APPENDIX A

APPENDIX A

DECLARATION OF PROFESSOR JOHN QUACKENBUSH, PH.D.

TABLE OF CONTENTS

Page

I. SUMMARY OF OPINIONS 1

II. BACKGROUND AND QUALIFICATIONS 4

III. MATERIALS RELIED UPON 8

IV. FRAMEWORK FOR MY ANALYSIS 8

- A. *Alice/Mayo* Framework 8
- B. Priority Date 9
- C. Person Of Ordinary Skill In The Art 10

V. BACKGROUND OF THE PATENTS 11

- A. The Patent Family 11
- B. Stanford's Patent Claim Strategy 12
- C. The Patent Claims 14
  - 1. The '652 Patent 14
  - 2. The '497 Patent 16
  - 3. The '607 Patent 17

VI. BACKGROUND ON THE TECHNOLOGY 20

- A. Nucleic Acids And Genes 20
- B. Genes, Polymorphisms, Homozygosity, And Heterozygosity 21
- C. Cell-Free Nucleic Acids 24
- D. Cell-Free Nucleic Acids Having Sequences Different From A Patient's Genotype Occur Naturally In An Individual's Circulation 26
  - 1. Cell-Free Nucleic Acids In Organ Transplant 26
  - 2. Cell-Free Nucleic Acids In Pregnancy 27
  - 3. Cell-Free Nucleic Acids In Cancer 28
  - 4. Cell-Free Nucleic Acids In Infectious Disease 29
  - 5. Scientists Have Long Used Cell-Free Nucleic Acids To Monitor Disease 30

VII. THE PATENT CLAIMS ARE DIRECTED TO WELL-UNDERSTOOD, ROUTINE AND CONVENTIONAL TECHNIQUES APPLIED TO OBSERVING A NATURAL PHENOMENON 31

A. The Claims of the Patents Are Directed to a Natural Law 31

B. Each Recited Laboratory Technique Was Conventional 34

1. Obtaining Or Providing a Biological Sample Containing Cell-Free Nucleic Acids From a Transplant Recipient 39

2. Genotyping The Transplant Donor And/Or Recipient To Establish Profiles Of Genetic Polymorphisms (Or SNPs), Including By Comprising Selectively Amplifying At Least 1,000 SNPs By PCR And Sequencing 41

(a) PCR, Including Selective Amplification, Was a Well-Understood, Routine and Conventional Method For Genotyping By 2009 43

(b) Arrays Were a Well-Understood, Routine and Conventional Method For Genotyping By 2009 46

(c) Sequencing Was a Well-Understood, Routine and Conventional Method For Genotyping By 2009 49

(d) The Patents Claim No Improvements To Or Nonconventional Uses Of The Conventional Genotyping Methods Available In 2009 51

3. Performing Multiplex Or High-Throughput Sequencing Of (Or Digital PCR On) The cfDNA To Detect The Genotyped Polymorphisms (Or SNPs) 52

(a) Background On Multiplex / High Throughput Sequencing Technology 53

(b) Commercially Available Multiplex And High-Throughput Sequencing Platforms As Of 2009 55

(c) Sensitivities And Error Rates Inherent In The Available Sequencing Instruments 62

(d) Digital PCR Methods Were Routine And Commercially Available In 2009 68

(e) The Patents Claim No Improvements To Or Nonconventional Uses Of The Conventional

Multiplex/High-Throughput Sequencing Or dPCR Methods Available In 2009 68

4. Quantifying The Transplant cfDNA In The Sample, Using The Polymorphisms / SNPs In The Sequences 69

C. The Claimed Combinations Of Steps For Detecting Multiple Genomes In A cfDNA Sample Were Conventional 72

1. The Patents' Written Description States That The Laboratory Techniques In The Claims As A Whole Were Conventional As Of 2009 72

2. POSAs Have Long Recognized That The Same Methods For Detecting Fetal, Cancer, And Infectious cfDNA Could Be Used To Detect Transplant cfDNA 76

3. POSAs Had Already Established The Claimed Laboratory Techniques By 2009 To Detect Multiple Genotypes In A Sample, Including In The Transplant Setting 82

(a) Cantor et al. – 2007 82

(b) Lo et al. – 2009 85

(c) Dhallan et al. – 2007 88

(d) Beck et al. – 2009 91

(e) Beck et al. – 2010 94

(f) Gordon et al. – 2009 96

(g) Holt et al. – 2009 98

D. The Dependent Claims Also Recite Well-Understood, Routine and Conventional Steps 100

1. Dependent Claims Reciting Different Types Or Numbers Of Polymorphisms 101

2. Dependent Claims Reciting Different Types Of cfDNA 103

3. Dependent Claims Reciting Certain Common Modifications, Error Rates, Or Quality Scores Associated With Multiplexed Or High Throughput Sequencing 104

