

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

PROGENITY, INC.,
Petitioner,

v.

NATERA, INC.,
Patent Owner.

Case IPR2021-00267
Patent No. 10,240,202

**DECLARATION OF PROFESSOR JOHN QUACKENBUSH IN SUPPORT
OF PATENT OWNER'S PRELIMINARY RESPONSE**

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

TABLE OF CONTENTS

	<u>Page</u>
I. ENGAGEMENT AND COMPENSATION	1
II. QUALIFICATIONS	2
III. RELEVANT TIMEFRAME FOR DETERMINING OBVIOUSNESS	8
IV. TECHNICAL INTRODUCTION	9
V. PROGENITY’S ASSERTED REFERENCES	17
A. Shimkets	17
1. Complications Associated With Applying Shimkets To Cell-Free Fetal DNA In A Maternal Blood Sample	34
(a) Complex-sample sequencing is not suitable for quantitative testing.....	34
(b) Shimkets uses its sequence-based karyotyping on cells and not samples of mixed genotypes	38
(c) Non-randomly fragmented DNA sequences can introduce systematic error	48
2. At the time of the invention, researchers, including Bianchi, had concluded that using fetal cells derived from maternal blood was “not clinically practical”	51
3. With the exclusion of large swaths of measured sequences, Shimkets analysis does not necessarily include any amounts of genetic material measured at multiple polymorphic loci. .	54
B. Huang	58
C. Bianchi.....	67

I, Professor John Quackenbush, Ph.D., hereby declare:

I. ENGAGEMENT AND COMPENSATION

1. My name is John Quackenbush (Ex. 2013). I have been retained by Natera, Inc. (“Natera”) to serve as an expert in this *inter partes* review proceeding concerning U.S. Pat. No. 9,424,392 (the “’392 Patent”), U.S. Patent No. 10,240,202 (the “’202 Patent”) and U.S. Patent No. 10,227,652 (the “’652 Patent”). I have been asked to provide my technical review, analysis, insights, and opinions regarding the references that form the basis for the Grounds of unpatentability set forth in the Petition for IPR2021-00266, IPR2021-00267 and IPR2021-00279.

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

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

4. In providing the testimony, I have considered each of the documents cited herein, the petition and the exhibits submitted therewith. I provide my testimony from the perspective of a person of ordinary skill in the art at the time of

the invention as Progenity has defined in its petition.¹ I am familiar with the knowledge and skill level of these people based on my nearly thirty years of experience in the field and frequent interactions with such people. I am also familiar with the state of the art at the time of the invention, as discussed in Section III below. Because the '392, '652 and '202 Patents share a common specification and because the arguments in the three IPRs are largely the same, my discussion of the patented specification cites only to the '392 Patent, which is submitted as Exhibit 1001 in each of the three IPR proceedings mentioned in Paragraph 1. If I discuss a term from the '392 Patent, I apply the plain and ordinary meaning of the term as understood by one of ordinary skill in the art at the time of the invention in light of the specification of the '392 Patent.

II. QUALIFICATIONS

5. I am a Henry Pickering Walcott Professor of Computational Biology and Bioinformatics and Chair of the Department of Biostatistics at the Harvard T.H. Chan School of Public Health. I am also Founding Director and co-Director of the MS Program in Computational Biology and Quantitative Genetics and Director of

¹ Progenity defines a person of ordinary skill in the art as one who “as of the filing date of the ' 392 patent would have possessed a post-undergraduate degree, *e.g.*, Ph.D. or Ph.D/M.D., in Molecular Biology, Molecular Genetics, Bioinformatics, Biostatistics, Obstetrics/Gynecology, Molecular Pathology, or a comparable discipline, and at least 2 years of experience relating to detecting chromosomal abnormalities.” IPR2021-00266 Pet. 9.

the Big Data to Knowledge Training Program, co-Director of the Cancer Genomics Training Program, and the Faculty Director for the Harvard Quantitative Biomedical Research Center. In addition, I hold appointments as Professor of Biostatistics and Computational Biology at the Dana-Farber Cancer Institute and at the Channing Division of Network Medicine of Brigham and Women's Hospital.

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

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

for the National Research Council, the National Academy of Sciences, and the Institute of Medicine.

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

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

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

11. I have studied, taught, practiced, and conducted research involving the use of genomic technologies, including DNA sequencing and DNA microarrays, and involving computational analysis of data related to genomic technologies. I have expertise in genomic technologies, computational analysis, and modeling of a plant, animal, and microbial systems with an emphasis on the study of human diseases, such as cancer.

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

methods for data analysis and application of those methods to the biological and biomedical data sets, including those generated using techniques such as genomics, functional genomics, and epigenomics. The beauty of all of these techniques is that their development and use is, largely although not entirely, independent of their application area. So methods developed on one area are often used in other areas.

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

14. I have also done work relevant to reproductive health, an area that became personally important to me as my wife and I, then in our forties, became enmeshed in the use of assisted reproductive technology. This gave me first-hand exposure to a wide range of techniques, including pre-implantation prenatal screening and the use and limitations of technologies directed toward prenatal

detection of genetic abnormalities including aneuploidies. Following the birth of my son in 2006 at Brigham and Women's Hospital in Boston, I also worked with, and mentored, physician scientists at Harvard, published a genomic analysis of diminished ovarian reserve, and presented the President's plenary lecture at the 63rd American Society for Reproductive Medicine Annual Meeting in 2007 in which I described the use of genomic technology in reproductive, prenatal, and neonatal health.

15. I have received numerous awards and honors including being recognized as a White House Open Science Champion of Change, an Australia Fellow, and a Distinguished International Advisor for the International Conference of BioInformatics and BioEngineering sponsored by the Institute of Electrical and Electronics Engineers.

16. I have published more than 300 peer reviewed scientific papers that have been cited more than 79,000 times. I have had continuous grant support for my research from the National Institutes of Health and other organizations for nearly thirty years. In 2018, I received an NCI Outstanding Investigator Career Award, a seven-year research grant recognizing a lifetime record of substantial contributions to the field.

17. A copy of my Curriculum Vitae, attached as Exhibit 2023, contains further details on my education, experience, publications, and other qualifications to render an expert opinion in this matter.

18. I note that I rely on a number of journal articles and other scientific materials in this declaration. In my opinion, these documents were published on the respective dates noted in each of the publications; some scientific papers may be published online before appearing in print and in these instances I consider the earliest date the work appears to be the publication date. In my opinion the journals and other venues in which these materials were published are reputable and these materials are the type of materials that experts in the fields of study relating to the methods of the '392 patent would reasonably rely on.

III. RELEVANT TIMEFRAME FOR DETERMINING OBVIOUSNESS

19. I understand that obviousness must be evaluated “at the time of the invention.” The '392 Patent claims priority to seven provisional applications filed between November 26, 2005 and September 22, 2006. Ex. 1001, cover (60), 1:9-18. The '392 Patent also claims priority to U.S. Pat. No. 8,532,930 filed on November 22, 2006. *Id.*, cover (63). The '202 and '652 Patents are respectively child and grandchild patent of the '392 Patent. *See* EX1002, pp.1-2 (63); EX1003, cover (63).

20. For the purpose of this declaration, I will use November 26, 2005 as the time of the invention. My testimony is directed to state of the art as of that date, even

if I do not always use a past tense. I may refer to timeframe at or before November 26, 2005 as “the relevant timeframe.” The relevant state of art and knowledge of the POSITA, however, would not change between November 26, 2005 and November 22, 2006.

IV. TECHNICAL INTRODUCTION

21. The '392 Patent is titled: System and Method for Cleaning Noisy Genetic Data from Target Individuals Using Genetic Data from Genetically Related Individuals.” Ex. 1001, Title. The '392 Patent relates generally to systems and methods “for determining the genetic data for one or a small set of cells, or from fragmentary DNA, where a limited quantity of genetic data is available.” Ex. 1001, Title, Abstract. The '392 Patent describes and claims methods for measuring and analyzing genetic data for “detecting the presence or absence of a chromosomal abnormality in a fetus” based on “a sample comprising DNA from the fetus and DNA from the mother of the fetus.” Ex. 1001, claims 1(a) and 11(a). Such mixed DNA samples may be a maternal blood sample containing small quantities of cell-free DNA. One form of fetal chromosomal abnormalities may be a fetal aneuploidy such as a trisomy on chromosome 13, 18 or 21.

22. Normal humans have 23 chromosome pairs in every diploid cell (22 pairs of nearly identical chromosomes called autosomes, and a pair of sex chromosomes—XX in females, XY in males), with one copy coming from each

parent. Aneuploidy, the state of having an extra or missing chromosomes, is believed to be responsible for a large percentage of failed embryonic implantations, miscarriages, and genetic diseases. *See, e.g.*, Ex. 1001, 6:66-7:5. Accordingly, early detection of abnormalities in chromosome numbers is highly desirable.

23. The field of prenatal genetic diagnosis seeks to alert physicians and parents to abnormalities, such as aneuploidy, in developing fetuses. The '392 Patent teaches that its methods have particular application in the field of prenatal genetic diagnosis where they enable reliable detection of fetal chromosomal copy number. The '392 Patent teaches that its methods enable effective and reliable non-invasive prenatal genetic testing, where, for example, a small quantity of fetal genetic material is isolated from maternal blood in order to analyze the developing fetus. *See, e.g.*, Ex. 1001, 8:33-39.

24. In the 2005 timeframe, prenatal testing of developing fetuses was performed using genetic material (e.g., DNA) obtained from conventional and invasive techniques available at that time, such as amniocentesis and chorionic villus sampling ("CVS"). *See, e.g.*, Ex. 1001, 1:31-49. Amniocentesis and CVS make use of fetal genetic material obtained from, respectively, amniotic fluid and placenta. Samples derived from amniocentesis and CVS in prenatal testing contain fetal but not maternal DNA and the data interpretation is therefore relatively straightforward. Such techniques (amniocentesis and CVS), however, are invasive and pose a risk to

the developing fetus. *See id.* Thus, at the time, alternatives to these limited and invasive procedures were needed.

25. At the time of the development of the inventions claimed in the '392 Patent, researchers had begun investigating the use of a small amount of genetic material, such as extracellular fetal DNA in maternal circulation or a single embryonic cell, to perform prenatal or preimplantation diagnostic testing. If fetal cells or fetal extracellular DNA could be isolated and tested to determine the genetic state of the developing fetus, this would alleviate the need to perform much more invasive and risky procedures to obtain DNA from the fetus directly.

26. One problem with developing new prenatal and pre-implantation diagnostic tests at the time of the inventions claimed in the '392 Patent was that the amount of fetal genetic material available, from either a pre-implantation embryo or maternal blood sample, was small. *See* Ex. 1001, 1:50-62.

27. At the relevant time, a single cell could be safely removed from a pre-implantation embryo. However, a single human diploid cell contains minute amounts of DNA, recognized to be about 7 picograms ($7 \text{ pg} = 7 \times 10^{-12} \text{ g}$). *See, e.g.,* Ex. 2016 [Tiersch 1989], 707-08 (reporting ~7 pg DNA per cell). This is consistent with my understanding that a single copy of the human genome (a single genome equivalent of DNA) would have a mass of approximately 3.6 pg.

28. Fetal cells in maternal blood were known to be rare and difficult to isolate. *See, e.g.,* Ex. 2013 [Bischoff2003], Abstract (noting “five independent laboratories were unable to identify any intact male cells from the plasma of 38 women known to be carrying male fetuses. ... [W]e believe that recovery of these cells is difficult and not clinically practical.”). Even when fetal DNA is present in an intact cell, one difficulty is identifying which cells are fetal from among the (much more numerous) maternal cells in maternal blood.

29. Fetal extracellular DNA, commonly referred to as cell-free fetal DNA, consists of fragments of fetal DNA that circulate freely in the maternal blood. Fetal extracellular DNA was known to be present in maternal circulation in varying, but extremely limited quantities and at a tiny fraction of the background cell free maternal DNA present. *See* Ex. 1008, S93 (reporting a study by Lo which observed average fraction of fetal cell-free DNA in maternal plasma to be 3.4% in “early pregnancy” and 6.2% in “late pregnancy”). In typical maternal blood samples, the exact fraction of fetal DNA is unknown, and can vary from sample to sample. Bianchi, for example, reported that in her studies, “there was a wide distribution of fetal DNA levels in both [trisomy] cases and controls, which decreased the average separation between these two groups.” Ex. 1008, S97. When fetal DNA is extracellular, part of the difficulty is identifying cell-free DNA of fetal origin against the overwhelming background of maternal DNA.

