of the unique tagged nucleic acid molecules comprises a segment of the segments of the genomic nucleic acids from step (a) and a nucleic acid tag of the nucleic acid tags.” I discuss this above in my overview of these references, *see supra* §§ XI.B.2, XI.C.2, XI.D.2, XI.E.2, and I rely on that disclosure for this claim element. Below, I discuss further the relevant disclosures in these references.

140. Consider first McCloskey, which teaches using an “encoding oligonucleotide with five distinct informational regions (FIG. 1): (1) the batch-stamp, unique to each experiment, specifies the DNA source such as the patient or sample identification, and the date of reaction; (2) *the random barcode distinguishes among sequences arising from different cell or allele copies*; (3) the

primer-binding site facilitates specific PCR amplification; (4) the genomic target-binding site targets the locus to be encoded prior to amplification; and (5) the extra 5' bases serve as an internal sentinel of recording.” EX1008 at 762. The “random barcode” listed as component (2) above is the same type of tag in the ’510 patent that distinguishes among molecules and allows one to count molecules accurately after amplification. This is what McCloskey is referring to when it mentions using the barcodes to “detect template redundancy” and being able to measure “template diversity.” *Id.* at Abstract. Each encoding oligonucleotide is attached to a genomic fragment before PCR, “enabl[ing] classification of each PCR-derived sequence as valid, contaminant, or redundant.” *Id.* at Abstract. Allowing one to distinguish sequences from different “cell or allele copies,” the introduction of the “random barcode” created the claimed “unique tagged nucleic acid molecules.” Importantly, because the barcodes are “randomly generated,” and/or because individual barcodes are added to the target nucleic acids through a stochastic process consisting of a solution-phase reaction wherein the collection of tags and target molecules are mixed together in solution such that particular tag molecule that hybridizes to any given nucleic acid molecule will be randomly selected based on which tag happens to interact with a particular target, McCloskey discloses “randomly tagging” the target nucleic acid molecules.

141. FIG. 1 of McCloskey (replicated below) shows the five distinct regions of the encoding nucleotide, including the “randomly generated barcode [that] identifies allele copy.” *Id.* at 763. The claimed “unique tagged nucleic acid molecules” are shown beginning in step (c) below:



142. As another example, consider Porreca, which discloses a concept little different from that of the barcode in McCloskey wherein one uses “differentiator tag sequences to *uniquely tag individual nucleic acid molecules.*” EX1017 at 5:3-5.

143. Porreca teaches using, among other things, “molecular inversion probe” (MIP) technology to reduce bias associated with genomic capture and sequencing, especially during detection of a “plurality of genetic loci in parallel.” *Id.* at 4:15,

3:4. Porreca’s probes are synthesized to contain “differentiator tag sequences” to “uniquely tag individual nucleic acid molecules.” i.e., to “permit the detection of bias based on the occurrence of combinations of differentiator tag and target sequences observed in a sequencing reaction.” *Id.* at 5:3-8, 23:63-67. In this technique, the “number of each differentiator tag sequence is counted, for example, by next-generation sequencing.” *Id.* at 60:65-67; *see also id.* at 61:6-10 (describing a “sequencing technique...used to sequence one or more captured targets, and the sequences are used to count the number of barcodes that are present.”). Porreca further explains that during library preparation, “each nucleic acid molecule may be tagged with a unique differentiator-tag sequence in a configuration that permits the differentiator-tag sequence to be sequenced along with the target nucleic-acid sequence of interest.” *Id.* at 29:28-32. Regardless, given Porreca’s teaching of using “differentiator tag sequences to ***uniquely tag individual nucleic acid molecules***,” *Id.* at 5:3-5, Porreca teaches the claimed “unique tagged nucleic acid molecules.” Further, like McCloskey, Porreca contemplates that the tags may be random sequences and/or that they can be introduced by stochastic processes such as ligation, which is a solution-phase reaction wherein the collection of tags and target molecules are mixed together in solution such that the particular tag that will react with and be added to individual nucleic acid molecules will be a random selection based on which tag happens to interact with a particular target. *See id.* at 29:7-11,

29:37-61, 30:31-40, 31:10-36. As such, Porreca teaches that the nucleic acid molecules are “randomly tagged.”

144. Chee also teaches “the use of nucleic acid sequences as the unique identifiers” to tag and study “one or more desired sequences” of interest. EX1015 at 11:12-14, 2:9. In general Chee teaches the concept of creating “regions of interest *coupled with unique identifiers.*” *Id.* at 23:2-3.

145. Chee uses the unique tagged sequences generated to identify bias in the final count of sequences or Regions of Interest (ROIs). In Chee, the “methods of the invention can identify technical bias in, e.g., amplification, as the unique identifier will be amplified with the ROI of each construct and thus the number of starting molecules can be inferred by association of each initial ROI present in a sample with its unique identifier.” *Id.* at 10:42-46. Regardless, in teaching “regions of interest coupled with unique identifiers,” Chee also teaches the claimed “unique tagged nucleic acid molecules.” Further, Chee also contemplates that the tags may be random sequences and/or that they can be introduced by stochastic processes such as ligation or primer extension, which are solution-phase reactions wherein the collection of tags and target molecules are mixed together in solution so that the particular tag molecule that will react with or hybridize to any given individual nucleic acid molecule will be a random selection based on which tag happens to

interact with a particular target. *See id.* at 3:58-65, 11:12-22. As such, Chee teaches that the nucleic acid molecules are “randomly tagged.”

146. Finally, Brenner explicitly teaches an approach that provides “methods and compositions for *counting molecules in a sample, wherein each molecule is labeled with a unique oligonucleotide tag.*” EX1014 at Abstract. These tags are then used for molecular counting. *See, e.g., id.* at 5:16-23 (stating that the “invention provides a method of counting molecules that are uniquely labeled with tags...substantially every molecule to be counted in a sample, e.g. the number of single stranded DNA molecules of a particular genetic locus in a sample of genomic DNA, is associated with a probe having a different tag, so that the process of counting multiple copies of the same molecule is transformed into a process of counting the number of different kinds of associated tags”); *id.* at 9:23-24 (disclosing a process wherein “binary tags are counted in each band to obtain an estimate of fragment copy number in the genomic DNA.”). Nevertheless, in teaching molecules that are “each labeled with a unique oligonucleotide tag,” Brenner again teaches the claimed “unique tagged nucleic acid molecules.” Further, Brenner also contemplates that the tags may be random sequences and/or that they can be introduced by stochastic processes such as ligation or primer extension, which are solution-phase reactions wherein the collection of tags and target molecules are mixed together in solution so that the particular tag molecule that will react with or

hybridize to any given individual nucleic acid molecule will be a random selection based on which tag happens to interact with a particular target. *See id.* at 2:4-9, 38:25-29, 6:52-55, 7:16-19, 9:31-33, 17:56-64, 18:2-5; FIG. 1G. As such, Brenner teaches that the nucleic acid molecules are “randomly tagged.”

147. Thus McCloskey, Porreca, Chee, and Brenner each disclose “randomly tagging the segments with nucleic acid tags, thereby generating a plurality of unique tagged nucleic acid molecules, such that each of the unique tagged nucleic acid molecules comprises a segment of the segments of the genomic nucleic acids from step (a) and a nucleic acid tag of the nucleic acid tags.” As reflected in each of these four references, this was a standard technique that was part of the skill artisan’s toolkit.

148. A POSA would be motivated to combine the counting approach in Lo with the generally known method of tagging (as reflected in McCloskey, Brenner, Chee and Porreca). There are several reasons for this.

149. First, the method of tagging prior to amplification and counting as described above in §§ X.B-D and ¶¶ 138-147 was well-known and was part of the standard tool kit of a POSA, since it offered a known solution to a known problem: the need to improve the accuracy of counting molecules, particularly in situations where there were limited sample DNA molecules and/or it was important to be more accurate in counting. A POSA would have been motivated to use the tagging set

forth in Chee, McCloskey, Brenner, and Porreca for the simple reason that this technique represented one of only a small number of options for improving the accuracy of counting and was well within the technical grasp of a skilled artisan. Indeed, I am not aware of any other method that was utilized for improving the accuracy of counting DNA molecules in the relevant time frame.

150. Second, adding the tagging feature to Lo would be of particularly high interest to a POSA because Lo is an approach ***based on*** counting, and a POSA would desire to count as accurately as possible. Using the tagging with Lo would enable one to subsequently use amplification without worry about amplification bias (which could be addressed by using the tags). The ability to use amplification would allow one to better capture the full range of molecules from the chromosome or chromosome region of interest, which would further improve the overall counting accuracy. Simply put, a POSA, recognizing that Lo is based on molecular counting, would naturally think to include tagging so as to perform Lo's counting method in as accurate of a manner as possible.

151. Third, and consistent with the foregoing, multiple of the references I rely upon expressly contemplate using the tagging method specifically to determine copy number information. For example, Brenner discusses using tagging and subsequent amplification to for the "measurement of genetic copy number variation." EX1014 at 4:6-7. Porreca discloses a method for "detecting nucleic acid

deletions or insertions in regions containing nucleic acid sequence repeats” for cancer detection “which has been associated with micro satellite instability (MSI) involving an increase or decrease in the genomic copy number of nucleic acid repeats at one or more microsatellite loci.” EX1017 at 55:48-50-56:54-58. Yet additional prior art references expressly contemplate using tagging and amplification with counting for determining copy number information. *See* EX1021 at Abstract, 1:55-2:6, 21:62-65, 5:56-6:46; *see also id.* at 18:30-49 (explaining that the disclosed method can be used to “increase the confidence in calling copy number variations”); *id.* at 21:62-65; EX1022¹ ¶ 2 (“determining the frequency of copies of gene segments for karyotyping is an important diagnostic option for diseases such as cancer); *id.* ¶ 13 (“The task is solved by labeling all nucleic acids of a nucleic acid mixture, regardless of whether they are double-stranded or single-stranded, with a ‘random tag’ (RAT) derived from a mixture of synthetic oligonucleotides, wherein the mixture contains sufficient RATs with different sequences to ensure that the resulting random tag nucleic acid (RAT-NS) combinations are practically unique.”). Thus, at least four prior art references provide express teaching to include tagging with an approach for determining copy number information based on molecular counting, as is taught in Lo.

¹ Certified translation of German original EX1040. *See* EX1041.

152. Fourth, adding the tagging features described in McCloskey, Brenner, Porreca, and Chee would help solve the signal-to-noise challenge described in Lo. Specifically, Lo explains that “fetal DNA is present in low absolute concentration” in maternal plasma, creating a low signal problem. *See* EX1010 ¶ 13. Additionally, fetal DNA “represent[s] a minor portion of all DNA sequences in maternal plasma and serum,” meaning it is diluted within a large background of maternal DNA, which contributes to high noise in the sample. *See id.* A POSA would be motivated to combine the tagging techniques known in the art with the amplification techniques described in step [1C] to increase the signal-to-noise ration and thereby improve the sensitivity and accuracy of detecting aneuploidy described in Lo. Lo emphasized that it was “desirable to have methods that allow the noninvasive detection of fetal chromosomal aneuploidy by maximizing the amount of genetic information that could be inferred from the limited amount of fetal nucleic acids which exist as a minor population in a biological sample containing maternal background nucleic acids.” *Id.*

153. A POSA would have had a reasonable expectation of successfully making this combination because the machinery for preparing and attaching the tags was well-known and straightforward.

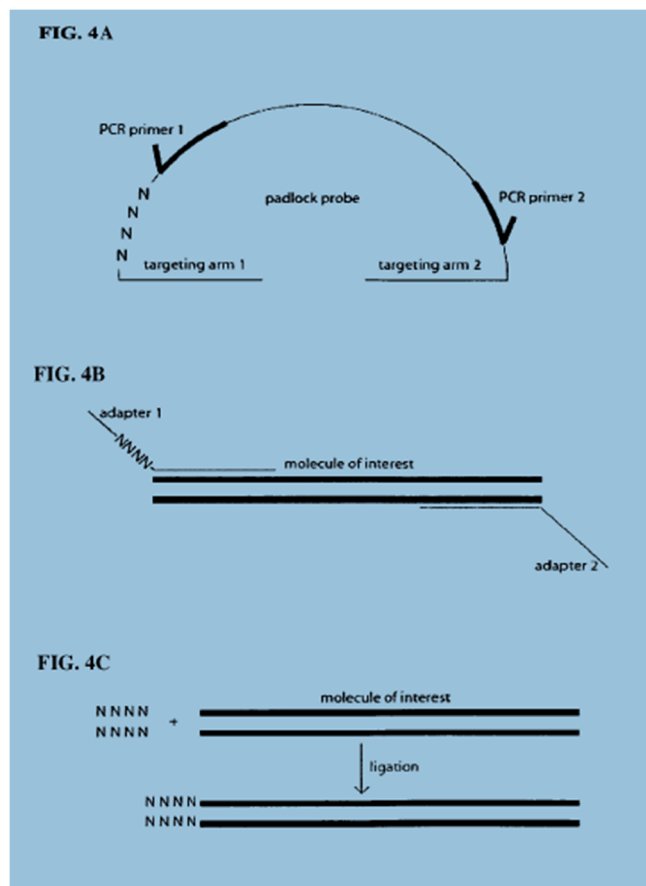
154. A POSA would know, for instance, that it is straightforward to design a large number of random sequences to achieve the claimed uniqueness. Brenner,

for instance, describes using oligonucleotide tagging sets which are “preferably at least ten times the number of target polynucleotides to be counted,” and “where labeling by sampling is employed, preferably, the size of sets of oligonucleotide tags of the invention are at least 100, and in another aspect, at least 1000; and in still another aspect at least 10,000.” EX1014 at 25:36-46. Brenner similarly contemplates that the tags should be “maximally different,” *id.* at 21:14-18, and teaches that there were well known ways of making the tags “maximally different.” Brenner explains, for instance, that there were “distance measures” known in the art for designing such sequences. Brenner cites several literature references to this effect, stating “exemplary distance measures for sequences applicable to the invention are well known and are widely disclosed in the literature, as exemplified by the following references: Gusfield, Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology (Cambridge University Press, 1997); Navarro et al, Flexible Pattern Matching on Strings (Cambridge University Press, 2002); Sankoff et al, editors, Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison (Center for the Study of Languages, 1999); and the like.” *Id.* at 25:58-66.

155. Chee likewise uses a “large number of unique sequences that can be generated from the random assembly of nucleotides [which] provides a high probability that each individual construct will be uniquely identified. The number

of nucleotides in the identifier is preferably chosen so that the number of potential and actual sequences represented is much greater than the total number of oligonucleotide constructs in the set.” EX1015 at 11:26-38. And Porreca similarly discloses a method of creating a “large library of unique differentiator tag sequences...using degenerate, random-sequence polynucleotides of defined length.” EX1017 at 29:37-39. Thus, methods were available for preparing a large number of different oligonucleotide tags.

156. So too were methods for attaching the tags. For example, Porreca describes various methods for combining tags with target sequences. Consider, for instance, FIG. 4A-C of Porreca below:.



EX1017 at FIG. 4A-4C.

157. Brenner likewise details a general procedure for attaching an oligonucleotide tag to one end of a polynucleotide (see FIGS. 2A-2B below), including methods for attachment chemically or enzymatically. Consider for instance Brenner, Examples I and II, which describe the construction of metric and binary tag sets, respectively:

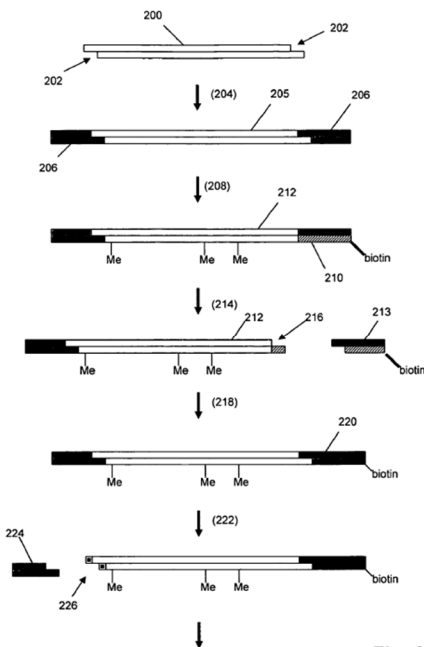


Fig. 2A

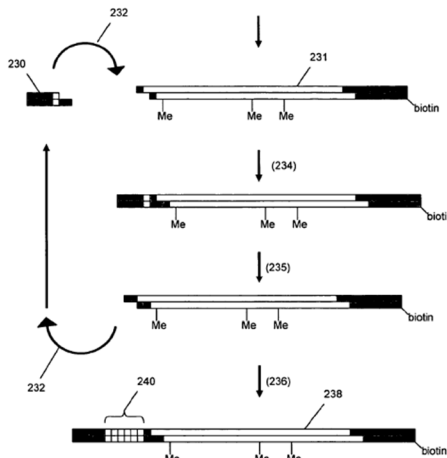


Fig. 2B

EX1014 at FIG. 2A-2B.

158. In any event, there can be no question of reasonable expectation of success here because, as I explain above, the '510 patent admits that various methods of tagging existed in the art. *See* EX1001 at 2:20-22 (describing labeling by ligation); *id.* at 2:23-26 (describing labeling through specific primers); *id.* at 21:38-40 (“Tags may be associated with nucleic acid molecules by using, singly or in combination, ligation, nick-ligation, terminal transferase, hybridization, priming, and the like.”).

159. In sum, it is my opinion that a POSA would be motivated to combine the well-known tagging of McCloskey, Brenner, Chee, and Porreca, with the counting method of Lo to increase accuracy in the final genomic fragment counts.

Likewise, it is my opinion that a POSA would have a reasonable expectation of success, based on well-known methods of bar-coding as practiced in the cited references herein.

- c) **[1C] subjecting the unique tagged nucleic acid molecules to a polymerase chain reaction (PCR), thereby generating copies of the unique tagged nucleic acid molecules;**

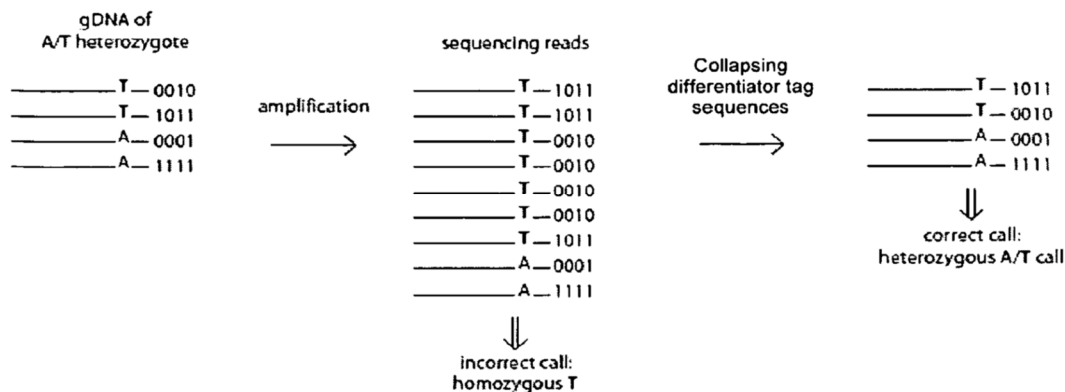
160. As discussed immediately above in connection with element [1B], it is my opinion that it would have been obvious to tag the DNA fragments utilized in Lo's method to improve the accuracy of molecular counting. This technique (as disclosed in McCloskey, Brenner, Chee, and Porreca) is a method that would have been within the common knowledge of the skilled artisan. The skilled artisan would have further understood that this approach includes PCR amplification of the tagged fragments. Indeed, this is disclosed in each of McCloskey, Brenner, Chee, and Porreca. I discuss this above in my overview of these references, *see supra* §§ XI.B.2, XI.C.2, XI.D.2, XI.E.2, and I rely upon that disclosure for this claim element. Below, I discuss further relevant disclosure in these references.

161. As discussed above, McCloskey teaches tagging with an encoding nucleotide containing five distinct regions. In a specific example involving the FMR1 locus, for instance, McCloskey tags the relevant DNA molecules after which "standard PCR amplification to the human locus FMR1" was applied. EX1008 at

766. Thus, McCloskey discloses PCR amplification of unique tagged nucleic acid molecules.

162. Porreca likewise teaches the use of PCR methods that “may include amplification of a single target nucleic acid and multiplex amplification (amplification of multiple target nucleic acids in parallel). EX1017 at 16:12-15. FIG. 5 of Porreca (see below) shows a method for amplification of tagged target DNA fragments, followed by counting of the tagged reads to reduce amplification bias.

FIG. 5



Id. at FIG. 5; *see also id.* at 15:64-67 to 16:1-2 (describing that in “amplification-based (e.g., PCR-based or LCR-based, etc.) preparative methods, genomic loci (target nucleic acids) are isolated directly by means of a polymerase chain reaction or ligase chain (or other amplification method) that selectively amplifies each locus

using a pair of oligonucleotide primers.”); *id.* at 58:43-49 (“PCR...is performed to isolate and amplify the circular targets from the starting nucleic acid pool.”); *id.* at 58:50-53 (“repetitive regions are surrounded by non-repetitive unique sequences, which can be used to amplify the repeat-containing regions using for example, PCR.”). Thus, Porreca further discloses PCR amplification of unique tagged nucleic acid molecules.

163. Brenner describes a technique for amplification after sequence tagging wherein, for example, “*oligonucleotide tags may be amplified using emulsion PCR, e.g. as also disclosed in the cited references*, so that clonal populations of each oligonucleotide tag of a sample are formed on beads, *which are then sequenced.*” EX1014 at 17:53-55. Brenner emphasizes in the disclosure of the method of metric tagging to adjust the tagging lengths to streamline the subsequent amplification, stating that although “each genome is in a one-to-one correspondence with a single length of an oligonucleotide (i.e. a metric tag), the physical lengths of the metric tags are not the same and *since it is desirable to be able to PCR the tags* preferably the metric tags should be the same length. Thus, appropriate length of oligonucleotide are added to each to make them all the same.” *Id.* at 17:14-20. Additionally, Example 1 of Brenner describes a detailed technique using PCR to amplify 64 plasmid DNA samples containing tags. *Id.* at 29:27-54.