30. If fetal cells or fetal extracellular DNA could be isolated from maternal blood and tested to determine the genetic state of the developing fetus, this would alleviate the need to perform much more invasive and risky procedures to obtain DNA from the fetus directly. In order for a small amount of genetic material, such as DNA from a single cell or from cell-free DNA to be used for genetic diagnosis, that DNA generally must be amplified. As the '392 Patent itself teaches, using a small amount of starting DNA and using standard amplification techniques can result in significant problems with genetic data derived from that material. These problems can be caused when that DNA is amplified to quantities suitable for analysis:

There are numerous difficulties in using DNA amplification in these contexts. Amplification of single-cell DNA (or DNA from a small number of cells, or from smaller amounts of DNA) by PCR can fail completely, as reported in 5-10% of the cases. Other sources of error that may arise in measuring the embryonic DNA by amplification and microarray analysis include transcription errors introduced by the DNA polymerase where a particular nucleotide is incorrectly copied during PCR, and microarray reading errors due to imperfect hybridization on the array. The biggest problem, however, remains allele drop-out (ADO) defined as the failure to amplify one of the two alleles in a heterozygous cell. ADO can

affect up to more than 40% of amplifications and has already caused PGD misdiagnoses.

Ex. 1001, 3:9-25.

31. Accordingly, with such small quantities of genetic material, the measured genetic data often contained incorrect, unclear, spurious, and missing measurements. *Id.*, 59:34-39. Thus, the resulting genetic data was “noisy.” *Id.*, 8:15-24. As the ’392 Patent explains, “[f]or example, using current techniques, uncleaned measured genetic data from DNA amplified from a single cell may contain between 20% and 50% unmeasured regions, or allele dropouts.” *Id.*, 59:41-44. As the ’392 Patent teaches, generating genetic data from a small amount of genetic material is challenging for other reasons as well, including amplifying extremely dilute DNA concentrations in small volumes and reliably dissociating DNA from proteins across the whole genome. *Id.*, 3:2-6.

32. A single-nucleotide polymorphism (“SNP”) is a substitution of a single nucleotide that occurs at a specific position in the genome, where each variation is present at some level in the population. *See, e.g.*, Ex. 1010, 376. With the discovery of a single nucleotide polymorphisms, researchers began investigating the possibility of using SNP-based analyses to determine chromosomal copy number deviations. *See, e.g.*, Ex. 2017 [Weiss 6th Ed.], 3060 (analyzing SNPs to characterize DNA copy number changes including amplification and homozygous deletion from

a subset of lung and breast carcinoma cell lines and lung tumors). At the same time, as the '392 Patent reports, “[s]everal techniques [were] in development to measure multiple SNPs on the DNA of a small number of cells” but, as the '392 Patent correctly explains, “[i]t is apparent that the techniques will be severely error-prone due to the limited amount of genetic material which will exacerbate the impact of effects such as allele drop-outs, imperfect hybridization, and contamination.” Ex. 1001, 4:4-19.

33. The '392 Patent discloses and claims the use of SNP-genotyping or high throughput DNA sequencing technologies to analyze at least 100 loci on a chromosome of interest from the genetic material of an embryo or fetus. The '392 Patent then teaches determining the true state of the underlying genetic material by “creating a set of hypotheses about the potential states of the DNA, and testing to see which one has the highest probability of being true given the measured data.” *See, e.g., id.*, 9:7-9. The “relative probability of each of the hypotheses given the measurements of the target individual’s genetic data” is calculated. *See id.*, 26:61-65. The “relative probabilities associated with each hypothesis” are then used to determine the number of instances of any given chromosome or chromosomal segment. *See id.* Ultimately, the “hypothesis that is calculated to be most likely is...expected to correspond to the true state of the genetic material in the target DNA sample.” *See, e.g., id.*, 39:61-64.

34. This process of assessing multiple alternative hypotheses to estimate the probability of which is most likely given the data is fundamentally different than simply comparing a single measurement to a “control” distribution. The latter depends on examining a single distribution of measurements, about which we can only say what the likelihood is that a particular measurement came from the distribution by chance, not whether there is a far greater likelihood it represents something else. For example, we might be able to say that there is only a 5% chance that a measurement represents a diploid chromosome, but we would not be able to draw any conclusion about the likelihood that the measurement therefore shows the chromosome is aneuploid. To make an aneuploidy call, we need to compare the measurement to a distribution of known aneuploidy cases and to make sure that there is no systematic error that causes the observed effects.

35. The probability computation at issue in the '392, '202 and '652 Patents is based on performing multiple statistical tests of different models and asking which model best fits the data, then presenting the results of such an analysis with an assessment of the likelihood of the various alternate hypotheses. For example, one might determine that there is a 25% chance that a measurement represents a diploid chromosome but a 78% chance that it represents a triploid chromosome. One would be able to make such a statement because both alternatives would have been tested.

36. The '392 Patent teaches that one can accurately determine, from a small amount of genetic material, and “noisy” genetic data produced therefrom, the number of copies of a chromosome of interest in the genome of, for example, an embryo or fetus, despite the low quality of that data. *See, e.g., id.*, 2:1-4 (explaining that “the genotyping of the fetal tissue with high fidelity is not currently possible”). The '392 Patent teaches that its methods enable such a determination through the use of detailed statistical methods which permit a person of ordinary skill in the art to “detect and correct some or all of the[] errors” associated with analyzing a small amount of genetic material. *Id.*, 59:38-39.

V. PROGENITY’S ASSERTED REFERENCES

37. I discuss below the references asserted by Progenity. For each reference, I will first discuss the analytical methods disclosed by the reference and then issues a skilled artisan would consider in order to apply the method to fetal DNA in a maternal blood sample to detect fetal chromosomal abnormalities.

A. Shimkets

38. Shimkets refer to International Publication Number WO 2005/039389 to Shimkets *et al.*, filed as International Application PCT/US2004/034890 on October 22, 2004. Ex. 1006, cover (10), (21) and (22). Shimkets is titled “Sequence-Based Karyotyping.” *Id.* at 1:1. It was published on May 6, 2005. *Id.*, cover (43). Shimkets describes a sequence-based method for “the determination of karyotypes

of genomes of individual[] cells and organisms.” Ex. 1006, 1:7-8 (Field of the Invention).

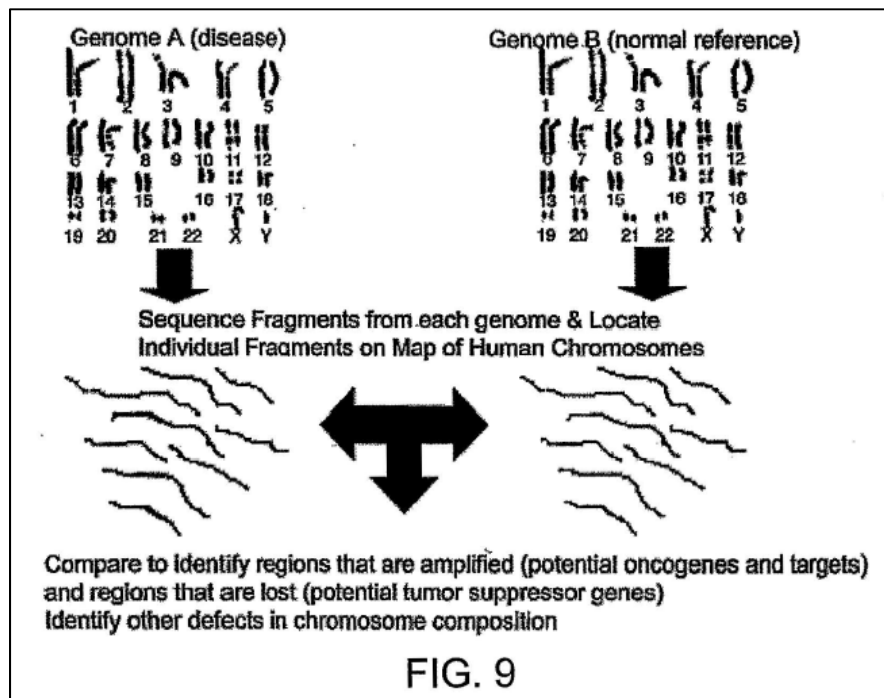
39. Karyotyping is a well-established process by which *chromosomes from a cell* are visualized under a microscope and analyzed (generally by taking photographs and analyzing the picture) to determine the chromosome complement of an individual, including the number of chromosomes and any apparent abnormalities. *See* Ex. 1006, 7:12-17 (“The term ‘karyotype’ refers to the genomic characteristics of an individual cell or cell line of a given species, e.g., as defined by both the number and morphology of the chromosomes.”). As was generally understood at the time of the invention and as is consistent with Shimkets’s definition of “karyotype,” Shimkets’s Sequence-Based Karyotyping method is performed on DNA extracted from a cell instead of on cell-free DNA such as cell-free fetal DNA circulating in maternal blood. Because the Karyotyping method is performed on DNA extracted from a cell, the sample used for testing and measurements is a DNA sample of a single genetic source, and not a sample of mixed genetic sources. This differs from maternal blood samples having both fetal and maternal DNA.

40. The brief invention summary of the application provides a high-level description of the method (Ex. 1006, 2:20-33):

The current invention provides for a method of karyotyping a genome of a test cell (e.g., eukaryotic or prokaryotic) by generating a pool of fragments of genomic DNA by a random fragmentation method, determining the DNA sequence of at least 20 base pairs of each fragment, mapping the fragments to the genomic scaffold of the organism, and comparing the distribution of the fragments relative to a reference genome or relative to the distribution expected by chance. The number of a plurality of sequences mapping within a given window in the population is compared to the number of said plurality of sequences expected to have been sampled within that window or to the number determined to be present in a karyotypically normal genome of the species of the cell. A difference in the number of the plurality of sequences within the window present in the population from the number calculated to be present in the genome of the cell indicates a karyotypic abnormality.

41. As described above and in various examples, Shimkets's Sequence-Based Karyotyping method generally consists of the following steps. *See* Ex. 1006, 2:20-33, 22:23-30:5. First, DNA from a cell is isolated. The DNA is then randomly fragmented using mechanical methods such as sonication or fluid shearing or enzymatically using DNase I. The ends of the DNA are then "polished" using Pfu DNA polymerase. Synthetic DNA adaptors are then ligated to the DNA ends; these adaptors include "universal" sequences for PCR amplification or DNA sequencing. DNA segments that are likely to have adaptors, and falling into an

appropriate size range, are isolated and treated with modifying enzymes, including but not limited to polymerase, ligase and kinase, to repair “nicks” in the DNA. These fragments are then captured, amplified, and sequenced to a length of 20 bases or more. The sequence data from the test cell is mapped to the genome and grouped into windows across the genome to generate a test distribution map. The depth of sequence within those windows is measured and compared to data from a control cell or to reference sequence data to determine if there are regions either over-represented or under-represented, possibly indicating the presence or absence of copy number variants, including aneuploidy. Shimkets provides a schematic depiction of the overview of the Sequence-Based Karyotyping in Figure 9. Ex. 1006, 5:18.



42. Concerning the data analysis, Shimkets state that “[t]he plurality of DNA sequences obtained are mapped to a genomic scaffold to create a distribution of mapped sequence to a region of the genome.” *Id.*, 3:14-15. Shimkets defines the test cell distribution (described as a chromosomal map density) “as the number of mapped sequences (fragments) by the number of possible map locations present in a given chromosome.” *Id.*, 3:22-24. “The number of possible map locations is defined by the size of the observation window and the length of the chromosome.” *Id.*, 3:24-25. This test distribution is compared to a reference distribution from a reference cell and differences between the test distribution and the reference distribution are identified. The reference distribution can be “a database of mapped sequences from previously tested cells.” *Id.*, 3:28-30. In this comparison, the amounts of genetic material (DNA) are not measured at a particular locus of a chromosome or chromosome segment. Rather, the sequences are binned according to the observation windows and the aggregate numbers of sequences for the test sample and reference in a particular observation window are compared.

43. Note that not all measured data is used in creating the test distribution and by extension the reference distribution:

Sequences that did not map to the genome, either due to incompleteness of the genomic scaffold or issues of sequencing quality, were removed from considering. Filtering was also performed

to remove DNA sequences which mapped to multiple genomic locations (within repeated sequences).

Id., 58:23-26; *see also* 17:22-24 (“Filtering is performed to remove DNA sequences within repeated regions and to remove the rare DNA sequences not present in the human genome.”); 62:25-31 (explaining removal of sequence reads), 14:22-24 (“A reference distribution is produced by applying the same method used to produce the test distribution with the exception that the DNA molecule that is subjected to Sequence-Based Karyotyping is from a reference cell.”).

44. As a result of the data filtering, a significant percentage of the sequence reads is not mapped to genomic scaffold for creating test/reference distribution. In the only actual experiment described in Shimkets, the authors generated 354,451 total sequence reads for the reference sample GM12911 and 487,310 total sequence reads for the test sample DiFi.² Of these, using the criteria “of at least 95% identity over 90% of the read length” and that the sequences be not “mapped to more than one position,” only 125,684 reads for the GM12911 normal cell and 203,352 reads for the DiFi cells were uniquely mapped to the Genome. *Id.*, 61:14-17. That means

² DiFi is a rectal carcinoma cell line whose karyotype is tetraploid, contains many marker chromosomes (small chromosomal fragments), and shows numerous episomal particles (additional DNA fragments). *See* Ex. 2018, Abstract. This makes DiFi a poor candidate for one exploring whether a cell has a normal complement of chromosomes. *See* discussions in paragraphs 48-50.

only 35.5% of the reference sequences and 41.7% of the test sequences were used in the analysis (or a discard rate of 64.5% and 58.3% respectively for the reference and test cells).

45. Shimkets measures the karyotypic difference between the test cell and the reference cell by a difference in the mapped sequence distributions between the test distribution and the reference distribution. *Id.*, 3:31-33. Shimkets identifies an aberration if the difference between the test and reference distributions is deemed statistically significant, as measured by a p-value below a threshold. *Id.*, 4:1-7. Specifically, Shimkets states that “[s]tatistical significance is determined by method known in the art. An alteration is statistically significant if the p-value is at least 0.05. The p-value[] is a measure of probability that a difference between groups during an experiment happened by chance. ($P(z \geq z_{\text{observed}})$).” *Id.*, 4:1-3.

46. A p-value in this context only indicates the probability that the difference between two sets of data occurred by chance. In and by itself, the p-value does not indicate the probability of the presence or the probability of the absence of a chromosomal abnormality in the test sample. As one prior karyotyping paper warned, the limiting feature of the karyotyping analysis “was not sensitivity for detecting the alteration” in a cell sample. Ex. 2012 [Wang] at 16158, Table 1 notes. “What was of more concern was the positive predictive value (PPV), that is, the probability that a detected mutation represents a real mutation.” *Id.* The Wang paper

is the digital karyotyping paper referenced by Shimkets. Ex. 1006, 59:8-10 & Fig. 1, 107:5-6 (Wang is reference # 17).

47. Shimkets does not appear to have a high sensitivity to copy number changes. For example, in comparing DNA from a colon cancer cell line (DiFi) to a normal cell line (GM12911), both of which would provide large quantities of DNA from a single genetic source (“pure” DNA), Shimkets detects (or calls) aneuploidy “when the test distribution to reference distribution is greater than 1.5 or less than 0.75.” Ex. 1006, 4:8-10, 16:30-17:4. Thus, in the only experiment it performed, there were instances where the measured test to reference distribution was 25% higher than expected for a diploid chromosome, but no aneuploidy call was made.³ See Fig. 1 in which the data points in the red box are labeled “Intermediate Content” and for which no aneuploidy call was made; these are likely chromosomes known to be diploid and the decision not to make a call may indicate a bias to match the known result. Further, Shimkets provides no data as to whether calls could be made at all if the DiFi DNA were diluted into a significantly larger quantity of GM12911 DNA to mimic a test condition more similar to maternal blood samples having both fetal and maternal DNA.

³ These results are difficult to interpret since the chromosomal content of DiFi is aberrant for every chromosome and the characterization of DiFi presented by Olive is inconsistent with the results presented in Shmkets.

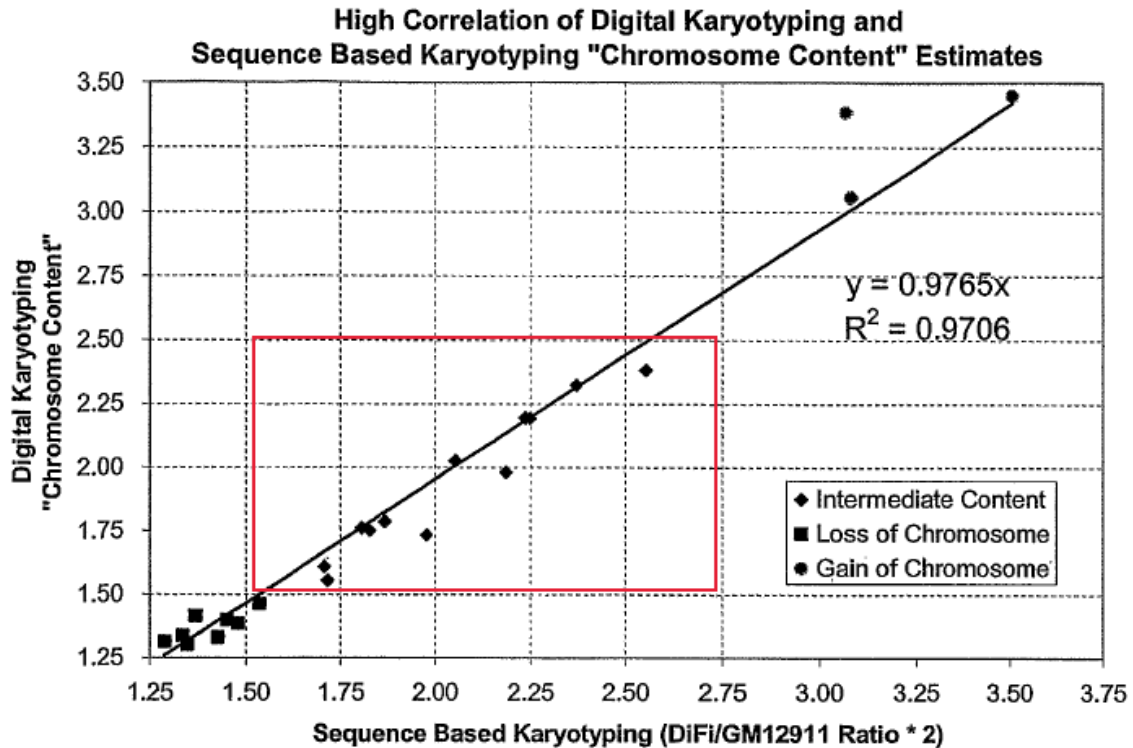


Figure 1

48. And a POSITA would conclude it cannot: Shimkets reliance on analysis of the DiFi cell line makes interpretation of any of his results challenging at best. DiFi is a rectal carcinoma cell line reported by Olive and her colleagues aberrant for (nearly) every chromosome. Ex. 2018. The cell line has a tetraploid karyotype (meaning four copies of each chromosome instead of the usual two) based on flow cytometry data indicating the presence of more than double the usual number of chromosomes. *Id.*, Abstract. It is also reported to contain many marker chromosomes (small chromosomal fragments), and shows numerous episomal particles (additional independent DNA fragments). *Id.* All of these make DiFi a

poor candidate for one exploring whether a cell has a normal complement of chromosomes.

49. Olive's estimates of ploidy are based on the observation that the genome is highly rearranged with unclear numbers of chromosomal copies but with likely none of the chromosomes demonstrating euploidy:

Consistent with the flow cytometry evaluation of Passages 1 and 10, the karyotype of DiFi cells at Passage already showed a predominance (>50%) of near tetraploid (Fig. 5). Many marker chromosomes were present. The loss copies of chromosomes 5, 9, 18, 21, and X was noted without concomitant appearance of these in marker chromosomes. Trisomies were observed for chromosomes 6, 19, and 20. Chromosomes 7 and 2 contained additional segments of their respective chromosomes that were fused to form m3 and m4, respectively. In addition, two copies (of the four expected) of chromosome 14 absent and seemed to form the m1O [iso(14q)] marker. For chromosomes 12 and 17 only a single normal copy was present, and copy of chromosome 8 present seemed to be a derivative of normal form [der(8)t(8;?) (p22;?)]. The long arm of chromosome seemed to be involved in the formation of marker m1I [iso(Chromosome 12 was present as three copies of the three copies of the m7 marker [t(12;13)], and two copies of chromosome 17 were involved in the formation of the m14 marker. In addition, the DiFi line demonstrated a fairly high frequency of episomal particles (data not shown).

Id. at 243-44.

50. This peculiar number of chromosomes calls into question Shimkets analysis as no chromosomes are identified at having four copies (twice that expected). *See* EX1006, Fig. 1. Although Oliver reports reduced numbers for some chromosomes in a “normal” form, the apparently chromosomes seem to form small “marker chromosomes” by fusing their content with fragments from other chromosomes. If we are to take Olive’s report of trisomies for three chromosomes (6, 19, and 20), then one has to question whether Shimkets’s method is reliable as Shimkets only found amplification on chromosomes 7, 13, 20 and some segments of chromosome 5, 10, 14 and 18 (Ex. 1006, 59:24-29). For chromosomes 6 and 19, Shimkets appeared to have detected some amplification, but did not make the aneuploidy call. Ex. 2012 [Wang], Table 2 (2X amplification for chromosomes 6 and 19 are 2.4 and 2.2 respectively); Ex. 1006, Fig. 1 (2X amplification for data points at 2.4 and 2.2 are ~2.55 and 2.25, respectively). This means that Shimkets’s method is not sufficiently sensitive or accurate. These failures for the analysis of “pure” DNA from the DiFi cell line raises serious doubts that this method could be employed in a situation in which the aberrant chromosomal signal is diluted by the presence of a large excess of DNA from cells with “normal” numbers of chromosomes. Given the clear failure of this method to identify known aberrations

in chromosomal copy number, there would be no reason a POSITA would consider this a viable method for applications to cell-free DNA analysis.

51. It should be noted that Shimkets largely provides laboratory protocols for the generation of data but provides no details on the analytical methods to be used. But in the only experiment that it performed, it seems that Shimkets is trying to replicate the analysis performed by Wang (reference 17 in Shimkets). For example, it plotted its determined DIFI-to-GM12911 ratio against Wang's test on the same species using digital karyotyping; and it included Wang's test data as insets for its Figures 2 and 3. *Compare* Ex. 1006 Figs. 2 & 3 insets with Ex. 2012 [Wang], Fig. 2 plots for Chr2 and Chr 7. Even their data exclusion rules are similar:

Shimkets	Wang
<p>“Filtering is performed to remove DNA sequences within repeated regions and to remove the rare DNA sequences not present in the human genome.”</p> <p>“Sequences that did not map to the genome, either due to incompleteness of the genomic scaffold or issues of sequencing quality, were removed from consideration. Filtering was also performed to remove DNA sequences which mapped to multiple genomic locations (within repeated sequences).”</p> <p>Ex. 1006, 17:22-24, 58:23-26.</p>	<p>The experimental tag with the same sequence as virtual tags were termed filtered tags and were used for subsequent analysis. The remaining tags corresponded to repeated regions, sequences not present in the current genome database release, polymorphisms at the tag site, or sequencing errors in the tags or in the genome sequence database.</p> <p>Ex. 2012 at 16157.</p>

52. One feature of Shimkets's karyotyping method is "random" fragmentation of DNA material from the test cell. *See* Ex. 1006, 2:20-25 ("generating a pool of fragments of genomic DNA by a random fragmentation method"); 3:1-3 ("The genome of the cell is karyotyped by randomly fragmenting the DNA from a cell and a sequencing at least a portion of each fragment."); 9:18-20 ("It is a discovery of the present inventors that the genome of an organism can be sampled by random fragmentation and sample sequencing to determine karyotypic properties of a cell, tissue, or organism"). According to Shimkets, thanks to this random fragmentation feature, its sequence-based karyotyping method is "limited in resolution only by the number of fragments an operator wishes to sequence, rather than a systematic limitation imposed by the method of sequence fragmentation." *Id.*, 19:15-18.