164. Finally, Chee describes “using techniques such as PCR or emulsion PCR” for amplification in a solution of target sequences. EX1015 at 13:40-41. Chee teaches the reason or motivation for choosing this technique, that “methods of the invention can identify technical bias in, e.g., amplification, as the unique identifier will be amplified with the ROI of each construct and thus the number of starting molecules can be inferred by association of each initial ROI present in a sample with its unique identifier.” *Id.* at 10:41-46. Chee further discloses that to “truly minimize the amount of amplification needed, the labeling and sorting of individual molecules can be employed prior to the amplification step in the present methods.” *Id.* at 15:64-66. Thus, Chee further discloses PCR amplification of unique tagged nucleic acid molecules.

165. A POSA would have been motivated to use amplification with tagging for several reasons. Most important, amplification is part and parcel of the methods set forth in Brenner, McCloskey, Chee, and Porreca, which is all reflective of the knowledge of the skilled artisan. Indeed, a POSA would naturally combine tagging with subsequent amplification because a key advantage of the tagging is to permit one to subsequently perform amplification and hence ensure that one captures all the molecules of interest while maintaining an avenue to avoid amplification bias. Thus, all the reasons set forth above in connection with element [1B] as to why a POSA

would be motivated to apply tagging to Lo's method apply equally to the use of amplification.

166. Further, exemplary samples in Lo's method were bodily fluids from a subject, such as blood or urine samples. EX1010 ¶ 14. As a POSA would understand and as Lo confirms, with such samples there is a limited amount of DNA overall and even less DNA from the fetus that one is interested in testing. *See id.* ¶¶ 13, 16, 18. Therefore, a POSA would have naturally been motivated to perform amplification to increase the overall amount of DNA for testing.

167. Yet additional motivation to utilize amplification is present within Lo itself. Specifically, Lo teaches multiple methods of enriching nucleic acid sequences prior to sequencing and the motivation for doing so, namely for cost-saving and efficiency:

- Lo teaches that in “some embodiments of the present invention, the random sequencing may be preceded by procedures to enrich a biological sample with particular populations of nucleic acid molecules sharing certain common features.” EX1010 ¶ 47.
- Lo explains that it would be “possible and cost-effective to enrich the pool of nucleic acids to be sequenced prior to random sequencing of a fraction of the enriched pool” *Id.* ¶ 77.

- Lo teaches that since the fetal DNA found in the maternal plasma is generally of shorter length than the maternal chromosomal fragments, one method to accomplish enrichment would be to “fractionate the nucleic acid sequences in the sample according to molecule size, e.g. by gel electrophoresis or size exclusion columns or by microfluidics-based approach.” *Id.*
- Lo also teaches the use of hybridization techniques to accomplish enrichment of potentially aneuploid chromosomes and one or more normal chromosomes “for example onto oligonucleotide microarrays,” followed by random sequencing. *Id.* ¶ 79. This is a known cost-saving technique to those skilled in the art.
- Lo describes that pre-sequencing enrichment in fetal cell free DNA could be accomplished through suppression of the maternal DNA relative to fetal DNA with the use of chemical treatments such as formaldehyde. *Id.* ¶ 78.
- Lo teaches the concept of sub-selecting some desired sequences from the pool of nucleic acids prior to sequencing. “For example, hybridization based techniques such as oligonucleotide array could be used to first sub-select for nucleic acid sequences from certain chromosomes, e.g. a potentially aneuploid chromosome and other

chromosome(s) not involved in the aneuploidy tested. Another example is that a certain sub-population of nucleic acid sequences from the sample pool is sub-selected or enriched prior to sequencing.” *Id.* ¶ 72.

168. The enrichment taught by Lo is an improvement over previous whole-genome sequencing methods and is well-known to those of skill in the art, providing enhanced efficiency and cost-savings, as described above. The amplification as taught by McCloskey, Brenner, Porreca or Chee is another form of enrichment, and therefore provides the same benefit but with the additional increase in counting accuracy by virtue of the tagging.

169. In my opinion, a POSA would be motivated to combine the amplification and tagging of McCloskey, Brenner, Chee, and Porreca, with the enrichment method of Lo to increase efficiency and cost savings without the loss of accuracy in the final genomic fragment counts. Indeed, Lo expressly contemplates enrichment, and PCR is a well-established method to provide enrichment.

170. Further, a POSA would have a reasonable expectation of success because there is no question that PCR amplification was routine and widely used well before the priority date. Indeed, each of Brenner, Chee, Porreca and McCloskey disclose the use of PCR. And the '510 patent effectively admits that PCR was a well-known and routine technique. Specifically, the specification explains that

“[g]enerally...the [amplification] steps are performed multiple times (e.g., polymerase chain reaction)” EX1001 at 13:50-54. The Abstract also references “subjecting the tagged nucleic acid molecules to amplification by polymerase chain reaction.” *Id.* at Abstract. In other words, the ’510 patent rightly characterizes PCR as a well-known prior art method that people of skill in the art had no trouble utilizing.

d) [1D] generating tag associated sequence reads by sequencing the copies of the unique tagged nucleic acid molecules of step (c);

171. Lo in combination with the knowledge of the skilled artisan (as reflected, for example, in McCloskey, Brenner, Chee, and Porreca) renders obvious “generating tag associated sequence reads by sequencing the copies of the unique tagged nucleic acid molecules of step (c).” As discussed above in connection with claim elements [1B] and [1C], a POSA would have been motivated to utilize tagging and amplification of sequences in connection with Lo’s method for the detection of copy number information. Having been motivated to use tagging and amplification with a reasonable expectation of success, a skilled artisan would have further proceeded according to Lo’s standard method, which involves sequencing of nucleic acid molecules for the purpose of determining copy number information.

172. Lo describes a process, wherein in “step 120, at least a portion of a plurality of the nucleic acid molecules contained in the biological sample are

sequenced. The portion sequenced represents a fraction of the human genome. In one embodiment, the nucleic acid molecules are fragments of respective chromosomes. One end (e.g. 35 basepairs (bp)), both ends, or the entire fragment may be sequenced. All of the nucleic acid molecules in the sample may be sequenced, or just a subset may be sequenced. This subset may be randomly chosen.” EX1010 ¶ 55.

173. Figure 6 of Lo (reproduced below) shows the results of sequencing tagged DNA fragments (referred to as sequence tags) from maternal plasma samples in Experiment I.

Sample	Total no. of sequenced tags	Total no. of sequenced basepairs	Proportion of the human genome (%)	U0 count	U0 count / total count (%)	No. of U0 sequenced basepairs	U0 Proportion of the human genome (%)
Sample 1 T21	1.12E+07	4.02E+08	13.4	1.93E+06	17.3	6.94E+07	2.31
Sample 2 T21	1.04E+07	3.73E+08	12.4	1.80E+06	17.4	6.50E+07	2.17
Sample 3 T21	8.90E+06	3.20E+08	10.7	2.09E+06	23.4	7.51E+07	2.50
Sample 4 T21	1.02E+07	3.69E+08	12.3	2.23E+06	21.7	8.02E+07	2.67
Sample 5 euploid	1.06E+07	3.83E+08	12.8	2.12E+06	19.9	7.62E+07	2.54
Sample 6 euploid	9.58E+06	3.45E+08	11.5	1.91E+06	20.0	6.88E+07	2.29
Sample 7 euploid	9.55E+06	3.44E+08	11.5	2.01E+06	21.0	7.22E+07	2.41
Sample 8 euploid	9.09E+06	3.27E+08	10.9	2.09E+06	23.0	7.53E+07	2.51

FIG. 6

Id. at FIG. 6.

174. Experiment II explains that “only the uniquely mappable sequenced tags, termed U0 in nomenclature from the Eland software, were used to demonstrate the presence of overrepresentation in the amount of chromosome 21 sequences in the maternal plasma specimens from pregnancies each carrying a fetus with trisomy 21, as described in example I. As shown in FIG. 6, *U0 sequences only represent a subset of all the sequenced tags generated from each specimen and further represent an even smaller proportion*, some 2%, of the human genome.” *Id.* at ¶ 112; *see also id.* ¶¶ 56, 73 (disclosing single molecule sequencing technologies including “the Roche 454 platform, the Applied Biosystems SOLiD platform, the the [sic] Helicos True Single Molecule DNA sequencing technology, the single molecule, real-time (SMRT™) technology of Pacific Biosciences, and nanopore sequencing”).

175. The only difference between this claim step and the method set forth in Lo is that this claim step involves sequencing tagged and amplified nucleic acid molecules whereas Lo directly sequences the original sample molecules. Yet, as set forth above in connection with elements [1B] and [1C], a skilled artisan would have found it obvious to include tagging and amplification with Lo’s method. Having been motivated to modify Lo’s method to use tagging and amplification, a POSA would have necessarily sequenced the tagged and amplified molecules, according to the methods and techniques set for in Lo that I summarize above.

176. In sum, it is my opinion that a POSA would have been motivated to combine the tagging and amplification of McCloskey, Brenner, Chee, and Porreca, as described above in claim elements [1B] and [1C] with the sequencing method of Lo, to increase the accuracy of the counting method, for the purpose of determining copy number information. Likewise, it is my opinion that a POSA would have had a reasonable expectation of success, given how well-known the methods of sequencing were as practiced in the cited references herein. Indeed, as noted above, Lo proposes to use well-known commercial available sequencing platforms, such as Roche 454, SOLiD, Helios, and Pacific Biosciences, similar to the disclosure of the '510 specification.

- e) **[1E] assigning each of the unique tagged nucleic acid molecules to a location on a genome by mapping each of the tag associated sequence reads of step (d) to a location in the genome; and**

177. In my opinion, a POSA would have been motivated to perform this claim step in conjunction with Lo's method for determining copy number information.

178. To begin, Lo teaches the process of assigning nucleic acid sequences to a reference by mapping. Lo teaches, for instance, that "based on the sequencing (e.g. data from the sequencing), a first amount of a first chromosome (e.g. the clinically relevant chromosome) is determined. The first amount is determined from sequences identified as originating from the first chromosome. *For example, a*

bioinformatics procedure may then be used to locate each of these DNA sequences to the human genome.” EX1010 ¶ 58.

179. Further, Lo teaches that “[e]ach genomic DNA or plasma DNA sample was sequenced using one lane of a flow cell. The short sequence tags generated *were aligned to the human reference genome sequence* and the chromosomal origin was noted. The total number of individual sequenced tags aligned to each chromosome were tabulated and compared with the relative size of each chromosome as expected from the reference human genome or non-disease representative specimens.” *Id.* ¶ 70.

180. The Lo reference provides, in Example I, details regarding a study wherein maternal plasma DNA was collected from women with aneuploid and euploid pregnancies. After DNA from each maternal plasma specimen was sequenced, “[s]equenced reads were compiled using Solexa Analysis Pipeline. All reads were then *aligned to the repeat-masked reference human genomic sequence, NCBI 36 assembly.*” *Id.* ¶ 93.