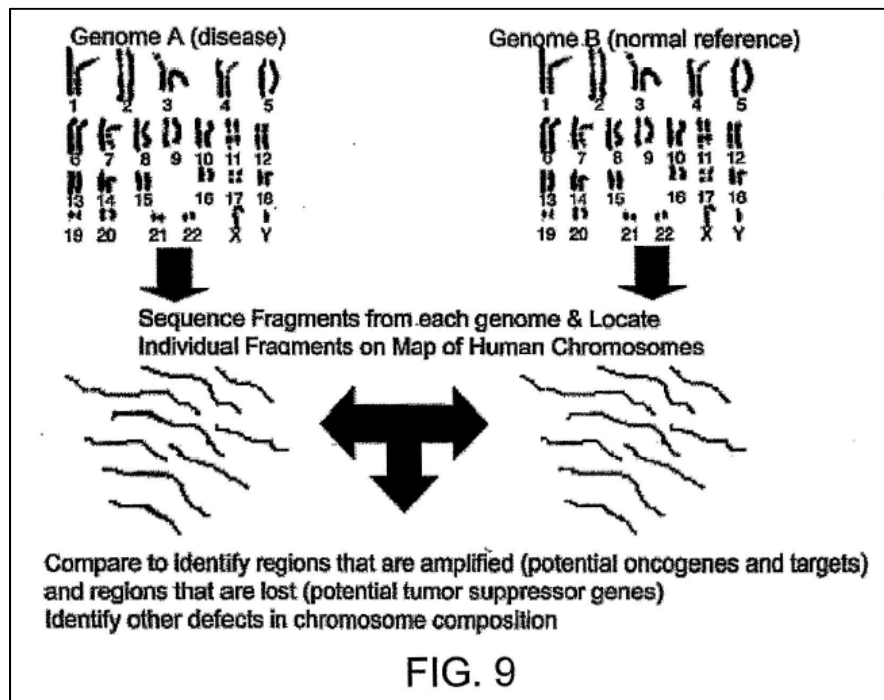
53. Shimkets explains that "[r]andomly fragmenting a DNA refers to the physical fragmentation (e.g., also called breakage or digestion) of a large molecule DNA into multiple smaller DNA molecules in a non-sequence specific manner." *Id.*, 11:18-20. Shimkets refers to "random fragmentation" also as "non-sequence specific fragmentation" and distinguishes it "from sequence specific fragmentation which may involve, for example, restriction endonuclease digestion." *Id.*, 11:20-22.

54. Shimkets' sequence-based karyotyping is a method for analyzing "test" cells. Ex. 1006, 2:20-22 ("method of karyotyping a genome of a test *cell*"); 3:1-3 ("The genome of *the cell* is karyotyped"); 3:11-12 ("The cell is a cancer cell, an embryonic cell, or a fetal cell"); 60:24-28 (performing sequence-based karyotyping "on DNA from the DiFi colorectal cancer cell line, and from lymphoblastoid cells of a normal individual"). For example, Shimkets describes comparing the data from a test cell and a reference cell from the same species (*id.*, 3:9) indicating that the method requires a test and control cell or at the very least data from control cells. DNA extracted from a single test cell or a single cell line would be understood as containing genomic material from a single source. For example, a skilled artisan would understand that a test sample derived from the DiFi cell line is to contain only the DNA of the specific cancer cell line. Similarly, a skilled artisan would understand that a reference sample obtained from lymphoblastoid cells of a single normal individual is a sample containing only normal cells with only genomic material from the individual and therefore would be expected to have two copies of each autosome and either XX or XY sex chromosomes in each cell.

55. As examples, Shimkets's authors describe their test cells as belonging to "cancer cell, an embryonic cell, or a fetal cell." *Id.*, 3:11. For fetal cells, the authors tell us that "[t]he cell is isolated from amniotic fluid or is derived from in

vitro fertilization.” *Id.*, 3:13. A skilled artisan would understand fetal cells isolated from amniotic fluid as fetal cells free of maternal DNA material and that they provide a sample with only the fetal DNA and not a sample with also the mother’s DNA. Similarly, a fetal cell derived from *in vitro* fertilization (or a pre-implantation embryo) would be understood by a skilled artisan as a cell having only the fetal DNA and not one also having DNA from the mother, because an embryo is commonly viewed as containing only the fetus’s genetic material.

56. Shimkets does not state that its karyotyping method is suitable for sequencing cell-free DNA or DNA from different cell lines with different genotypes. That is, the Sequence-Based Karyotyping described in Shimkets are directed toward “pure” samples rather than mixed samples such as those that occur in non-invasive prenatal testing (NIPT) analysis. This can be seen, for example, from the overview illustration below in which “pure” Genome A sample and “pure” Genome B reference are separately sequenced, analyzed and compared. Ex. 1006, Fig. 9. This can also be seen from the test cell examples I discussed above where tests are performed on individual cells or cell lines containing only genetic material from only a single source.



57. In addition to Sequence-Based Karyotyping, Shimkets also refers to five other methods without providing any needed details on how they would be implement, one of which is the “complex sample sequencing” method. Ex. 1006, 18:6-10 (listing six methods including Sequence-Based Karyotyping and complex sample sequencing), 5:16-26 (listing figures for the six different methods). Sequence-based karyotyping and complex-sample sequencing are the only two methods that Shimkets suggests are suitable for diagnosis. *Id.*, Figure 4B (listing the two methods under “DIAGNOSTICS”, reproduced below). Given these disclosures, a skilled artisan would understand that complex-sample sequencing and Sequence-Based Karyotyping are two different types of methods.

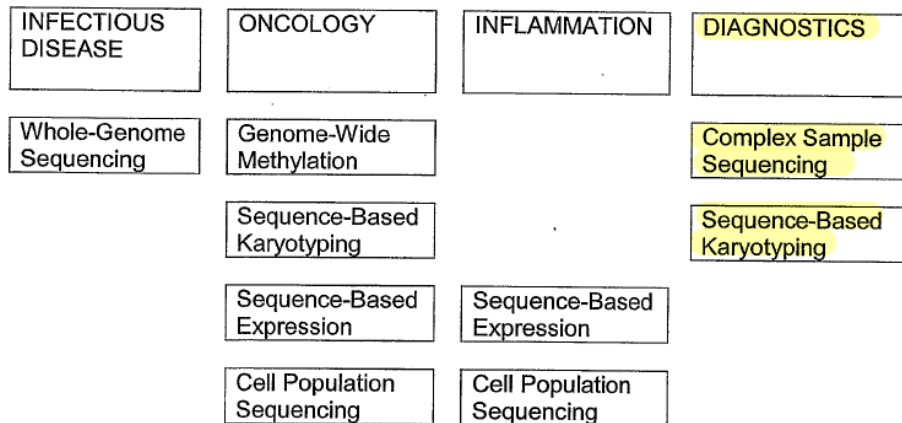


FIG. 4B

58. Unlike Sequence-Based Karyotyping, complex sample sequencing is only described as a qualitative test “for detection of pathogens in blood, water, air, soil, food” etc. *Id.*, 21:10-12, FIG. 12. That detection involves “preparing a mixed DNA and cDNA sample” and “sequencing random fragments from the DNA and RNA in the sample.” *Id.*, 21:12-15. BLAST is used to “assign sequences to known genomes for pathogen.” *Id.*, 21:17-18. Shimkets also suggests organizing the pathogens “into an evolutionary tree to indicate known agents and/or new agents or strains (e.g., virus or bacteria).” *Id.*, 21:18-20. One advantage of the complex sample sequencing method is its supposed ability “to identify unknown pathogenic agents and microorganisms” and “for difficult diagnostic cases, such as the detection of *M. tuberculosis*.” *Id.*, 2:19-24.

59. Shimkets does not suggest using complex-sample sequencing for quantitative determinations such as its test-to-reference distribution-based

aneuploidy detection or other chromosomal copy number changes. Ex.1006, 21:10-24. That is, the complex-sample sequencing method qualitatively determines the identity of species to which the sequences may belong, but does not describe the quantitative assessment the number of individuals of each species present. Shimkets also does not suggest using the method for identifying more complex organisms, such as a mixture of DNA from different mammals. Instead, complex-sample sequencing targets simpler live forms like bacteria, virus and microorganisms. *Id.*, 21:10-24.

60. Having described the operation and features of Shimkets, I will now turn to obstacles that a skilled artisan would have faced if they were to apply Shimkets to either cell-free fetal DNA circulating in maternal DNA or DNA extracted from fetal cells harvested from maternal blood.

1. Complications Associated With Applying Shimkets To Cell-Free Fetal DNA In A Maternal Blood Sample

(a) Complex-sample sequencing is not suitable for quantitative testing

61. I understand that Progenity states that “Shimkets suggested applying its sequencing-based karyotyping methodology to a cell-free blood sample.” IPR2021-00266 Pet. 19, 20. It is apparent that Progenity has mistaken Shimkets’s disclosures related to complex-sample sequencing for sequence-based karyotyping. For

example, the characteristics that Progenity imputes to the method are described in connection with the complex-sample sequencing method described in 21:10-24:

Petition	Shimkets
<p>“Shimkets teaches its methodology of measuring up to one million or more loci could be applied to detect an unknown organism in a blood sample.” IPR2021-00266 Pet. 22.</p> <p>“Shimkets suggests applying the invention to a ‘complex sample’ to identify ‘all organisms in a sample without any prior knowledge,’ including by generating one million sequence reads for a cell-free blood sample.” IPR2021-00266 Pet. 28</p>	<p>“Complex sample sequencing in accordance with the invention can be used for detection of pathogens in blood ..., and for identification of all organisms in a sample without any prior knowledge. ... According to one embodiment, a cell-free sample (e.g., blood, water, air, food or soil) can be used to generate 1 million sequence reads.” Ex. 1006, 21:10-17</p>

62. The complex-sample sequencing method described in 21:10-24, however, as I have noted above, is a qualitative test, not a quantitative test. Specifically, identifying a pathogen or microorganism involves a qualitative determination of the identity of the DNA sequence(s), but not the quantity of the sequence(s)). The latter, e.g., the frequency of DNA sequences in a test sample compared to a reference, is needed to detect a change in chromosomal copy numbers (based on Progenity’s mapping of the elements for the various claimed steps). Thus, even if complex-sample sequencing could be applied to more complicated life forms to identify their species, the most it could do, when applied to a maternal blood sample containing a mixture of fetal and maternal DNA, was

to identify the presence of two individuals (as in a forensic sample) or the presence of a male fetus (and any pathogens that might be in the blood). Complex sample sequencing method, however, is not described as capable of being able to perform any quantitative determination, not even the relative fractions of fetal and maternal DNA content. Shimkets only associates Sequence-Based Karyotyping with aneuploidy detection and copy number change detection. EX1006, 16:27-28 (“Sequence-Based Karyotyping methods of the invention may be used to determine aneuploidy or copy number polymorphisms”).

63. I have also been asked to review Shimkets’s disclosures in 3:9-13, 3:14-20, 7:22-8:1, 16:14-17:4 to see whether they associate aneuploidy detection or copy number change detection with complex-sample sequence or testing using fetal DNA derived from blood samples. I do not see any such association.

64. In particular, the disclosure in 3:9-13 discusses the source of the test cells. EX1006, 3:9-13. For fetal cells, the passage does not mention fetal cells from blood as a potential source. Instead, it only suggests using fetal cells “isolated from amniotic fluid” or “derived from *in vitro* fertilization.” *Id.* Both amniotic fluid and pre-implanted embryos are sources of pure fetal DNA. The disclosure in 3:14-20 discusses mapping test DNA sequences “to a genomic scaffold to create a distribution of mapped sequence to a region of the genome” and that “1,000,000 or

more sequences” can be mapped. *Id.*, 3:14-20. This paragraph, however, does not mention the sample source. *Id.*

65. The disclosure in 7:28-8:1 discusses that “[c]hromosome abnormalities may be correlated with the presence of a pathological conditions... or with a predisposition to developing a pathological condition.” *Id.*, 7:28-8:1. The passage also does not discuss the source of the sample.

66. The paragraph in 16:14-19 discusses applying the sequence-based karyotype method “on a cancer cell or disease cell,” and does not mention using blood as a source for test samples. *Id.*, 16:14-19. The next paragraph discusses the source of the test cells and the potential application of the method to prenatal diagnosis of “embryo or fetus.” *Id.*, 16:20-26. The paragraph, however, only suggests using “a sample from amniotic fluid collection procedure or from a chorionic villus sampling procedure.” *Id.* A skilled artisan would understand that these procedures do not involve collecting DNA from a maternal blood sample. Further, both procedures collect “pure” fetal cells and not a mixed sample having both maternal and fetal DNA. The paragraph in 16:26-17:4 discusses applying sequence-based karyotyping for aneuploidy detection, but like previous passages, also does not disclose blood samples as a source of the genetic materials for this quantitative determination. *Id.*, 16:26-17:4.

(b) Shimkets uses its sequence-based karyotyping on cells and not samples of mixed genotypes

67. As I have discussed above, Shimkets' Sequence-Based Karyotyping operates on DNA extracted from cells and that the extracted DNA is expected to be the DNA only of the test cell and not of other sources. ¶¶ 54-56. For example, fetal cells isolated from amniotic fluid or derived from an embryo fertilized *in vitro* would be understood by a skilled artisan as containing only fetal DNA and not maternal DNA. ¶ 55.

68. Sequence-Based Karyotyping is described as suitable for analysis of pure DNA samples (*i.e.*, samples containing DNA from a single source) that can be compared to each other and is not suitable for detecting copy number changes of a target DNA when the target DNA counts for only a small fraction of the total DNA material in the sample to be analyzed. Shimkets relies on comparing sequence distribution density of DNA from a test cell to a reference cell so chromosomal trisomy compared to a euploid chromosome would have a ratio of 3:2, or 1.5. Shimkets recognizes that its method is not sensitive enough to reliably identify changes below a stringent threshold and in fact is not consistent with the characterization by Olive that there are three trisomic chromosomes in the DiFi cell line. For example, the method calls aneuploidy "when the test distribution to reference distribution is greater than 1.5 or less than 0.75." Ex. 1006, 4:8-10, 16:30-

17:4. That is, abnormality is only called if the observed deviation is 50% higher than that expected for a diploid cell. The arguments presented by Progenity never bridge the gap between cell lines that produce nearly unlimited quantities of pure DNA and cfDNA which is available in very limited quantities and in which the target DNA is highly diluted in the background of normal DNA.

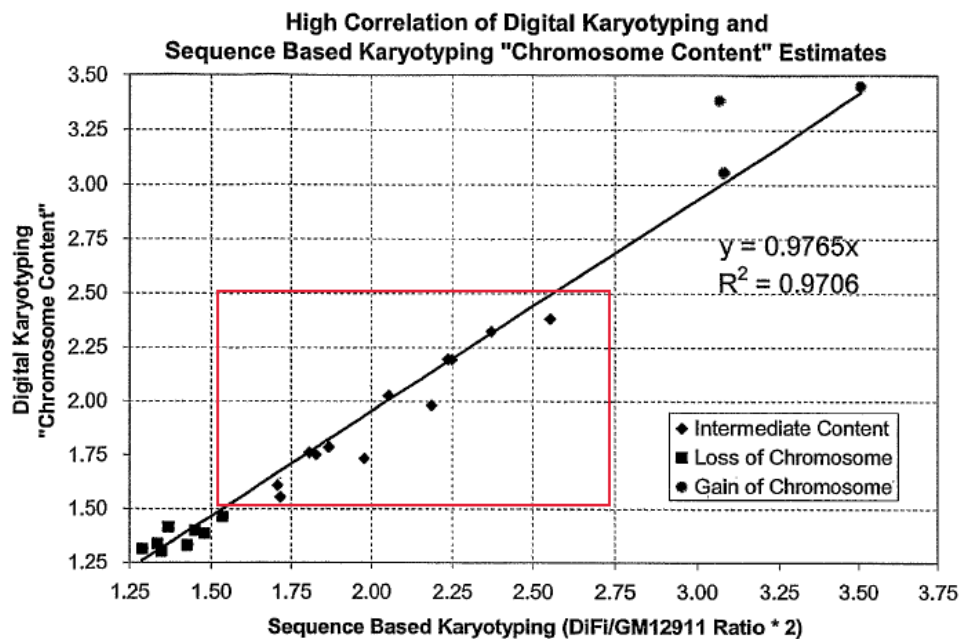


Figure 1

69. As Progenity and its expert acknowledge, a maternal DNA sample would contain fetal "DNA of limited abundance from a complex background." IPR2021-00266 Pet. 29. For example, Bianchi reports the following fetal DNA concentrations:

	Mean fetal fractions in serum (range)	Mean fetal fractions in plasma (range)
early pregnancy	0.13% (0.014-0.54%)	3.4% (0.39-11.9%)
late pregnancy	1.0% (0.032-3.97%)	6.2% (2.33-11.4%)

Ex. 1008, S94.

70. A technical brief by Wei of unknown date submitted by Progenity reports a median fetal fraction of 11.8% for 19 male pregnancy plasma samples. Ex. 1026 at 337-338 and Table. 1. The fetal DNA fractions ranged between 1.98% and 47.69%. Fifteen (15) of the nineteen (19) samples had fetal DNA fractions below 20%, three had fetal DNA fractions between 20-30% and one had an unusually high fetal fraction at 47.69%.

Table 1. Patient population.

Case	Mat ^a age, years	Gravida ^b	Para ^c	EGA, weeks	Indication	Amnio or CVS	Karyotype	Copies/mL		Fetal DNA, %
								β -actin	SRY	
1	36	4	3	17.1	AMA	Amnio	46,XY	1140	22.6	1.98
2	38	3	0	15.8	AMA	Amnio	46,XY	2990	38.4	1.28
3	34	1	0	15.2	AMA	Amnio	46,XY	860	40.1	4.66
4	35	2	1	15.4	AMA	Amnio	46,XX	680	0.0	
5	39	2	1	19.8	AFP+	Amnio	46,XX	640	0.0	
6	34	3	1	9.8	PA	CVS	46,XX	1190	0.0	
7	39	4	1	18.0	AFP+	Amnio	46,XY	1390	96.5	6.94
8	35	2	1	15.4	AMA	Amnio	46,XY	990	69.7	7.04
9	32	2	0	16.0	AFP+	Amnio	46,XX	840	0.0	
10	37	4	0	20.0	AMA	Amnio	47,XY,+21	880	256.5	29.15
11	40	2	1	11.3	AMA	CVS	46,XX	560	0.0	
12	29	1	0	29.3	U/S	Amnio	46,XY	1190	231.1	19.42
13	42	2	1	15.0	AMA	Amnio	46,XY	610	290.9	47.69
14	40	1	0	11.2	AMA	CVS	46,XY	930	94.4	10.15
15	39	3	1	11.2	AMA	CVS	46,XY	660	38.5	5.83
16	32	6	3	19.1	AFP+	Amnio	46,XY	760	125.5	16.51
17	38	1	0	16.1	AMA	Amnio	46,XX	490	0.0	
18	37	13	9	19.3	AFP+	Amnio	46,XX	880	0.0	
19	30	4	2	19.9	AFP+	Amnio	46,XY	670	81.2	12.12
20	34	2	0	17.6	AMA	Amnio	46,XY	510	91.1	17.86
21	23	2	1	23.4	U/S	Amnio	46,XY	510	57.4	11.25
22	39	3	1	15.0	AMA	Amnio	46,XX	940	0.0	
23	39	3	1	16.7	AMA	Amnio	46,XY	490	100.0	20.41
24	41	4	2	17.5	AMA	Amnio	46,XX	1080	0.0	
25	36	3	2	16.5	AMA	Amnio	46,XY	410	13.8	3.37
26	21	1	0	21.5	AFP+	Amnio	46,XX	400	0.0	
27	42	1	0	10.9	AMA	CVS	46,XY	880	103.6	11.77
28	37	3	2	17.1	AMA	Amnio	46,XX	500	0	
29	34	3	2	17.5	AFP+	Amnio	46,XY	830	189.1	22.78
30	30	6	3	19.0	AFP+	Amnio	46,XY	1260	164.5	13.06
Mean								872	111	
SD								474	78	

^a Mat, maternal; EGA, estimated gestational age; AMA, advanced maternal age (≥ 35 years at due date); AFP, α -fetoprotein; PA, previous aneuploid pregnancy (i.e., abnormal number of chromosomes); U/S, ultrasound.
^b Number of pregnancies.
^c Number of times delivered baby of ≥ 20 weeks.

71. The above data indicates (1) the fetal DNA fractions vary widely from sample to sample and (2) fetal fractions generally are only a small fraction of the total DNA material circulating in maternal blood.

72. At a 10% fetal fraction, a fetus with trisomy would theoretically cause the number of copies of chromosome to increase from 2.0 to 2.1 (90% maternal fraction \times 2 copies + 10% fetal fraction \times 3 copies). The test distribution to reference distribution would be about $2.1 \div 2 = 1.05$ in this case. At 20% fetal

fraction, the expected number of copies is 2.2 (80% maternal fraction \times 2 copies + 20% fetal fraction \times 3 copies) with a ratio of test distribution to reference distribution of $2.2 \div 2 = 1.1$. At 50%, that ratio would only increase to 1.25 ((50% maternal fraction \times 2 copies + 50% fetal fraction \times 3 copies) \div 2 = 1.25). Yet, all these values would fall under the Shimkets's threshold of aneuploidy call, which is "when the test distribution to reference distribution is greater than 1.5 or less than 0.75." EX1006, 4:8-12, 16:30-32.⁴ Given the threshold set by Shimkets, none of the expected ratios would be deemed to be indicative of aneuploidy: For instance, even for a test to reference distribution ratio of 1.25, Shimkets did not regard that value as indicating a gain in chromosome. *See* Ex. 1006, Fig. 1. Shimkets results are also inconsistent with characterization of the DiFi karyotype and cast further doubt on the efficacy of the method as discussed in paragraphs 48-50 above.

⁴ Indeed, even at the unrealistic 95% fetal fraction, that ratio would still be below 1.5 (5% maternal fraction \times 2 copies + 95% fetal fraction \times 3 copies) \div 2 = 1.475).

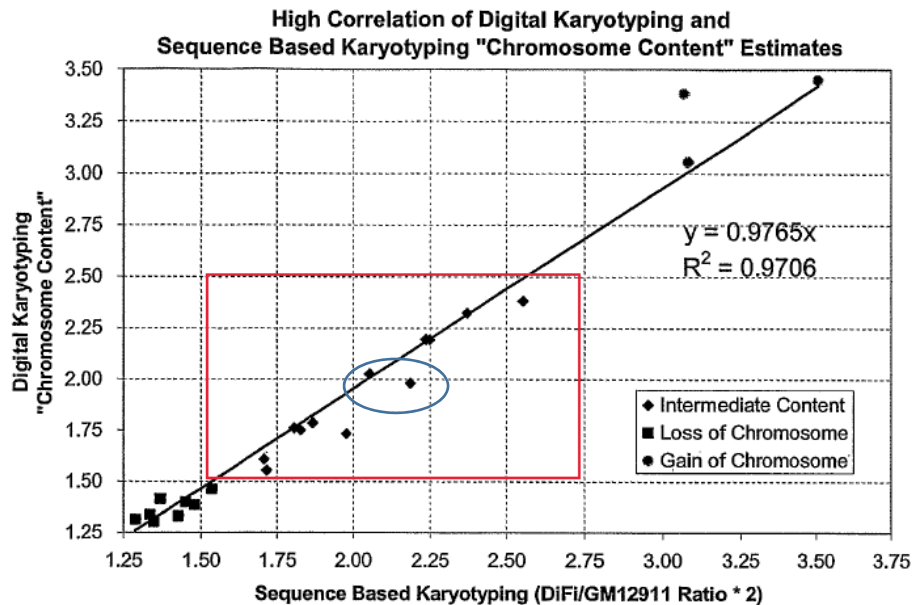


Figure 1

73. Theoretically, the threshold should be set as a function of the fetal fraction. But here is the dilemma: the actual fractions would not be known before the assay could be run and those measurements would themselves be determined experimentally subject to error. As an example, Chiu studied the effects of sample preparation on detected SRY concentrations (a gene on the fetal Y chromosome) and β -globin concentrations (a gene on chromosome 11, which is never aneuploid). The ratio of the two is typically taken as the fraction of the (male) fetal DNA material since only males carry a Y chromosome. Even with the optimized procedure, there was still significant data scattering in both concentrations for the same sample. *See* Ex. 2019 [Chiu].

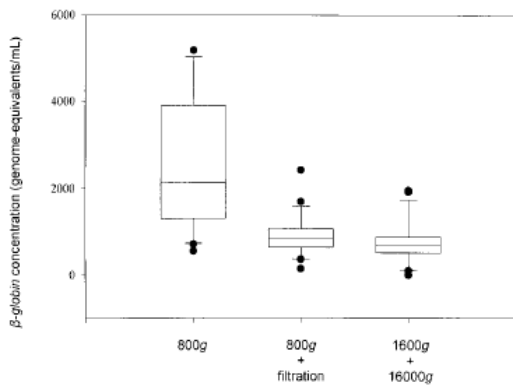


Fig. 3. Box plot of β -globin concentration of plasma aliquots prepared by protocols (as indicated on the y axis) for module 2. 800g and 1600g denote the speeds of centrifugation, and 16 000g denotes the speed of microcentrifugation.

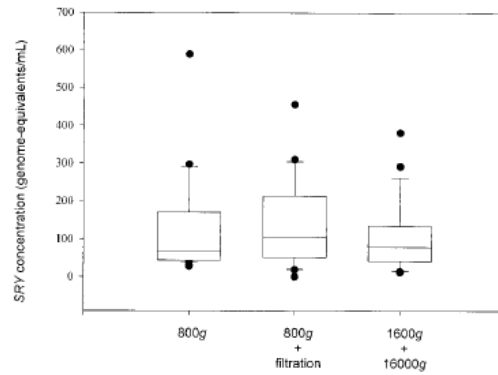


Fig. 4. Box plot of SRY concentration for module 2 presented in the same format as Fig. 3.