181. Figure 4A of Lo (reproduced below) shows the percentages of maternal plasma DNA aligned to each of the 24 chromosomes (22 autosomes and X and Y chromosomes) for two representative cases, one with a euploid fetus and the other carrying a trisomy-21 fetus:

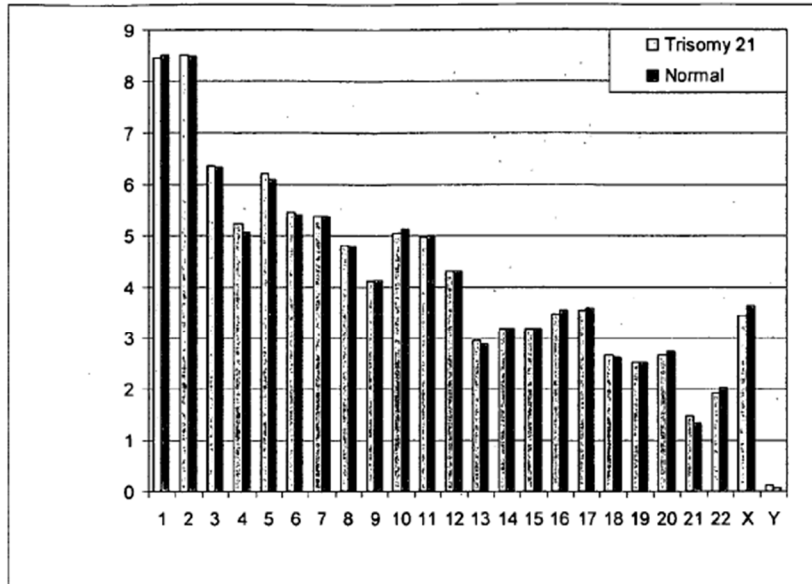


FIG. 4A

Id. at FIG. 4A.

182. The only difference between this '510 claim step and the method set forth in Lo is that this claim step involves mapping tagged and amplified nucleic acid molecules whereas Lo aligns the sequences of the original sample molecules. Yet, as set forth above in connection with elements [1B] and [1C], a skilled artisan would have been found it obvious to include tagging and amplification with Lo's method. Having been motivated to modify Lo's method to use tagging and amplification, a POSA would have necessarily sequenced and then aligned the sequences of the tagged and amplified molecules, according to the methods and techniques set forth in Lo that I summarize above. There can be no question that a POSA would have had a reasonable expectation of successfully doing this given that alignment methods were well-known and routine. Lo even reports actual data

showing that it was easily achievable to align sequence reads and use the aligned sequence reads to determine relative copy number information. *See id.* ¶¶ 95-96, 99-104; *id.* at FIGS. 3A, 4B, 5. The tag sequences would, of course, present no challenge in the alignment because a POSA would know the tag sequences would be easily trimmed (and analyzed separately) in the alignment process as was conventional and routine at the time.

- f) **[1F] in a plurality of locations in the genome, obtaining relative copy number information of the genomic nucleic acids from the sample based on the number of unique tagged nucleic acid molecules that have been assigned to each of the plurality of locations in the genome.**

183. In my opinion, in view of the knowledge of the skilled artisan, it would have been obvious to perform this claim step in conjunction with Lo's method. As I explain above, in connection with steps [1B] and [1C], a POSA would have found it obvious to use tagging and amplification in conjunction with Lo's method, which, as I explain immediately below, is based on counting molecules.

184. To begin, Lo is clear that its method is based on counting molecules to obtain relative copy number information. Lo discloses, for instance, that there "are a number of ways of determining the amounts of the chromosomes, including but not limited to counting the number of sequenced tags, the number of sequenced nucleotides (basepairs) or the accumulated lengths of sequenced nucleotides

(basepairs) originating from particular chromosome(s) or chromosomal regions.”

EX1010 at ¶ 60.

185. Lo then discloses how to utilize the sequence tags that were counted above to derive further types of relative copy number information. Specifically, Lo discloses that “the number of sequenced tags, i.e. absolute count, aligned to a particular chromosome *could be compared to the absolute count of sequenced tags aligned to other chromosomes*. Alternatively, the *fractional count of the amount of sequenced tags from chromosome 21 with reference to all or some other sequenced tags could be compared to that of other non-aneuploid chromosomes*.”

Id. ¶ 75. Lo also notes that because 36 base pairs “were sequenced from each DNA fragment, the number of nucleotides sequenced from a particular chromosome could easily be derived from 36 bp multiplied by the sequenced tag count.” *Id.*

186. Lo further teaches that detection of relative copy number information can be generated by sequencing a “random representative fraction of DNA molecules from the specimen pool” and comparing the “*number of different sequenced tags aligned to various chromosomal regions between specimens containing or not containing the DNA species of interest. Chromosomal aberrations would be revealed by differences in the number (or percentage) of sequences aligned to any given chromosomal region in the specimens*.” *Id.* ¶ 108.

Lo explains that any type of clinically relevant sequence that one may identify by

these means are “copy number variations,” which a POSA would understand encompasses variations in the number of regions of a genome relative to the normal that is expected and seen in other portions of the genome. *Id.* ¶ 37. Lo teaches the mathematical formula for doing this analysis and includes actual experimental data showing that his method works. *See id.* ¶¶ 95-96, 99-104; *id.* at FIGS. 3A, 4B, 5.

187. Thus, there is no question that Lo’s method is based on counting nucleic acid molecules to determine relative copy number information. Moreover, Lo teaches that this counting can be done at a “plurality of locations” in the reference genome. Specifically, Lo teaches that the analysis can be done at a plurality of chromosomes or a plurality of regions in the genome. *See, e.g., id.* ¶ 109 (“Changes (amplifications and deletions) in multiple chromosomal regions may be used. Thus, there would be an increased proportion of sequences aligned to the amplified regions and a decreased proportion of sequences aligned to decreased regions.”); *id.* at claim 5 (“5. The method of claim 1, wherein the first chromosome is chromosome 21, chromosome 18, chromosome 13, chromosome X, or chromosome Y”); *id.* ¶ 14 (“One or more cutoff values are chosen for determining whether a change compared to a reference quantity exists (i.e. an imbalance), for example, with regards to the ratio of amounts of two chromosomal regions (or sets of regions).”); *id.* ¶ 67 (“Also, dosage imbalance of a particular chromosome or chromosomal regions can be quantitatively determined. In other words, the dosage imbalance of the chromosome

or chromosomal regions is inferred from the percentage representation of the said locus among other mappable sequenced tags of the specimen.”); *id.* ¶ 71 (“Gains or losses of chromosomal regions or whole chromosomes were determined by comparing the tag counts with the expected chromosome size in the reference genome or that of a non-disease representative specimen.”); *id.* ¶ 108 (“The number of different sequenced tags aligned to various chromosomal regions is compared between specimens containing or not containing the DNA species of interest. Chromosomal aberrations would be revealed by differences in the number (or percentage) of sequences aligned to any given chromosomal region in the specimens.”).

188. Lo’s mapping to a “chromosomal region” is consistent with PO’s reading of “location” in the parallel district court litigation. *See* EX1020 (interpreting locations as any of “genes,” “genetic locus,” and “start/stop positions”). This count of molecules in a region represents “copy number information.

189. The only difference between this claim step and the method set forth in Lo is that this claim step involves counting “unique tagged nucleic acid molecules,” whereas Lo counts original sample molecules. Yet, as set forth above in connection with elements [1B] and [1C], a skilled artisan would have found it obvious to include tagging and amplification with Lo’s method. Having been motivated to modify Lo’s method to use tagging and amplification, a POSA would have naturally counted the

“unique” molecules by using the barcodes along with the alignment information. Indeed, as I explain above in connection with steps [1B] and [1C], the main reason for using the tagging and amplification in the first place is to increase the accuracy of the counting, which is accomplished by employing the tags to identify and count only the unique molecules, as is done in each of Chee, Brenner, McCloskey, and Porreca. To count anything other than the unique molecules would defeat the purpose of using the tagging and amplification that the skilled artisan would have been motivated to use.

190. In this regard, even in Lo’s method, he proposed to only count the molecules that could be uniquely mapped to a single, specific location. *See, e.g.*, EX1010 ¶ 85 (“For example, in example II, sequenced tags that were uniquely aligned to the repeat-masked reference human genome were used.”); *id.* ¶ 112 (“Furthermore, in this study, only the uniquely mappable sequenced tags, termed U0 in nomenclature from the Eland software, were used....”). In this way, Lo avoided including sequence reads that might distort the count of sequence reads mapped to a particular region. A skilled artisan using tagging and amplification in Lo’s method would have acted consistently with Lo’s interest in having accurate sequence counts and used only the unique sequence reads rather than counting duplicates and hence overinflating read counts.

191. Therefore, in my opinion, a POSA would understand that the Lo disclosure teaches this claimed element.

2. Claim 2: The method of claim 1, wherein the polymerase chain reaction is performed in the presence of region specific PCR primers.

192. As discussed above, Lo in view of the knowledge of the skilled artisan renders obvious claim 1. Claim 2, which recites the “method of claim 1, wherein the polymerase chain reaction is performed in the presence of region specific PCR primers” would have likewise been rendered obvious by specific knowledge of the skilled artisan, as reflected in disclosure in each of McCloskey, Brenner, Chee, and Porreca. Below, I discuss the relevant disclosure in each of these references and explain why this claim would have been obvious.

193. McCloskey performs PCR using region-specific primers—one primer that binds the synthetic tag added by the encoding oligonucleotide and a second primer that binds a sequence derived from the FMR1 genomic locus. EX1008 at 764. To achieve this, McCloskey’s method first tags FMR1 DNA using an encoding oligonucleotide that hybridizes to the locus via a 3’ genomic binding segment. *See id.* at 762, 764. A brief heat step melts the tag duplex, exposing a 5’ primer binding site. *Id.* at 764. Sequenase then extends the 3’ end of the oligonucleotide. *Id.* at 765. PCR is then performed using two primers: one that anneals to the 5’ tail “designed to have no complementarity with competing genomic fragments” and one

that anneals to the newly copied FMR1 sequence. *Id.* at 764. A POSA would therefore understand that McCloskey's amplification is carried out in the presence of region-specific PCR primers, satisfying the limitation of claim 2.

194. Brenner teaches multiple PCR variations, including "RT-PCR, real-time PCR, nested PCR, quantitative PCR, [and] multiplexed PCR." EX1014 at 38:58-60. A POSA would understand that the reference to "multiplexed PCR" is a teaching of using region specific primers. Specifically, multiplexed PCR refers to an approach wherein multiple regions are PCR amplified at the same time, which will involve multiple primers for DNA targets within the region(s) of interest. More generally, a POSA would understand that the methods taught in Brenner may routinely involve region-specific primers, as Lo specifically contemplates. *See* EX1010 ¶ 82 ("[M]olecular diagnostic approaches, such as those based on fluorescence in situ hybridization, quantitative real-time PCR, digital PCR...and so on, where gene loci to be targeted require prior pre-determination, thus requir[e] the use of locus-specific primers.").

195. Chee also teaches the use of "subset-specific primers" to selectively amplify defined groups of nucleic acid molecules. For example, Chee proposes organizing large sets of oligonucleotide sequences into smaller "subsets," which can be selectively amplified by its own pair of subset-specific primers. *See* EX1015 at 16:27-30 (explaining that "rather than isolating and using 10,000 sequences, it may

be preferable to isolate 100 groups of 100 sequences each. This can be accomplished using the aid of 100 pairs of subset-specific primers”). Chee also discloses the use of “ROI-specific primers” to convert single-stranded regions of interest into double-stranded molecules. *Id.* at 17:48-50.

196. Finally, Porreca also discloses region-specific primers in the context of “multiplex amplification (amplification of multiple target nucleic acids in parallel).” EX1017 at 6:22-23; 16:14-15.