74. Generally, when $f = \frac{A}{B}$, the standard deviation in f is

$$\sigma_f \approx |f| \sqrt{\left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2 - 2\frac{\sigma_{AB}}{AB}}$$

with σ_A and σ_B corresponding to the expected values A and B's standard deviation and σ_{AB} is the covariance or standard deviation for the multiplication of A and B.

(In the case of SRY and β -globin concentrations, A and B would be the median or mean of SRY and β -globin concentrations). Given the unsymmetrical scattering especially in the measured fetal DNA amounts, the error in the determined value of fetal DNA fraction may be over 100%. Progenity and Dr. Ding did not consider the difficulty in obtaining a reasonably accurate and certain fetal DNA fraction at the time of the invention and how that uncertainty affects the experimental results and analysis.

75. In addition to the uncertainty in fetal fractions, there is also uncertainty in the detection method itself. As Shimkets's Figure 1 shows, the two measured ratios in the blue circle may all be around 1.0 (the values on the x- and y-axis are both after a multiplication by 2).

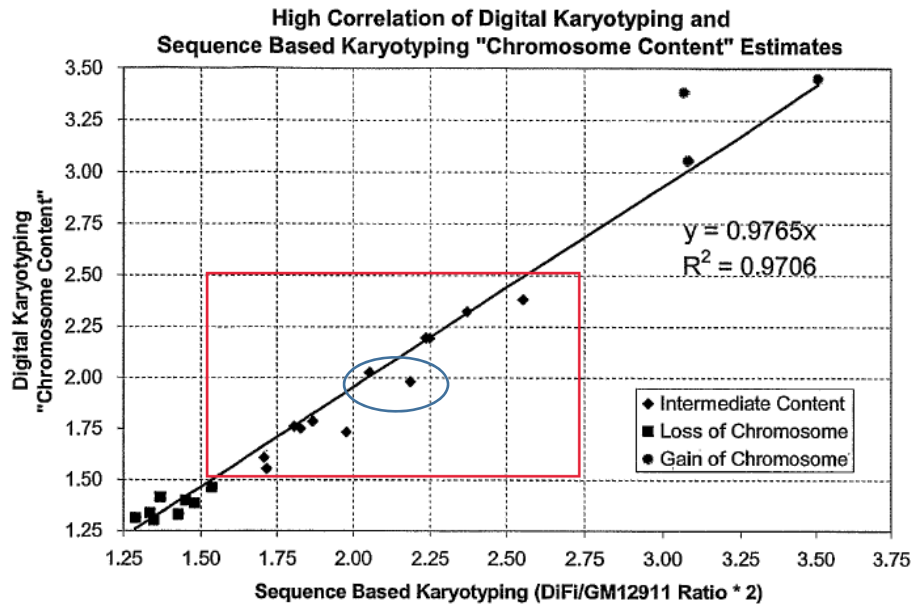


Figure 1

76. Shimkets, however, does not provide any information on the margin of error for its measurements. A skilled artisan would have recognized therefore the uncertainty in determining the fetal fraction, the wide variability in fetal fractions, and the unknown detectability error margin make it practically impossible to set meaningful threshold for using the sequence-based karyotyping method to detect fetal chromosomal abnormalities based on fetal DNA in a maternal blood sample.

77. I understand that Dr. Ding cites Shimkets's 54:20-30 in Paragraph 129 of his declaration to argue that "Shimkets ... provides methods for increasing the sensitivity of the assay, including for complex mixtures where sample DNA is limited, such as tumor biopsies and fetal/maternal blood and tissue samples" EX1005, ¶ 129. The cited paragraph is reproduced below:

To gain more sensitivity or to help in the analysis of complex mixtures, the protected primers can be modified (e.g., derivatized) with chemical moieties designed to give clear unique signals. For example, each protected primer can be derivatized with a different natural or synthetic amino acid attached through an amide bond to the oligonucleotide strand at one or more positions along the hybridizing portion of the strand. The chemical modification can be detected, of course, either after having been cleaved from the target nucleic acid, or while in association with the target nucleic acid. By allowing each protected target nucleic acid to be identified in a distinguishable manner, it is possible to assay (e.g., to screen) for a large number of different target nucleic acids in a single assay. Many such assays can be performed rapidly and easily. Such an assay or set of assays can be conducted, therefore, with high throughput efficiency as defined herein. Ex. 1006, 54:20-30.

78. As is clear from the above, the paragraph does not mention tumor, fetal/maternal blood or tissue samples. Instead, the paragraph concerns the design of "protected primers" used to hybridize template DNAs. A POSITA would also

understand that reading Shimkets as a whole, the reference to “complex mixtures” is to the complex-sample sequencing method used to detect different pathogen species in a sample. Ex. 1006, 21:10-24. As I have explained earlier, complex-sample sequencing is a qualitative determination method for identifying different (pathogen) species, and not a quantitative determination method suitable for detecting changes in chromosomal copy numbers based on comparing test distribution to reference distribution (a process not involving allele calls or specific loci). Shimkets only teaches that his complex-sample sequencing method can detect the presence or absence of specific target nucleic acid, not to quantitatively determine the presence of a chromosomal abnormality of the target in the mixture by, for example, measuring amounts of genetic material at multiple loci of the target species, computing the probability of the presence or the probability of the absence of a target DNA’s chromosomal abnormality by comparing the measured amounts to a threshold value or an expected reference distribution, and identifying the presence or absence of the target DNA’s chromosomal abnormality (such as aneuploidy) by selecting the probability most likely to be true.

79. Notably, nothing in this paragraph 54:20-30 states that, if used with a complex mixture, the relative fractions of the mixture could be determined. Instead, a POSITA would understand that the paragraph states that, with an appropriately designed primer (which Shimkets does not describe), it might become easier to

differentiate among different species or sequences. At any rate, without specifically describing the suitable primers for fetal DNA and maternal DNA in a maternal blood sample, this paragraph remains a goal, rather than an actionable plan. Shimkets, for example, states that unlike typical PCR amplification, its method “necessitates careful *de novo* primer design.” Ex. 1006, 73:27-31. No hint is given as to how such *de novo* design is to be done for a maternal blood sample containing fetal DNA or for applications that require more than identifying the presence or absence or a particular target sequence, such as for aneuploidy.

(c) Non-randomly fragmented DNA sequences can introduce systematic error

80. At the time of the invention, it was known that cell-free DNA fragments were generated by sequence-specific cleavage of fetal chromosomes. Ex. 2007 at 61. For example, in January/February 2005, Bischoff published a paper that explained that “[i]t is likely that circulating fetal DNA is released from fetal and/or placental cells undergoing apoptosis.” Ex. 2007 at Abstract. The article explained that the fetal DNA fragments in many systems “result[ed] from activation of endogenous Ca^{2+} and Mg^{2+} -dependent nuclear endonuclease.” *Id.* at 3. “The process of DNA cleavage [by the endonuclease] is not random” with “certain DNA sequences ...relatively more susceptible to cleavage.” *Id.* A 2015 paper confirmed that this is the case, establishing the inherent nature of “[n]on-random

fragmentation patterns in circulating cell-free DNA” Ex. 2011 [Ivanov2015], Title. This 2015 paper by Ivanov confirmed non-random fragmentation of fetal DNA. *Id.*, at 3.

81. In contrast, as I have noted earlier, random fragmentation is a key feature of Shimkets’s sequence-based karyotyping method. Ex. 1006, 2:20-25 (“generating a pool of fragments of genomic DNA by a random fragmentation method”); 3:1-3 (“The genome of the cell is karyotyped by randomly fragmenting the DNA from a cell and a sequencing at least a portion of each fragment.”); 9:18-20 (“It is a discovery of the present inventors that the genome of an organism can be sampled by random fragmentation and sample sequencing to determine karyotypic properties of a cell, tissue, or organism”). Shimkets even attributes certain advantage of its method to “random fragmentation.” *Id.*, 19:7-9. Shimkets equates “random fragmentation” to “non-sequence specific fragmentation” and distinguishes it from “sequence specific fragmentation” such as those expected to be observed in the cell-free fetal DNA generation process. *Id.*, 11:20-28; *see also* 11:18-20 (“Randomly fragmenting a DNA refers to the physical fragmentation ... of a large molecule of DNA into multiple smaller DNA molecules in a non-sequence specific manner.”).

82. Compared to random fragmentation, preferential sequence-specific fragmentation can yield systematic deviation from the expected numbers. For

example, if certain sequences are preferentially generated at 150% the nominal frequency, then even without any aneuploidy, the observed test distribution at a 10% fetal fraction would be 2.1 (90% maternal fraction \times 2 copies + 10% fetal fraction \times 2 copies \times 150% frequency). That is the same as would be expected when trisomy occurs. To the extent Shimkets's method can detect a 5% change with confidence (which as I have explained above, it cannot), the test distribution over the genomic sequence would show a similar ratio to that in the case of trisomy. In such a case, even if Shimkets's p-value test showed a statistically significant alteration, that alteration would not reflect a probability of the presence of a fetal chromosomal abnormality or a probability of the absence of such an abnormality. That is, p-value may show a statistically significant alteration, but it does not show the cause of that alteration.

83. Theoretically, this systematic error can be accounted for if a skilled artisan knows which fragments are preferentially generated and at what preference. But as Bianchi noted in her 2004 paper, at the time of the invention, there was limited understanding on fetal DNA's biology: "To date, little is known about the molecular and biological characteristics of the fetal DNA present in maternal circulation." Ex. 1008, S94. I am not aware of any study at the time of the invention that managed to determine the fragmentation patterns of fetal DNA including the frequency at which each sequence was expected to be generated. Without such

knowledge, it would be impossible to account for the systematic error in the modeling. I note that Progenity and Dr. Ding did not address the complication associated with fetal DNA's non-random fragmentation pattern.

2. At the time of the invention, researchers, including Bianchi, had concluded that using fetal cells derived from maternal blood was “not clinically practical”

84. I am also asked to comment on the state of the art at the time of the invention regarding using fetal cells from maternal DNA for non-invasive pre-natal testing.

85. First, if fetal cells were isolated from maternal blood and the test was then performed on them, the sample would only include DNA from the fetus and not from the mother. Such a test therefore would not involve measurements “in a sample comprising DNA from the fetus and DNA from the mother of the fetus” as required by the claims. Ex.1001, steps 1(a) and 11(a).

86. Second, at the time of the invention, the researchers generally concluded that it was “not clinically practical” to use fetal cells from maternal blood. Ex. 2013 [Bischoff2003-Lancet 2003; 361:139-140]. Bischoff and co-workers reported in a 2003 *Lancet* paper:

Rare fetal cells can be recovered from maternal blood, which suggests that non-invasive prenatal diagnosis is possible. However, recovery and analysis of fetal cells from blood is complex, and sensitivity is low because of the rarity of these cells in the maternal circulation. An

alternative strategy, which suggested that intact fetal cells can be found in maternal plasma by use of simple enrichment methods, has been reported. We aimed to replicate this technique. However, *five independent laboratories were unable to identify any intact male cells from the plasma of 38 women known to be carrying male fetuses*. Although apoptotic *intact fetal cells* could contribute to the detection of fetal DNA in maternal plasma, *we believe that recovery of these cells is difficult and not clinically practical*.

Ex. 2013 at 139.

87. After reviewing the difficulties associated with the isolation of fetal cells in maternal circulation, the authors concluded that “non-invasive detection of fetal aneuploidy will probably require novel strategies to isolate and analyse intact fetal cells from maternal whole blood.” *Id.*, 140.

88. A 2003 patent application also commented on the implication of low amounts of fetal DNA from fetal cells: “*There may not be enough to reliably determine anomalies of the fetal karyotype* or assay for other abnormalities. Furthermore, most techniques are time consuming, require high-inputs of labor, and are difficult to implement for a high throughput fashion.” Ex. 2015, [0037].

89. It was therefore generally accepted by the time of the invention that fetal cells from maternal blood would not be a suitable source for NIPT and a skilled

artisan would not have gone through the trouble to develop protocols to isolate the rare fetal cells for NIPT given this general knowledge.

90. Of note, although Progenity specifically points to Dr. Bianchi's success in isolating fetal cells for study,⁵ Dr. Bianchi is one of the coauthors of the above-mentioned Lancet paper that concluded recovery of intact fetal cells from maternal blood was "not clinically practical." Ex. 2013 at 139.

Intact fetal cells in maternal plasma: are they really there?