197. As discussed above in step [1C], a POSA would be motivated to use amplification with Lo’s method. Further, a POSA would have been motivated to use region-specific primers for the PCR because Lo teaches already enrichment of specific target sequences prior to sequencing and counting. *See supra* ¶¶ 86, 165-168. The use of PCR with target-specific regions would be a straightforward way of combining the amplification and enrichment into a single step. Lo also identified specific chromosomal regions where an abnormal amount of that chromosomal region was relevant to a clinical diagnosis. EX1010 ¶ 4. Lo discusses using region-specific primers to amplify specific genomic regions in connection with digital PCR in its discussion of background methods that were available. *Id.* ¶ 10 (“Specific nucleic acid targets are amplified with sequence-specific primers to generate specific amplicons by digital PCR”). A POSA would recognize that one well-established

way to achieve such selective enrichment is by performing PCR with region-specific primers.

198. Further, a POSA would have a reasonable expectation of success in using region-specific primers because using region-specific primers with PCR amplification was commonplace, as described in McCloskey, Brenner, Chee, and Porreca. The use of region-specific primers allowed for successful targeted amplification of specific genome regions in those references.

199. In light of the above references, it would have been obvious for a POSA to perform PCR with region-specific primers in Lo.

3. Claim 9: The method of claim 1, wherein the unique tagged nucleic acid molecules are subjected to hybrid capture prior to the PCR or the amplified tagged nucleic acid molecules produced in step (c) are subjected to hybrid capture prior to step (d).

200. As discussed above, Lo in view of the knowledge of the skilled artisan renders obvious claim 1. Lo further teaches the features of claim 9, or, at a minimum, claim 9 would have been obvious in view of Lo.

201. Lo teaches an embodiment that uses hybrid capture prior to sequencing, wherein “the fraction of the nucleic acid pool that is sequenced in a run is further sub-selected prior to sequencing. For example, hybridization-based techniques such as oligonucleotide array could be used to first sub-select for nucleic acid sequences

from certain chromosomes, e.g. a potentially aneuploid chromosome and other chromosome(s) not involved in the aneuploidy tested.” EX1010 ¶ 72.

202. Moreover, Lo establishes why a POSA would be motivated to do this and that there would be a reasonable expectation of success. As Lo explains, “[a]s mentioned above and established in the example section below, only a portion of the human genome needs to be sequenced to differentiate trisomy 21 from euploid cases. Thus, it would be possible and cost-effective to enrich the pool of nucleic acids to be sequenced prior to random sequencing of a fraction of the enriched pool.” *Id.* ¶ 77; *id.* ¶ 79 (“The enriched pools of nucleic acids would then be subjected to random sequencing. This would allow the reduction in sequencing costs.”). Thus, as Lo makes clear, a POSA would be motivated to perform advance enrichment via hybrid capture because this would reduce the cost of the sequencing and, as Lo teaches, this would be also be possible using existing oligonucleotide array technologies.

203. Therefore, a POSA would find claim 9 obvious in view of Lo and the knowledge of the skilled artisan.

4. Claim 10: The method of claim 1, wherein the genomic nucleic acids are from a single species, two or more organisms, two or more species, or from a population of microbes.

204. As discussed above, Lo in view of the knowledge of the skilled artisan renders obvious claim 1. Lo further teaches the features of claim 10.

205. Lo discloses collecting and sequencing genomic DNA from pregnant women, which are humans and therefore a “single species.” EX1010 ¶ 92 (“Eight pregnant women were recruited for the study....Twenty milliliters of peripheral venous blood was collected from each subject....DNA was then extracted from 5-10 mL of each plasma sample.”). Additionally, maternal plasma samples contain genomic DNA of both fetal and maternal origin. *Id.* ¶ 6 (“First, fetal nucleic acids co-exist in maternal plasma with a high background of nucleic acids of maternal origin”).

206. Therefore, a POSA would find claim 10 obvious in view of Lo and the knowledge of a skilled artisan.

5. Claim 11: The method of claim 1, wherein the sample is a blood sample or bone marrow sample.

207. As discussed above, Lo in view of the knowledge of the skilled artisan renders obvious claim 1. Lo further teaches the features of claim 11

208. As already noted in ¶ 205, Lo’s protocol begins with a blood draw: “Twenty milliliters of peripheral venous blood was collected from each subject....The maternal plasma DNA was then used for massively parallel sequencing.” EX1010 ¶ 92.

209. Given Lo’s disclosure of collecting peripheral blood samples for downstream sequencing analysis, a POSA would find claim 11 obvious in view of Lo and the knowledge of a skilled artisan.

6. Claim 22: The method of claim 1, wherein the nucleic acid tags comprise a sufficiently large number of distinct nucleic acid tags.

210. As discussed above, Lo in view of the knowledge of the skilled artisan renders obvious claim 1. Claim 22, which recites the “method of claim 1, wherein the nucleic acid tags comprise a sufficiently large number of distinct nucleic acid tags” would have likewise been rendered obvious by specific knowledge of the skilled artisan, as reflected in disclosure in each of McCloskey, Brenner, Chee, and Porreca. Below, I discuss the relevant disclosure in each of these references and explain why this claim would have been obvious.

211. McCloskey’s encoding oligonucleotides include a seven-base “barcode” that is a “random sequence[s] of the four deoxyribonucleotides. The number of distinguishable barcodes in a population of oligonucleotides used in a reaction is determined by the number of random bases, n . This enables one to distinguish among 4^n allele copies per reaction. [McCloskey] use[s] seven nucleotides for the barcode, giving 16,384 [or 4^7] possible barcodes.” EX1008 at 764. A POSA would understand that increasing the length of the barcode sequence would allow for the creation of an even larger number of distinct tags. In other words, McCloskey teaches a recipe for generating a “sufficiently large number of distinct nucleic acid tags.”

212. Brenner teaches that the “size of mutually discriminable sets of oligonucleotide tags may [be]...50 to 1.6×10^7 .” EX1014 at 37:52-55; *see also id.* at 7:23-29 (noting that “tag adaptors collectively include a population of tags *sufficiently large*...at least ten times the number of fragments to be counted [and]...in another aspect...at least 100 times the number of fragments to be counted”). Teaching a number of tags that may be “100 times” the number of fragments to be counter, Brenner also teaches a “sufficiently large number of distinct nucleic acid tags.”

213. Chee teaches that the “diversity of the identifiers needs to be *sufficient* so that molecules that are not related won’t be mistaken as descendent ROIs.” EX1015 at 11:20-22, 35-36 (disclosing the use of 4^{20} (or around 10^{12}) possible sequences) (emphasis added). In one example calculation, Chee explains that one sequencing lane obtaining 10 million reads, with a 100-fold average construct redundancy would use a library of around “100 [thousand] separate constructs with unique identifiers.” *Id.* at 2:56-62. Thus, Chee also teaches a “sufficiently large number of distinct nucleic acid tags.”

214. Finally, Porreca teaches that “[f]or a set of targets, the number of differentiator tag sequences necessary to be confident (within some statistical bounds) that a certain differentiator tag sequence will not be observed more than once by chance in combination with a certain target sequence was determined. The

total number of unique differentiator tag sequences for a certain differentiator tag sequence length is determined as $4^{(\text{Length in nucleotides of the differentiator tag sequence})}$.” EX1017 at 73:10-17. Porreca further explains how the probability of two copies of a target sequence having the same tag varies with the number of total number of possible unique tags. *Id.* at 73:17-30 (“For a molecular inversion probe capture reaction that uses MIP probes having differentiator tag sequences, the probability of performing the capture reaction and capturing one or more copies of a target sequence having the same differentiator tag sequence is calculated as : $p=1-\frac{[N!/(N-M)!]}{[N]^M}$, wherein N is the total number of possible unique differentiator tag sequences and M is the number of target sequence copies in the capture reaction. Thus, by varying the differentiator tag sequence length it is possible to perform a MIP capture reaction in which the probability of capturing one or re copies of a target sequence having the same differentiator tag sequence is set at a predetermined probability value.”)

215. Porreca gives an example that “for a differentiator tag sequence of 15 nucleotides in length, there are 1,073,741,824 possible differentiator tag sequences. A MIP capture reaction in which MIP probes, each having a differentiator tag sequence of 15 nucleotides, are combined with 10000 target sequence copies (e.g., genome equivalents), the probability of capturing one or more copies of a target sequence having the same differentiator tag sequence is 0.05.” *Id.* at 73:31-38; *see*

also id. at FIG. 6. Thus, Porreca also teaches a “sufficiently large number of distinct nucleic acid tags.”

216. A POSA would have been motivated to use a sufficiently large number of tags because, as explained by Porreca above, a larger number of tags means a lower likelihood that DNA fragments with the same original sequence will end up with the same sequence when combined with a tag. A POSA would have had a reasonable likelihood of success because using a large number of tags was a well-known, conventional technique, as described above. Indeed, as taught in McCloskey, to get more tags, one simply needs to make the tag size longer, and techniques to make tags long enough to generate huge numbers of distinct sequences were commonplace, so a POSA would further have a reasonable expectation of successfully using longer tags as he or she would have using shorter tags.

7. Claim 26: The method of claim 22, wherein the nucleic acid tags comprise nucleotide sequences with constant portions and variable portions.

217. As discussed above, Lo in view of the knowledge of the skilled artisan renders obvious claim 22. Claim 26, which recites the “method of claim 22, wherein the nucleic acid tags comprise nucleotide sequences with constant portions and variable portions” would have likewise been rendered obvious by specific knowledge of the skilled artisan, as reflected in disclosure in each of McCloskey,

Brenner, Chee, and Porreca. Below, I discuss the relevant disclosure in each of these references and explain why this claim would have been obvious.

218. McCloskey discloses encoding oligonucleotides that include a fixed batch-stamp, a random barcode, and a primer-binding site. *See supra* § XI.B.2. As described above, the primer-binding site is constant for the target region, and the batch-stamp sequences is constant for the sample. The random barcode sequence is a variable portion of the tag.

219. Brenner discloses “[t]ag adaptor[s] compris[ing] primer-binding sites,...and sandwiched between such sites are the binary tags.” EX1014 at 8:31-33. The “[p]rimer binding site...has the same, or substantially the same, sequence whenever it is present.” *Id.* at 43:46-47. Brenner also discloses a variable “[o]ligonucleotide tag...sandwiched between restriction sites...which, in turn, are sandwiched between primer-binding sites.” *Id.* at 20:30-33. “The binary-tag component may take on any one of a very large number of species...for example, it may be a binary tag of any one of 2^{16} (~65,000) different sequences.” *Id.* at 14:30-33.

220. Chee describes embodiments where “the sets of the invention are populated by constructs comprising two or more amplification sites flanking the ROI and the unique identifier.” EX1015 at 2:65-67. The unique identifier is variable and

distinguishes between constructs; the amplification sites are constant and allow for amplification. *See supra* § XI.D.2..

221. Porreca likewise teaches adding tag sequences that contain both a variable differentiator sequence alongside a constant adaptor sequence. EX1017 at 31:21-27 (“Another method is to incorporate the differentiator tag sequence into a PCR primer, such that the primer structure is arranged with the common adaptor sequence followed by the PCR priming sequence (in 5’ to 3’ order). A differentiator tag sequence and adaptor sequence (which may contain the sequencing primer site) are incorporated as tags.”).

222. A POSA would have been motivated to use tags with both variable sequence portions and constant sequence portions as described in each of the references above. The variable portions allow for unique tagging of the regions of interest as described above in step [1B] above. *See supra* ¶¶ 148-152. The constant portions enable PCR amplification and sequencing, among other possible uses, as described in steps [1C] and [1D] above. *See supra* ¶¶ 165-168, 176. Adding both the constant and variable portions in a single tag would simplify the workflow and was commonplace.