*Farideh Z Bischoff, Sinuhe Hahn, Kirby L Johnson,
Joe Leigh Simpson, Diana W Bianchi, Dorothy E Lewis,
William D Weber, Katherine Klinger, Sherman Elias,
Laird G Jackson, Mark I Evans, Wolfgang Holzgreve,
Felix de la Cruz*

91. Further, Dr. Bianchi's study on isolated fetal DNA cells from maternal blood in the late 1990's was not reported as involving determining the sequences of the chromosomes. Ex. 1008, S93, S96. For example, she studied aneuploidy based on total fetal amounts, not the distribution maps of the sequences. The Bianchi paper further acknowledged that the "major limitation" in using fetal cells as a fetal DNA source was "the generally low number of fetal cells found in the maternal circulation (approximately one fetal cell per ml of maternal blood)." Ex.

⁵ IPR2021-00266 Pet. 23 ("Bianchi teaches that a fetal cell advantageously could be obtained from maternal blood.").

1008, S93. This is consistent with the sentiment expressed in the 2003 Lancet paper that it is not clinically practical to use fetal cells for prenatal testing.

92. Additionally, using fetal cells for NIPT would require intact fetal cells. The reason for this is that, if the cell is already damaged and test is performed on the remaining DNA, it cannot be known whether any observed deviation from the norm was caused by cell damage and leaked or contaminated DNA content or the existence of chromosomal abnormality.

93. In summary, at the time of the invention, a skilled artisan would have recognized the impracticality to use fetal cells from maternal blood as a source of fetal DNA material for NIPT.

3. With the exclusion of large swaths of measured sequences, Shimkets analysis does not necessarily include any amounts of genetic material measured at multiple polymorphic loci.

94. '392 Patent's Claim 1 requires, *inter alia*, “measuring the amount of genetic material at ***multiple polymorphic loci*** on a chromosome or chromosome segment of interest” and “computing ... the probability of the presence and the probability of the absence of a chromosomal abnormality in the fetus by comparing the amount from step (a) to [certain values].” EX1001, 60:63-66, 61:3-8. The claims contemplate the use of identified polymorphic sites in the measurements, and not incidentally included polymorphic sites during sequencing as in Shimkets. But even if the claims are to be interpreted as Progenity does—that is, they are

satisfied by incidentally included polymorphic sites, Progenity and Dr. Ding still have not explained why steps (b)-(d) are met.

95. As I have noted above, Shimkets exclude a large number of read sequences when creating test and reference distributions. ¶¶ 43-44, 51.

96. Progenity and Dr. Ding did not analyze whether the filtered data used for distribution generation necessarily included amounts of genetic material measured at multiple polymorphic loci. As I explain below, they do not necessarily have to and likely do not.

97. First, Shimkets's methodology does not depend on the presence or absence of polymorphic alleles. Instead, Shimkets maps DNA sequences from test samples and reference samples to one or more regions of a genomic scaffold to create a distribution of mapped sequences. *E.g.*, Ex. 1006, 3:14-27, 15:1-15. For example, "a 1 Mb stretch of genomic sequence may be fragmented into 10 map locations of 100 kb each (0-100, 101-200, 201-300, 301-400, 401-500, 501-600, 601-700, 701-800, 801-900, 901-1000). Any fragments which map[] to the same range of bases (e.g., 603 kb, 650 kb, 675 kb) would be considered to be mapped to the same location." *Id.*, 14:11-15; *see also* 15:1-15. If 10,000 fragments are mapped to these 10 map locations of equal size, "each location is expected to have a frequency of 1000 mapped fragments." *Id.*, 15:5-7. Shimkets observes that "[s]ome non-uniformness will be introduced by the fact that genomes contain regions of

repetitive sequence which are non-uniformly distributed throughout the genome.” *Id.*, 15:7-9. But Shimkets’s way of addressing this non-uniformity is to remove repetitive sequences from test distribution mapping. *E.g.*, 17:22-24 (“Filtering is performed to remove DNA sequences within repeated regions and to remove the rare DNA sequences not present in the human genome”); 61:15-16 (“Any sequences that mapped to more than one position were discarded.”). Reference distributions are generated similarly. *Id.*, 14:22-24, 15:1-15.

98. Put another way, Shimkets’s Sequence-Based Karyotyping method is indifferent to the presence (or absence) of polymorphic alleles, but instead focuses on the matching with common sequences. Shimkets’s analysis step therefore does not need to include data measured at any polymorphic locus.

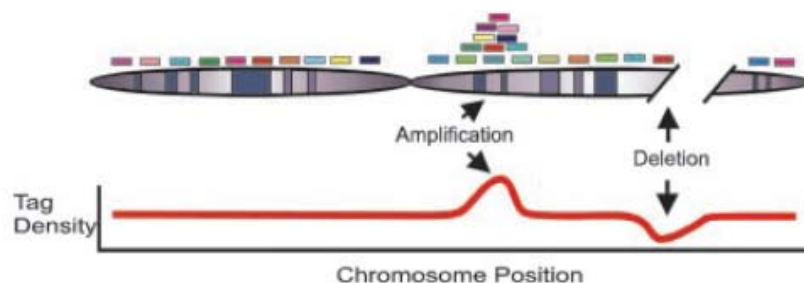
99. Second, a skilled artisan would also recognize that excluding amounts measured at polymorphic sites is unlikely to affect the outcome of Shimkets’s method. For example, Progenity noted the fact that “5.3 million common SNPs” had been identified by the time of the invention. IPR2021-00266 Pet. 26 (citing Ex. 1016). But 5.3 million common SNPs accounts for only 0.17% of the human genome:

$$\frac{\sim 5.3M \text{ common SNPs}}{\sim 3,200M \text{ base pairs per haploid genome}} \approx 0.17\%$$

100. In comparison, as I noted above, Shimkets removed 65% and 58% of the read sequences from mapping when generating test and reference distributions. Ex. 1006 at 61:11-17 (only 125,684 out of 354,451 reference fragments and 203,352 out of 487,310 test fragments were mapped to the Human Genome, corresponding to 65% and 58% of removed sequences for reference and test samples, respectively).

101. Third, as I explained in paragraph 51 above, Shimkets validates its karyotyping method by comparing its result with that obtained by the digital karyotyping method discussed in a paper by Wang *et al.* EX1006, 59:8-10, Fig. 1, Figs. 2-3 (insets in these two figures originated from Wang). A skilled artisan would understand that for an apples-to-apples comparison, Shimkets would have analyzed data in the same way as Wang.

102. Wang refers to its DNA sequences as tags, and uses tag density to detect copy number gain or loss. Ex. 2012 [Wang] at Fig. 1. (presented below).



Wang, Fig. 1. Showing the mapping of genomic tags (colored boxes) to a chromosome to generate tag density.

103. As I have noted in paragraph 51 above, Shimkets’ stated rules of exclusion generally matches that of Wang’s. Because Wang expressly excludes sequences with polymorphisms, Shimkets would also have removed sequences associated with polymorphic loci to facilitate apples-to-apples comparison.

Shimkets	Wang
<p>“Filtering is performed to remove DNA sequences within repeated regions and to remove the rare DNA sequences not present in the human genome.”</p> <p>“Sequences that did not map to the genome, either due to incompleteness of the genomic scaffold or issues of sequencing quality, were removed from consideration. Filtering was also performed to remove DNA sequences which mapped to multiple genomic locations (within repeated sequences).”</p> <p>Ex. 1006, 17:22-24, 58:23-26.</p>	<p>The experimental tag with the same sequence as virtual tags were termed filtered tags and were used for subsequent analysis. The remaining tags corresponded to repeated regions, sequences not present in the current genome database release, polymorphisms at the tag site, or sequencing errors in the tags or in the genome sequence database.</p> <p>Ex. 2012 at 16157.</p>

104. In summary, the explanation provided above shows why it is neither expressly disclosed nor necessarily needed for Shimkets to use amounts measured at polymorphic loci for its karyotyping.

B. Huang

105. U.S. Patent Publication No. 20050064476 to Huang *et al.*, with a title “Methods for Identifying DNA Copy Number Changes” was published on March

24, 2005. EX1007, cover. The experimental detail of Huang was also published as a paper entitled “Whole Genome DNA Copy Number Changes Identified by High Density Oligonucleotide Arrays” in *Human Genomics* in May 2004. *See* Ex. 2008 (“Huang’s Paper”).

106. Huang measures changes in DNA copy number by first genotyping each individual single nucleotide polymorphisms (SNP) that may be present in a genomic region of interest. *Id.*, [0161]. Huang uses the Whole-Genome Sampling Assay (WGSA) method as disclosed in a paper by Kennedy to reduce the complexity of the starting genome by fragmenting and amplifying only a particular segment of the genome of interest and to genotype SNPs that may be in that genomic region. *See, e.g., id.*, [0167]. The amplified fragments are then hybridized to an array of probes.

107. Huang measures the hybridization intensity for each probe element in the array and compares the intensity response of the target cell lines to that of a reference. EX1007, [0063]. The difference in intensity provides information on the change in copy number of DNAs, because higher intensity is generally correlated with higher copy number. *Id.*; *see also* [0081]-[0082]. For example, “[t]he relationship between log intensity and log copy number was found to be approximately linear and using control samples of known copy number the slope and y-intercept of the line may be estimated.” *Id.* As such, “copy number is

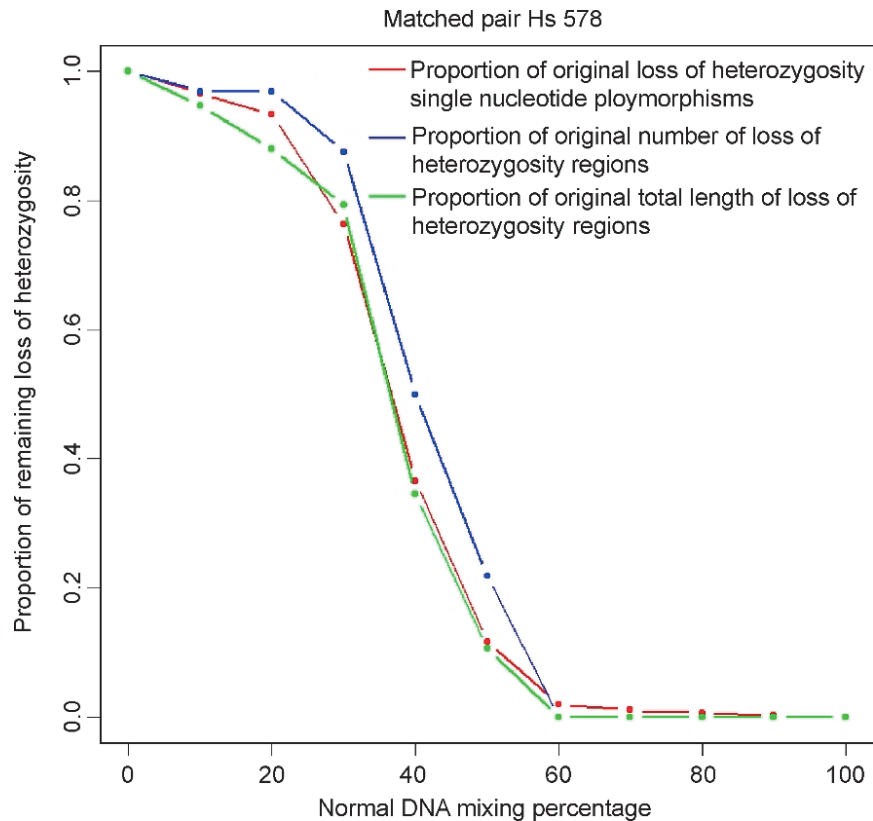
estimated by comparing an *intensity measurement for a SNP* in an experimental sample to a distribution of intensity measurements from the *same SNP in a plurality of reference samples.*” *Id.*, [0070] (emphasis added).

108. Huang also discloses an embodiment “for estimating genome wide copy number using high density oligonucleotide array.” *Id.*, [0083]. This method couples the analysis of LOH (loss of heterozygosity) with DNA gains and losses. *Id.*, [0083]. The intensity of the hybridized sample is compared to “a reference set consisting of more than one hundred normal individuals” and p-values are used to identify “statistically significant gains and losses.” *Id.*

109. Huang applied her method to “pure tumor samples or mixed samples, containing both normal and tumor DNA.” *Id.*, [0072]. Huang’s method, however, is severely limited when it comes to mixed samples. *Id.*, [0147]-[0150]. Huang analyzed mixed samples “for changes in LOH and for changes in the detection of copy number alterations.” *Id.*, [0148]. She mixed cancer cells and normal cells at a 10% incremental step of increasing fractions of normal cells. *Id.*

110. For LOH, she observed “nearly 60% loss of detection of LOH” “when the mixed normal DNA reaches 30 to 50% of the total.” *Id.* More than 98% of the SNPs undergoing LOH became undetectable “[w]hen normal DNA is present at 60% or greater.” *Id.* The loss in detectability is not linear with respect to the amount of contaminants (DNAs from normal cells), but drops precipitously between 30-

60% of the normal DNA fraction. This is seen in the graph below from Huang's technical paper. Ex. 2008, Fig. 5.