223. A POSA would have had a reasonable expectation of success in using tags with both constant and variable tag portions because such tags were well-known and commonplace, as described in the references above. .

8. Claim 28: The method of claim 26, wherein the length of the constant portions and the length of the variable portions are different.

224. As discussed above, Lo in view of the knowledge of the skilled artisan renders obvious claim 26. Claim 28, which recites “[t]he method of claim 26, wherein the length of the constant portions and the length of the variable portions are different” would have likewise been rendered obvious by specific knowledge of the skilled artisan, as reflected in disclosure in each of McCloskey, Brenner, Chee, and Porreca. Below, I discuss the relevant disclosure in each of these references and explain why this claim would have been obvious.

225. McCloskey discloses an encoding oligonucleotide comprising an 8-nucleotide batch-stamp, a variable 7-nucleotide random barcode, a fixed 21-nucleotide primer-binding site, and a fixed 22-nucleotide genomic target-binding site. *See* EX1008 at 764.

226. As discussed in ¶ 219, Brenner teaches a “[t]ag adaptor...compris[ing] primer binding sites...and sandwich[ed] between such sites are binary tags.” EX1014 at 8:31-33. The forward primer site that recurs throughout the adaptor is the 9-nt sequence GCATCAAAA (preceded only by optional biotin and two random bases used for cloning), e.g., “5’-biotin-NN...GCATCAAAA...” *See id.* at 46:40-45. Brenner further discloses that the “lengths of the binary tags...may vary widely.... In one aspect,...binary tags comprise from 14 to 17 dinucleotide

subunits.” *Id.* at 10:39-41, 44-45; *see also id.* at 37:52-53 (“The size of mutually discriminable sets of oligonucleotide tags may vary widely”).

227. Chee teaches that the “number of nucleotides in the identifier is preferably chosen so that the number of potential and actual sequences represented is much greater than the total number of oligonucleotide constructs in the set. For example, if the identifier comprises a 20-mer synthesized with a mixture of A, C, G, and T at each position, there are 4^{20} possible sequences.” EX1015 at 11:30-36. In contrast, the amplification sites “allow amplification of the construct and/or the ROI.” *Id.* at 2:29-30. A POSA would understand that the unique identifier length would vary based on the number on number of desired possible unique sequences while the amplification site length depended on the amplification primer length.

228. Porreca teaches that the variable “[d]ifferentiator tag sequences may be...up to 25, or more nucleotides in length.” EX1017 at 29:17-21. In contrast, the constant “[a]dapters often contain PCR primer sites (for amplification or emulsion PCR) and/or sequencing primer sites.” 31:3-5. A POSA would understand that the length of the differentiator tag sequence is chosen based on the number of target molecules, while the adapter sequence length is chosen based on PCR and sequencing considerations.

229. In the context of a sufficiently large set of tags, as required by claim 26, a POSA would have understood that such diversity in tag structure—specifically,

using different lengths for constant and variable portions—was both necessary and standard practice to balance identification power, synthesis feasibility, and compatibility with sequencing workflows.

230. A POSA would be motivated to use this approach and have a reasonable expectation of successfully doing so for the same reasons explained above in connection with claims 22 and 26. Further, a POSA would have been motivated to use variable and constant tag portions of different length because different considerations drive the length of the variable and constant portions. The variable portion length is driven by considerations related to ensuring a sufficient number of tags to distinguish sample molecules, whereas the length of the constant portion is driven by considerations related to whatever length molecules are needed for the mechanics of amplification and sequencing. The two purposes are unrelated, and hence a POSA would not at all think to use sequence of the same length, but would rather arrive at different lengths depending on the relevant considerations for each portion of the tag.

9. Claims 33 and 34: The method of claim 22, wherein the nucleic acid tags comprise [nucleotide sequences that are not known before the tagged nucleic acid molecules are generated/random nucleotide sequences].

231. Claims 33 and 34 recite tags having “nucleotide sequence that are not known before the tagged nucleic acid molecules are generated” and “random nucleotide sequences,” respectively. In its infringement contentions, PO contends

that both of these claim elements may be satisfied by the use of tags with random sequences. *See* EX1020.325, .331. As I explain above in connection with claim element [1B], each of Brenner, Chee, Porreca, and McCloskey teach the use of random tags, and the use of these random tags in combination with Lo would have been obvious. *See supra* ¶¶ 140-159.

B. Ground 2: The Combination of Lo and the Knowledge of a Skilled Artisan in Further View of Specific Disclosure in McCloskey, Brenner, or Chee Renders Obvious Claim 3.

- 1. Claim 3: The method of claim 1, wherein the unique tagged nucleic acid molecules differ at more than one nucleotide or wherein the sequences of the nucleic acid tags are known prior to said generating unique tagged nucleic acid molecules from the segments of genomic nucleic acids.**

232. As discussed above, Lo in view of the knowledge of the skilled artisan renders obvious claim 1. Claim 3, which recites the “method of claim 1, wherein the unique tagged nucleic acid molecules differ at more than one nucleotide or wherein the sequences of the nucleic acid tags are known prior to said generating unique tagged nucleic acid molecules from the segments of genomic nucleic acids” would have been rendered obvious by specific knowledge of the skilled artisan, as reflected in the disclosure in each of McCloskey, Brenner, or Chee. Below, I discuss the relevant disclosure in each of these references and explain why this claim would have been obvious.

233. McCloskey's encoding oligonucleotides include a seven-base "barcode" that is a "random sequence[s] of the four deoxyribonucleotides. The number of distinguishable barcodes in a population of oligonucleotides used in a reaction is determined by the number of random bases, n . This enables one to distinguish among 4^n allele copies per reaction. "[McCloskey] use[s] seven nucleotides for the barcode, giving 16,384 [or 4^7] possible barcodes." EX1008 at 764. A POSA would understand that, with 16,384 different possible barcoded tags that may attach to a given molecule, tagged molecules routinely carry barcodes that differ at multiple nucleotide positions.

234. McCloskey's encoding oligonucleotide tags also consist of a "batch-stamps" that "specifies the DNA source...and the date of reaction," and that each batch stamp sequence "is registered for a specific sample run...and is used only once." *Id.* at 762, 764. A POSA would therefore understand that each batch-stamp sequence is predetermined and recorded before the tagging reaction begins and is thus "known prior to generating the tagged genomic fragments" as claim 3 requires.

235. Brenner teaches oligonucleotide tags that "differ maximally from one another by some characteristic, such as sequence." EX1014 at 37:50-51; *see also id.* at 25:56-59 ("[S]election of tag sequences that are, in some sense, maximally different, e.g. exemplary distance measures for sequences applicable to the invention are well known and are widely disclosed in the literature."). Brenner further notes

that “tag sequences that are maximally different” are desirable because they remain distinguishable “even in the presence of sequencing errors.” *Id.* at 21:16-18. A POSA would be motivated, therefore, to define the sequences in advance in order to ensure that the tags are “maximally different” as taught by Brenner.

236. Chee teaches attaching “a diverse but *pre-defined* set of sequences” to ROIs, including via PCR primers or by ligation. EX1015 at 11:18-20 (emphasis added). In addition, Chee describes dividing tagged nucleic acid constructs into subsets and selectively amplifying them using “subset-specific primers.” *See supra* at ¶ 113. Because each subset corresponds to a defined group of tagged constructs, this approach further presumes prior knowledge of tag identity.

237. Further, a POSA would have been motivated to use tags that differ by more than one nucleotide to distinguish between original molecules, especially given the potential for sequencing errors, which could obscure accurate molecule counting. This is particularly critical in Lo’s context, in which reliable fetal chromosomal quantification is necessary for prenatal diagnosis. By making the tags differ by more than a single nucleotide, one is more likely to ensure that the tags can be distinguished from one another despite potential sequencing errors. Brenner, for instance, teaches precisely this, stating that “[p]referably, sequences of oligonucleotide tags are selected so that they can be distinguished even in the presence of sequencing errors. Thus, oligonucleotide tags (4120) are selected to be

maximally different.” EX1014 at 21:14-18. A POSA would have easily had a reasonable expectation of success in making the tagged nucleic acid molecules differ by more than one nucleotide. Indeed, Brenner teaches that there were well known ways of making the tags “maximally different.” Brenner explains, for instance, that there were “distance measures” known in the art for designing such sequences. Brenner cites several literature references to this effect, stating “exemplary distance measures for sequences applicable to the invention are well known and are widely disclosed in the literature, as exemplified by the following references: Gusfield, Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology (Cambridge University Press, 1997); Navarro et al, Flexible Pattern Matching on Strings (Cambridge University Press, 2002); Sankoff et al, editors, Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison (Center for the Study of Languages, 1999); and the like.” *Id.* at 25:58-66.

238. Additionally, a POSA would see value in using pre-defined tag sequences—as taught by Chee—to simplify data interpretation, improve error correction, and support traceability of samples across sequencing workflows. A POSA would have a reasonable expectation of successfully doing this because it simply amounts to selecting tags in advance, using, for instance, several known procedures, including those described in Brenner.

239. Therefore, a POSA would have found it obvious to use tags that differ at multiple nucleotide positions or to use tag sequences that were known prior to tagging based on the disclosures of such tags in either McCloskey, Brenner, Chee, or Porreca.

C. Ground 3: The Combination of Lo and the Knowledge of a Skilled Artisan in Further View of Specific Disclosure in Brenner or Porreca Renders Obvious Claim 4.

1. Claim 4: The method of claim 1, wherein the segments of the genomic nucleic acids are produced by mechanical shearing, heating, or sonicating the genomic nucleic acids from the sample.

240. As discussed above, Lo in view of the knowledge of the skilled artisan renders obvious claim 1. Claim 4, which recites the “method of claim 1, wherein the segments of the genomic nucleic acids are produced by mechanical shearing, heating, or sonicating the genomic nucleic acids from the sample” would have likewise been rendered obvious by specific knowledge of the skilled artisan, as reflected in disclosure in each of Brenner and Porreca.

241. At the outset, I note that Lo’s method was generally performed using cell-free DNA from a pregnant woman. Such samples would generally not be appropriate for shearing, heating, or sonicating because this may reduce the amount of an already limited DNA sample. Nevertheless, as I explain above, Lo’s counting method for determining copy number had been applied prior to Lo in several other contexts, including with DNA samples that may need to be sheared, heated, or

sonicated prior to DNA sequencing. *See supra* ¶¶93-95.. This is taught, for instance, in Brenner and Porreca.

242. Brenner discloses generating target polynucleotide from source DNA “by shearing or by cleavage with one or more nucleases.” EX1014 at 18:33-34. Porreca likewise teaches fragmenting genomic DNA prior to sequencing using mechanical methods. For example, Porreca states that genomic DNA can be “[m]echanically shear[ed], using either a nebulizer, BioRuptor, Hydroshear, Covaris, or similar instrument” before sequencing. EX1017 at 71:16-17. Porreca explains that “DNA should be sheared into fragments several hundred basepairs in length” in order to prepare the sample for sequencing. *Id.* at 71:17-19.