Ex. 2008 (Huang Paper), Figure 5. Loss of heterozygosity (LOH) analysis on mixed samples.

111. Huang observed the same trend in the loss of detectability on the relative percentage of copy number alterations. EX1007, [0150]. At 30% normal DNA, the detectable signals remaining from original total for copy number alternation was 57.55% using meta-analysis and 21.43% using individual analysis. *Id.* At 60% normal DNA, “most of the signals are undetectable.” *Id.*

112. Huang concludes that “detection of LOH and copy number alteration using the WGS assay and algorithm can tolerate a mixed sample containing up to 20 to 30% normal DNA.” *Id.* That is, Huang can tolerate up to 30% contaminants.

113. In comparison, the contaminating maternal DNA in a maternal blood sample would far exceed the 30% threshold, and most likely in the 80-90% range. For example, Bianchi reports the following fetal DNA concentrations:

	Mean fetal fractions in serum (range)	Mean fetal fractions in plasma (range)
early pregnancy	0.13% (0.014-0.54%)	3.4% (0.39-11.9%)
late pregnancy	1.0% (0.032-3.97%)	6.2% (2.33-11.4%)

Ex. 1008, S94.

114. A technical brief by Wei submitted by Progenity reports a median fetal fraction of 11.8% for 19 male pregnancy plasma samples. Ex. 1026 at 337-338 and Table. 1. The fetal DNA fractions ranged between 1.98% and 47.69%. Fifteen (15) of the nineteen (19) samples had fetal DNA fractions below 20%, three had fetal DNA fractions between 20-30% and one had an unusually high fetal fraction at 47.69%. *Id.* Even at delivery, (a time point not relevant to NIPT), the median concentration of fetal DNA in maternal plasma is only 14.3% (2.3-60%). Ex. 1008, S96.

115. All these data demonstrates that the fraction of contaminating maternal DNA would generally be above the 20-30% upper tolerance limit found suitable for Huang's cancer detection. Even if detection of DNA derived from cancer cells could be directly transferred to detection of cell-free fetal DNA, a skilled artisan would not have reasonably expected that Huang's method would work for detecting fetal chromosomal abnormalities in a crude maternal blood sample: the amounts of contaminating maternal DNA would just be too overwhelming and far exceed the 20-30% limit.

116. A POSITA would also recognize that even with amplification, the starting fetal DNA amounts may be too small for reliable SNP array genotyping.

117. For example, amplifications are generally designed for a selected size range of DNA fragments. The disclosed size ranges in Huang are "400-800 or 400 to 2000 base pairs." EX1007, [0058]; *see also id.* [0114] ("Amplification methods may be optimized to amplify a subset of these fragments, for example, the fragments that are 400 to 800 basepairs."). Huang states that "[f]ragments that are outside the selected size range are not efficiently amplified." *Id.*

118. According to the Chan paper submitted by Progenity, however, the fetal DNA fragments were generally much smaller than 400 base pairs. For example, Chan reports that "[t]he median percentages of fetal-derived DNA with sizes >193bp and >333 bp were 20% and 0%, respectively, in maternal plasma." Ex.

1049 at abstract. Even with ligation of primers (typically adding fewer than 100 bps), the vast majority of the fetal DNA fragments would be below the lower threshold for the size range selected for Huang's amplification procedure. This would further suppress the fraction of fetal DNAs in the products after amplification.

119. Another problem, as I have explained in the IPR2019-01201 proceeding for the related '592 patent was that at the relevant time, "various researchers had determined that very low amounts of DNA could not be reliably used for SNP genotyping arrays (even with intervening amplifications)." IPR2019-01201, Paper 54 (FWD) at 23 (P.T.A.B. Dec. 11, 2020) ("IPR2019-01201 FWD").

120. Huang used, for example, 25 ng of adaptor-ligated genomic DNA for PCR amplification. EX1007, [0127]. In contrast, according to Bianchi, the average amount of fetal DNA concentration was 24.2 GE/ml in a maternal blood with normal pregnancy and 41.2 GE/ml in the case of trisomy 21. Ex. 1008, S97. Another researcher observed that the absolute amount of fetal DNAs was the same in serum and in plasma. Ex. 2019 [Chiu]. At 3.6 pg per GE, 25 ng of genomic material would imply 287 mL of maternal blood for a normal pregnancy and 168 mL for a trisomy 21 case. In contrast, typical blood draw would be in the 10-15 mL range. *See* Ex. 2019 [Chiu] at 1608-09 (10-12 mL per pregnant woman).

121. Generally, less than 5 mL of the blood sample is used for amplification—(see Ex. 2019 at 1608-09, which used a fraction of the 4mL of plasma sample for real-time PCR) —that would only be 121 GE for a normal case and 206 GE for a trisomy 21 case.

122. As I stated earlier, one GE is approximately 0.0036 ng (3.6 pg). Thus, 121 GE of fetal DNA from 1 ml of a maternal blood sample would yield 0.435 ng of genomic fetal DNA material. If trisomy 21 does have elevated fetal DNA amounts in a maternal blood sample (a result which Bianchi acknowledged was not always reproducible, Ex. 1008 at S96), the amount of fetal genomic materials would be 0.74 ng @ 41.2 GE/mL. These amounts are 35-60 fold less than the 25 ng Huang used for her amplification procedure.

123. This amount is also much less than the 3ng minimal threshold that Lovmar recommended for a successful SNP genotyping using MDA or the 10 ng amount that Bergen recommended for “observ[ing] optimal TaqMan® SNP genotyping performance from MDA wgaDNA.” Ex. 2020 [Lovmar] at 9; Ex. 2021 [Bergen] at 4.

124. The amounts are even below the threshold stated in Progenity’s own exhibits. For example, Lasken discloses that for whole genome amplification as in Huang, “each SNP determination requires 2 and 50 ng of gDNA” and often “many micrograms of each DNA sample.” Ex. 1035 at 533. Lasken warns that “[w]hen

less DNA is available or the template is degraded, special care must be taken in the interpretation of genotyping results because amplification will be prone to allele drop-out owing to stochastic effects.” *Id.*

125. Dietmaier (Ex. 1036), which was also raised in the IPR2019-01902 proceeding addresses amplification of DNA from tumor cells and does not use the amplification product in a SNAP array, as the Board previously found. IPR2019-01902 FWD at 21. Zheng, which cites Dietmaier, uses “**20-50 ng** of whole blood or bone marrow genomic DNA,” and addresses cancer cells (buccal cells from childhood leukemia cases). Ex. 1037 at 2. Paunio did perform its method on four single blastomeres isolated from a four-cell-stage human embryo, but the paper did not involve SNP arrays or DNA fragmentation size similar to those used by Huang. Ex. 1038.

126. It is true that Huang mentions an amplification procedure that was supposedly “successfully applied to as few as 1-10 copies of human genomic DNA.” IPR2021-00266 Pet. 43 (citing EX1007, [0075]). The referenced MDA method, however, was not used with short fragments like fetal DNA fragments and was based on amplification from tissue culture cells having much higher purity in target DNA than possible with circulating cell-free DNA in maternal blood. Ex. 2022 [US 6,617,137] at 49:50-53. Further, an amplification method such MDA would amplify both fetal and maternal DNA present in cell-free DNA, increasing

any potential signal as well as the much more sizable noise present in the maternal genome.

C. Bianchi

127. Bianchi is a review entitled “Circulating Fetal DNA: Its Origin and Diagnostic Potential—a Review” published in 2004. Ex. 1008, cover. The article provides a review of the developments in fetal DNA research.

128. Bianchi reports that “both intact fetal cells and cell-free nucleic acids circulate freely within maternal blood.” Ex. 1008, S93. According to Bianchi, “the generally low number of fetal cells found in the maternal circulation (approximately one fetal cell per ml of maternal blood)” limits fetal cell studies because “in many cases [analysis of maternal blood] requires the use of sophisticated fetal cell enrichment procedures.” *Id.* Bianchi also notes a “main limitation” to studying cell-free fetal DNAs: “the availability of uniquely fetal gene sequences that identify and/or measure the presence of fetal DNA in both male and female fetuses.” *Id.*, S93–S94.

129. Bianchi acknowledges that as of her paper in 2004, “little [was] known about the molecular and biological characteristics of the fetal DNA present in maternal circulation.” *Id.*, S94. However, Bianchi states that “there is indirect evidence to suggest that the circulating fetal DNA fragments are small (<450 bp).” *Id.* In fact, according to a paper by Chan, only 20% of circulating cell-free DNA

has a fragment size about 193 base pairs (bp) and none with a fragment size above 313 bp. Ex. 1049, Abstract. Bianchi also concludes that “the majority of the cell-free fetal DNA in the maternal circulation probably originates from the placenta with an additional (minor) contribution from haematopoietic cells and possibly from the fetus itself via direct transfer.” Ex. 1008, S96.

130. Bianchi documents the wide variability in the percentages of cell-free fetal DNAs in maternal blood, which increased the difficulty in discerning whether any change is statistically significant. This included her own experiments. Ex. 1008, S97 (observing “there was a wide distribution of fetal DNA levels in both [trisomy 21] cases and controls, which decreased the average separation between these two groups”).

131. Bianchi discusses the potential of using cell-free fetal DNA for “Noninvasive diagnosis of fetal aneuploidy.” *Id.* But her method did not involve measuring the amount of genetic material at polymorphic loci. Instead, it was based on the observation that fetal DNA levels in maternal circulation might increase by a statistically significant amount if the fetus suffered from aneuploidy. *Id.*, S96-98. For example, in one study, using archived blood samples and case-control model that matched for fetal gender, gestational age and duration of freezer storage, Bianchi’s team observed “adjusted mean serum fetal DNA concentration were 41.2 GE/ml in the Down syndrome [trisomy 21] cases, and 24.2 GE/ml in the euploid

controls ($P=0.002$).” *Id.*, S97. Researchers also observed “significantly elevated” fetal DNA levels in trisomy 13 cases. *Id.* They, however, did not observe any statistically significant difference between trisomy 18 cases and controls. *Id.* Of note, not all groups studying fetal DNA levels in maternal plasma or serum observed an elevated level of fetal DNA concentration in trisomy 21 cases. *Id.*, S96. The studies for trisomy 13 and 18 cases do not appear to have been replicated. *Id.*, S97.


132. Bianchi reports that when fetal DNA levels were combined with then-existing biomarkers for Down syndrome, the detection rate increased “from 81 to 86 per cent.” *Id.*, S97. Used alone, the detection rate was only 21% with a 5% false-positive rate. *Id.* She posits that the wide variability in data contributed to the reduced detectability, and posited that “with larger studies the screening performance of fetal DNA [would] improve.” *Id.* There is no evidence in the record that the wide variability can be addressed with a larger population; and I am not aware of such studies.

133. Bianchi expresses the hope that “amplified fetal nucleic acids will *ultimately* permit a non-invasive fetal genome scan as part of routine prenatal screening.” Ex. 1008, S99. But apart from her total fetal content measurements, Bianchi did not provide any other viable path forward. Her approach, examining the fetal content as a whole, however, does not reliably detect or predict the

presence or absence of fetal chromosomal abnormality. That problem was solved by the Natera inventors after their years of dedicated research.

This declaration and my opinions herein are made to the best of my knowledge and understanding, and based on the material available to me, at the time of signing this declaration. I declare under penalty of perjury under the laws of the United States of America that all statements made of my own knowledge are true and that all statements made on information and belief are believed to be true. I understand that willful false statements and the like are punishable by fine or imprisonment, or both (18 U.S.C. 1001) and may jeopardize the validity of the application or any patent issuing thereon.

Date: March 8, 2021

Signed:  _____

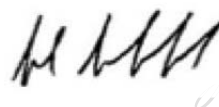
John Quackenbush, Ph.D.

presence or absence of fetal chromosomal abnormality. That problem was solved by the Natera inventors after their years of dedicated research.

This declaration and my opinions herein are made to the best of my knowledge and understanding, and based on the material available to me, at the time of signing this declaration. I declare under penalty of perjury under the laws of the United States of America that all statements made of my own knowledge are true and that all statements made on information and belief are believed to be true. I understand that willful false statements and the like are punishable by fine or imprisonment, or both (18 U.S.C. 1001) and may jeopardize the validity of the application or any patent issuing thereon.

Date: March 8, 2021

Signed: _____



Digitally signed by John
Quackenbush
Date: 2021.03.08 16:45:50
-05'00'

John Quackenbush, Ph.D.