243. A POSA would understand that some sequencing systems require DNA fragments of a specific size. Mechanical shearing, heating, and sonication were all known methods of creating smaller DNA fragments. A POSA would have thus been motivated to use these fragmentation techniques to adapt larger DNA molecules for sequencing as required by certain sequencing systems as described in step [1D]. *See, supra*, ¶¶ 176. Further, given that mechanical shearing, heating, and sonication were known and routine techniques in DNA preparation workflows, employing any of these methods to generic nucleic acid fragments would have been an obvious design choice and had a reasonable expectation of success.

D. Ground 4: The Combination of Lo and the Knowledge of a Skilled Artisan in Further View of Specific Disclosure in McCloskey, Brenner, or Porreca Renders Obvious Claims 5-8.

1. Claim 5: The method of claim 1, wherein the nucleic acid tags comprise a sample tag or a sample tag set.

244. A POSA would understand that Lo in view of the knowledge of the skilled artisan renders obvious claim 1. Claim 5, which recites “wherein the tags further comprise a sample tag or a sample tag set” is rendered obvious in further view of specific disclosure in Brenner, McCloskey, or Porreca. Below, I discuss the relevant disclosure in each of these references and explain why this claim would have been obvious.

245. As of the priority date, “sample tags” were well-known and used regularly. Briefly, prior to sequencing, the DNA from each sample is given a tag of a known sequence that identifies that DNA as arising from a particular sample. Each sample has its own unique tag. The DNA from multiple samples can then be pooled, and the pool of DNA can be sequenced at the same time. Sequencing generates the sequence reads for the sample tags as well as the attached molecules. Following sequencing, one can figure out which DNA sequences corresponded to a particular sample by identifying the sequences which contain the known sample tag sequence. Such sample tags were commercially available for widely used sequencing platforms as of at least 2009. *See, e.g.*, EX1039.002 (“The MID Adaptors include a IO-nucleotide sequence tag on Adaptor A which is unique for each MID. When different

libraries are prepared with different MIDs, they can be amplified by emPCR and sequenced together, in a multiplex fashion; the sequencing reads can be deconvoluted by the data analysis software after the sequencing Run, such that the reads from each of the pooled libraries are identified by their MID tag and correctly assigned.”). Each of Brenner, McCloskey, and Porreca discloses the use of a “sample tag” or a “sample tag set” in this manner. By allowing samples to be pooled and sequenced together at the same time, a POSA would be motivated to use such tags in order to improve efficiency and reduce costs.

246. Brenner discloses the use of “indexing” as one option for formulating oligonucleotide tags. As Brenner explains, “[s]uch indexing can also be used to analyze multiple samples using high throughput DNA sequencing instruments, where tags from different patients have different indexing sequences.” EX1014 at 21:29-32. A POSA would understand that this is a disclosure of the standard technique of using a tag on each molecule from a sample to label it as coming from that particular sample, sequencing the samples as a pool, and then deconvoluting the sequence data using the sample tag. .

247. McCloskey similarly teaches using a “batch stamp” tag alongside the “barcode” to identify specific samples. As McCloskey explains “[w]e have used batch-stamps of different lengths in different experiments. We typically use between six and eight base pairs to encode sample identity and date of reaction.... Each

different nucleotide sequence yields a unique batch-stamp. The assigned batch stamp is registered for a specific sample run on a specific date and is used only once.” EX1008 at 764; *see also id.* at 761 (“Batch-stamps specify the date and sample identity, and barcodes detect template redundancy.”); *id.* at 763 (“[T]he batch-stamp, unique to each experiment, specifies the DNA source such as the patient or sample identification”). A POSA would understand that this is a disclosure of the standard technique of using a tag on each molecule from a sample to label it as coming from that particular sample, sequencing the samples as a pool, and then deconvoluting the sequence data using the sample tag.

248. Finally, Porreca also discloses the use of sample tags, specifically pointing to McCloskey for this purpose. Porreca teaches that “in some embodiments, one or more of the techniques described herein may be combined in a single assay protocol for evaluating multiple patient samples in parallel.” EX1017 at 21:3-5. To this end, Porreca explains that “barcodes may be included, for example, to uniquely identify individual samples (e.g., patient samples) that may be mixed together. (See, e.g., USPTO Publication Number US 2007/0020640 A1 (McCloskey et al.).” *Id.* at 31:5-9. This, again, is an unambiguous disclosure of the use of a “sample tag.” A POSA would understand that this is a disclosure of the standard technique of using a tag on each molecule from a sample to label it as

coming from that particular sample, sequencing the samples as a pool, and then deconvoluting the sequence data using the sample tag.

249. A POSA would be motivated to utilize sample tags with Lo's sequencing and counting method. Sample tags provide a means to distinguish samples from each other and would permit more efficient analysis by allowing sequencing of multiple samples in a single sequencing run. Sample tags could also easily be included alongside the unique distinct tags which distinguish between nucleic acid molecule from the same sample. As reflected in Brenner, McCloskey, and Porreca the use of sample tags in this manner was a standard technique at the time and, accordingly, a POSA would have had a reasonable expectation of successfully using sample tags with Lo's method.

2. Claim 6: The method of claim 5, wherein the unique tagged nucleic acid molecules are pooled with a plurality of unique tagged nucleic acid molecules having a different sample tag prior to the PCR or prior to said sequencing of the amplified tagged nucleic acid molecules product of step (c).

250. A POSA would understand that Lo in view of the knowledge of the skilled artisan renders obvious claim 5. Claim 6, which recites "[t]he method of claim 5, wherein the unique tagged nucleic acid molecules are pooled with a plurality of unique tagged nucleic acid molecules having a different sample tag prior to the PCR or prior to said sequencing of the amplified tagged nucleic acid molecules

product of step (c)” is rendered obvious in further view of specific disclosure in Brenner, McCloskey, or Porreca.

251. Claim 6 recites nothing more than the standard manner of using a “sample tag” with sequencing, which I explain above in connection with claim 5. That is, the purpose of including a sample tag is to allow pooling of fragments from multiple samples prior to sequencing. Thus, for the same reasons that claim 5 would have been obvious, so too would claim 6 have been obvious.

3. Claim 7: The method of claim 6, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads.

252. As explained above, claim 6 would have been obvious to a POSA. Claim 7, which recites “[t]he method of claim 6, further comprising deconvoluting the tag associated sequence reads by grouping the tag associated sequence reads” is rendered obvious in further view of specific disclosure in Brenner, McCloskey, or Porreca.

253. Claim 7 recites nothing more than the standard manner of using a “sample tag” with sequencing, which I explain above in connection with claim 5. That is, the purpose of including a sample tag is to allow for pooling of samples prior to sequencing and by necessity the tag associated sequence reads must be deconvoluted after sequencing so as to group reads accurately within each sample.

Thus, for the same reasons that claims 5 and 6 would have been obvious, so too would claim 7 have been obvious.

4. Claim 8: The method of claim 1, wherein the nucleic acid tags are double-stranded.

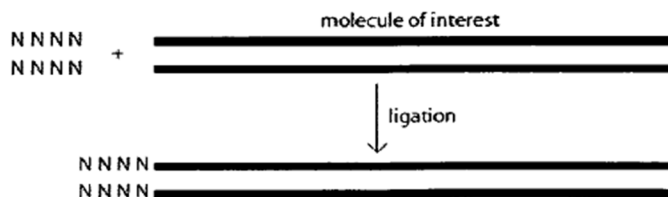
254. As discussed above, Lo in view of the knowledge of the skilled artisan renders obvious claim 1. Claim 8, which recites “[t]he method of claim 1, wherein the nucleic acid tags are double-stranded” is rendered obvious by disclosure in each of Brenner, McCloskey, and Porreca.

255. Brenner discloses “a set of 64 double stranded composite tags” and specifies that “each composite tag contains a double stranded metric tag, both of whose complementary single strands are in the range of from 28 to 91 bases in length.” EX1014 at 28:38-41.

256. McCloskey discloses a double-stranded hairpin loop tag, with a randomly generated barcode embedded in the loop region. EX1008 at 764. McCloskey explains that the tag’s secondary loop structure serves a functional purpose: it prevents the random barcode from hybridizing with the target sequence. *Id.* at 764-65; *see also id.* at 764 (noting that “only the genomic target-binding site is available for annealing [with the target] at 37°C, while the remaining bases are unavailable due to secondary structure”).

257. Porreca also discloses double-stranded oligonucleotide tags which ligate to both strands of a double-stranded “molecule of interest.” EX1017 at FIG. 4C.

FIG. 4C



Id. at FIG. 4C.

258. A POSA would be motivated to use double-stranded tags as one of two options for tagging (choosing between single- and double-stranded tags) and the option that is often used for ligation. Double-stranded DNA would thus have been obvious to try, in particular because DNA exists in nature in double-stranded form. A POSA would have a reasonable expectation of success because Brenner, McCloskey, and Porreca collectively demonstrate double stranded tags' utility, compatibility with PCR, and effectiveness in molecular tagging protocols.

E. Ground 5: The Combination of Lo and the Knowledge of a Skilled Artisan in Further View of Specific Disclosure in Brenner, Chee, or Porreca Renders Obvious Claims 12, 15, 17, 23, and 35-36.

- 1. Claim 12: The method of claim 1, wherein the unique tagged nucleic acid molecules are unique based on a combination of sequence information in the nucleic acid**

tags and sequence information in the segments of the genomic nucleic acids.

259. A POSA would understand that Lo in view of the knowledge of the skilled artisan renders obvious claim 1. Claim 12, which recites the “method of claim 1, wherein the unique tagged nucleic acid molecules are unique based on a combination of sequence information in the nucleic acid tags and sequence information in the segments of the genomic nucleic acids” is rendered obvious in further view of disclosure in Brenner, Chee, and Porreca.

260. Brenner teaches that “substantially every molecule to be counted in a sample, e.g. the number of single stranded DNA molecules of a particular genetic locus in a sample of genomic DNA, is associated with a probe having a different tag.” EX1014 at 5:17-21. Brenner also teaches that “a size of tag population is selected that ensures with *high probability* that substantially every target polynucleotide will have a unique tag. In one aspect, such probability is at least 90 percent; in another aspect, such probability is at least 95 percent; and in another aspect, such probability is at least 99 percent.” *Id.* at 18:2-7. As Brenner makes clear, one need not necessarily use a number of tags that exceeds the number of target polynucleotides. There can be fewer tags than there are targets. In such circumstances, it will be the combination of the tag and the information in the genomic nucleic acid that renders the tagged molecule unique.

261. Chee teaches “unique identifiers” that are not necessarily absolutely unique, but that have an “extremely high probability of being unique in [its] association with a particular ROI in a construct.” EX1015 at 9:60-63. Chee describes that “a feature of the invention [is] that the different constructs have unique identifiers, including constructs that may have the same ROI, so the specific, correct ROIs can be distinctly identified and distinguished from the other constructs in the set.” *Id.* at 5:27-31. Chee further teaches that the “sets of constructs are designed so there is a large excess of possible identifiers relative to the ROIs in a given set. Thus, the chances of a having constructs with given identifiers associated with more than one ROI is very low, and the chances of the same identifier being associated with the same ROI in different constructs is extremely low.” *Id.* at 9:54-60. Thus, Chee, like Brenner, makes clear that not every nucleotide must have a different tag, such that uniqueness is conferred not solely by the tags, but by the combination of the tag and genomic nucleic acid.

262. Porreca similarly teaches a method “wherein the isolated nucleic acids are identified as independently isolated if they comprise unique ***combinations*** of target nucleic acid and differentiator tag sequences.” EX1017 at 5:41-43. Porreca explains that “[d]ifferentiator tag sequences are typically of a predefined length, which is selected to control the likelihood of producing unique target:differentiator tag sequences in a preparative reaction.” *Id.* at 29:12-15. Porreca uses these unique

target:differentiator sequences for counting. *Id.* at 72:67-73:2 (“For any target:differentiator tag combination observed more than once, all such reads are ‘collapsed’ into a single read before consensus basecalls are determined.”). Thus, in Porreca, one uses the “combination” of the target and tag to confer uniqueness.

263. A POSA would be motivated to adopt this approach wherein uniqueness is conferred by the combination of the target and tag. Indeed, while a POSA may strive to have a large number of tags to confer uniqueness, it may not in all possible cases be cost-effective to generate such a large number of tags. As set forth in Brenner, Chee, or Porreca, a POSA would recognize that uniqueness could still be established based on the combination of tag and target nucleic acid even if there were somewhat fewer tags than targets. There would be no question of reasonable expectation of success here given the wide diversity of sample molecules in a typical DNA sample and the ability of a POSA to generate large numbers of tags simply by increasing tag length.

264. I note that, even in circumstances where all or substantially all of the genomic nucleic acid segments are expected to be uniquely tagged, a POSA would still be motivated to combine the information from the tag and the genomic nucleic acid itself to ensure uniqueness. Both the tags and genomic nucleic acid segments are necessarily sequenced and analyzed in this process; as such, the information from both portions of the tagged nucleic acids is readily available and would naturally be

combined by a POSA. Indeed, this was a commonplace step at the time and thus a POSA would have been motivated to perform it and have a reasonable expectation of success.

2. Claim 15: The method of claim 12, wherein the nucleic acid tags comprise nucleotide sequences with constant portions and variable portions.

265. As discussed above, Lo in view of the knowledge of the skilled artisan renders obvious claim 12. Claim 15, which recites “[t]he method of claim 12, wherein the nucleic acid tags comprise nucleotide sequences with constant portions and variable portions” is rendered obvious in further view of specific disclosure in Brenner, Chee, or Porreca.

266. As described above for claim 26, Brenner teaches tag structures that incorporate constant regions, such as primer binding sites, alongside variable regions for indexing or sample identification. EX1014 at 43:51-53, 14:30-33; *see supra* ¶ 219.

267. Similarly Chee describes embodiments with a variable unique identifier which distinguishes between constructs and constant amplification sites. EX1015 at 2:65-67; *see supra* ¶ 220.

268. Porreca likewise teaches adding tag sequences that contain both a variable differentiator sequence alongside a constant adaptor sequence. EX1017 at 31:21-27; *see supra* ¶ 221.

269. A POSA would have been motivated to use tags with both constant and variable regions, in order to increase the uniqueness of tagged molecules, while still maintaining ease of amplification and sequencing as described above in reference to claim 26. *See supra* ¶ 222.

270. A POSA would have had a reasonable expectation of success in using tags with both constant and variable tag portions because such tags were well-known and commonplace, as described in the references above and for the same reason that there would have been a reasonable expectation of success with respect to claim 26.

3. Claim 17: The method of claim 15, wherein the length of the constant portions and the length of the variable portions are different.

271. As discussed above, Lo in view of the knowledge of the skilled artisan renders obvious claim 15. Claim 17, which recites “wherein the length of the constant portions and the length of the variable portions are different” is rendered obvious in further view of specific disclosure in Brenner, Chee, or Porreca.

272. As discussed above with references to claim 28, each of Brenner, Chee and Porreca discloses tags with constant and variable portions, wherein the lengths of each type of tag region varies. *See, supra*, ¶¶ 226-228.

273. Brenner, for example, discloses variable binary tags that “comprise from 14 to 17 dinucleotide subunits” which are sandwiched between constant primer tags that are 9 nucleotides in length. EX1014 at 10:39-41, 44-45; *see also id.* at

37:52-53 (“The size of mutually discriminable sets of oligonucleotide tags may vary widely”).

274. Chee likewise teaches that the “number of nucleotides in the identifier is preferably chosen so that the number of potential and actual sequences represented is much greater than the total number of oligonucleotide constructs in the set.” EX1015 at 11:30-33. Chee gives an example of an identifier 20 nucleotides in length with 10^{12} possible sequences. *Id.* at 11:33-36. In contrast, the amplification sites “allow amplification of the construct and/or the ROI.” *Id.* at 2:29-30. A POSA would understand that the unique identifier length would vary based on the number on number of desired possible unique sequences while the amplification site length would depend on the amplification primer length.

275. Porreca also teaches that the variable “[d]ifferentiator tag sequences may be...up to 25, or more nucleotides in length.” EX1017 at 29:17-21. In contrast, the constant “[a]dapters often contain PCR primer sites (for amplification or emulsion PCR) and/or sequencing primer sites.” 31:3-5. A POSA would understand that the length of the differentiator tag sequence is chosen based on the number of target molecules, while the adapter sequence length is chosen based on PCR and sequencing considerations.

276. In the context of uniqueness conferred by a combination of the tag and genomic nucleic acid segments, as required by claims 12 and 15, a POSA would

have understood that such diversity in tag structure—specifically, using different lengths for constant and variable portions—was both necessary and standard practice to balance identification power, synthesis feasibility, and compatibility with sequencing workflows. A POSA would have been motivated to have tags with constant portions and variable portions with different lengths and reasonably expected success for the same reasons set forth above with respect to claim 28. *See supra* ¶ 230.

4. Claim 23: The method of claim 22, wherein the unique tagged nucleic acid molecules are unique based on a combination of sequence information in the nucleic acid tags and sequence information in the segments of the genomic nucleic acids.

277. A POSA would understand that Lo in view of the knowledge of the skilled artisan renders obvious claim 22. Claim 23, which recites “wherein the unique tagged nucleic acid molecules are unique based on a combination of sequence information in the nucleic acid tags and sequence information in the segments of the genomic nucleic acids” is rendered obvious in further view of specific disclosure in Brenner, Chee, or Porreca.

278. Claim 23, much like claim 12, recites nothing more than the standard manner of using tags to ensure unique identification of tagged nucleic acids after amplification. As I explain above for claim 12, each of Brenner, Chee, and Porreca

disclose unique tagged nucleic acid molecules based on a combination of tag and genomic nucleic acid. *See, supra*, ¶¶ 260-262.

- Brenner teaches that “a size of tag population is selected that ensures with **high probability** that substantially every target polynucleotide will have a unique tag. In one aspect, such probability is at least 90 percent; in another aspect, such probability is at least 95 percent; and in another aspect, such probability is at least 99 percent.” EX1014 at 18:2-7.
- Chee teaches that “the chances of a having constructs with given identifiers associated with more than one ROI is very low, and the chances of the same identifier being associated with the same ROI in different constructs is extremely low.” EX1015 at 9:56-60.
- Porreca similarly teaches a method “wherein the isolated nucleic acids are identified as independently isolated if they comprise unique **combinations** of target nucleic acid and differentiator tag sequences.” EX1017 at 5:41-43. Thus, in Porreca, one uses the “combination” of the target and tag to confer uniqueness.

279. As discussed above in claim 12, a POSA would naturally attempt to create a sufficiently large number of tags to ensure uniqueness of the tagged nucleic acid molecules but would still understand, based on Brenner, Chee, and Porreca, that

uniqueness may be conferred due to the *combination of tag + genomic nucleic acid*. See *supra* ¶¶ 263-264. A POSA would be further motivated to adopt this approach because it may not in all possible cases be cost-effective to generate such a large number of tags. For the same reasons as stated with respect to claim 12, there would be no question of reasonable expectation of success here given the wide diversity of sample molecules in a typical DNA sample and the ability of a POSA to generate large numbers of tags simply by increasing tag length.

5. Claims 35-36: The method of [claim 22/claim 1] wherein the nucleic acid tags are substantially unique nucleic acid tags.

280. In its infringement contentions, PO contends that this claim element is satisfied by “nonredundant” tags and tag sets that are “non-unique.” See EX1020.340, .342, .345, .347. To the extent the specification defines “substantially unique tags,” PO does not apply this definition. As I explain above, for the purposes of this Petition, I apply the interpretations reflected in PO’s infringement contentions, which treat this claim element as merely encompassing tag sets that are “non redundant” in that not all the tags are the same (*i.e.*, “non-unique”).

281. As I explain above in connection with claim 12, Brenner teaches that “a size of tag population is selected that ensures with high probability that substantially every target polynucleotide will have a unique tag. In one aspect, such probability is at least 90 percent; in another aspect, such probability is at least 95 percent; and in another aspect, such probability is at least 99 percent.” EX1014 at

18:2-7. Likewise, Chee teaches that “sets of constructs are designed so there is a large excess of possible identifiers relative to the ROIs in a given set.” EX1015 at 9:54-56. And Porreca teaches that the total number of tag sequences is 4^n , where n is the number number of nucleotides in the tag. EX1017 at 73:14-17. Porreca goes on to contemplate using more than a billion different tags. *Id.* at 73:31-32. Thus, each of Chee, Brenner, and Porreca contemplate using large numbers of tags that will all be different from one another and, in that regard, are nonredundant

282. A POSA would be motivated to use tags having this property because it would help ensure overall uniqueness of different the tag-target constructs in advance of amplification and would hence increase the accuracy of the counting. Moreover, there is no question of reasonable expectation of success because this approach would not require any changes to the normal process for creating tags and would have simply required manufacturing a large set of tags as needed compared the number of sample molecules.

F. Ground 6: Claims 18-21 And 29-32 Are Obvious In Further View Of The Knowledge Of The Skilled Artisan

283. Claims 18-21 and 29-32 recite tags with variable portions that are constructed from particular palettes of nucleotides:

- Claims 18 and 29 recite using only “two species of nucleotides.”
- Claims 19 and 30 recite using only “three species of nucleotides.”

- Claims 20 and 31 recite tags with variable portions “constructed by a set of dinucleotides.”
- Claims 21 and 32 recite tags with variable portions “constructed by a set of trinucleotides.”

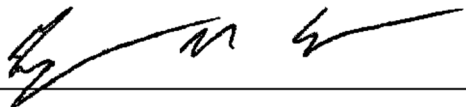
284. As I explain above, the claims from which these eight claims depend are all invalid as obvious. The use of tags with variable portions constructed from the claimed palettes of oligonucleotides also would have been obvious approaches that a POSA would have been motivated to explore. Indeed, using fewer than all possible species of oligonucleotides would simplify the synthesis process and simplify the data analysis because variable portions of tags could be distinguished from other tag regions by virtue of only including a limited repertoire of nucleotides. Similarly, a POSA would have been motivated to consider variable tag portions constructed from dinucleotides and trinucleotides because this would allow one to effectively have an expanded set of units for synthesis of tags without introducing any new nucleotides. This is because while there are only four nucleotides, there are 16 possible di nucleotides and 64 possible trinucleotides. To simplify the design of unique tags, a POSA would be motivated to use di- and trinucleotides as the basic building blocks. *See, e.g.*, EX1037 at 2:61-3:44. Further, a POSA would have had a reasonable expectation of successfully pursuing these options because the

technology for synthesis of oligonucleotides was well-developed as of 2010. *See, e.g.*, EX1038.

DECLARATION

I declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are belief 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: July 25, 2025

By: 
Gregory Cooper, Ph.D.