4. Dependent Claims Reciting Organ or Species Limitations ... 108

5. Dependent Claims Reciting Conventional PCR Or Amplification Methods 108

6. Dependent Claims Reciting Different Sample Types 109

7. Dependent Claims Reciting Further Genotyping Prior To Quantifying Cell-Free Nucleic Acids 110

8. Dependent Claims Reciting Certain Concentrations Of Cell-Free Nucleic Acids In The Sample 110

E. Concluding Opinions Regarding Conventionality of the Recited Combination of Laboratory Techniques..... 111

VIII. THE PATENT METHODS DO NOT CREATE ANY NEW PREPARATION..... 112

I. SUMMARY OF OPINIONS

1. In my opinion, the methods claimed in U.S. Patent Nos. 8,703,652 (the "'652 Patent), 9,845,497 (the "'497 Patent"), and 10,329,607 (the "'607 Patent", and collectively the "Patents") were, as of at least November 6, 2009, well-understood, routine, and conventional approaches to detecting a natural phenomenon regarding foreign genetic material (here, quantities of organ donor-derived cell-free nucleic acids) in a patient's body (here, an organ transplant recipient's circulation). In the case of the '652 Patent, the claims further recite routine and conventional approaches to correlating increasing amounts of that foreign genetic material over time to a disease (here, transplant rejection or failure).

2. As recited in the Patent claims, and as confirmed by the Patents' common written description, the presence and amount of foreign cell-free nucleic acids in a patient's body are a natural phenomenon. They are a natural consequence of biological processes, *i.e.*, death of the cells in the transplanted organ, and release of the cells' contents into the recipient's body. The correlation of amounts of those foreign cell-free nucleic acids to disease also is a natural phenomenon, *i.e.*, it is an observation that the more the transplanted organ is rejected or fails, the more of its cells die, and the more its cells' contents are released into the transplant recipient's body. These phenomena are not something the Patent inventors created; they are the result of biological processes occurring in a transplant recipient's body.

3. As recited in the Patents' written description, as well as established by the references discussed below, the correlation between circulating nucleic acids that differ from a patient's normal genotype and solid organ transplant rejection has been known since prior to at least November 6, 2009. B0001-B0024 ('652 Patent) at B0012-B0013, 6:56-8:21. For example, citing publications from 2006 and 2007, the Patents' written description shows it had already been established that "for heart transplant patients, donor-derived DNA present in plasma can serve as a potential marker for the onset of organ failure." *Id.* at B0013, 8:18-21.

4. The methods of the Patent claims are directed solely to detecting these natural phenomena using well-understood, routine and conventional procedures known to a person of skill in the art ("POSA") as of at least November 6, 2009. This is confirmed by the Patents' shared written description, which states "[t]he practice of the present invention employs, unless otherwise indicated, conventional techniques of immunology, biochemistry, chemistry, molecular biology, microbiology, cell biology, genomics and recombinant DNA, which are within the skill of the art." B0001-B0024 ('652 Patent) at B0012, 5:36-40. Regarding the language, "unless otherwise indicated," neither the written description nor the claims identify any nonconventional techniques for use in the claims, nor anything nonconventional about the way the techniques are used together in the claims. The claimed laboratory techniques were already, by November 6, 2009, standard means

for detecting foreign or abnormal cell-free nucleic acids in other contexts, including naturally occurring fetal DNA in the blood of a pregnant woman, naturally occurring mutated tumor DNA in the blood of a cancer patient, and naturally occurring bacterial or viral DNA in the blood of a sick person. *E.g.*, B0001-B0024 ('652 Patent) at B0012-B0013, 6:56-7:36.[1] Moreover, the Patents and prior and contemporaneous literature acknowledge that the same techniques used to quantify foreign nucleic acids in the fetal, cancer, and infectious disease contexts applied equally to the quantification of foreign nucleic acids in the transplant context.

5. Furthermore, it is my opinion that the claims here are not directed to the creation of any new or unnatural preparation. Rather, the techniques described in the Patent claims begin with obtaining from a patient a sample that already contains naturally occurring cell-free nucleic acids, and end with an observation of those same naturally occurring cell-free nucleic acids (including as being a marker for organ transplant rejection).

6. It is my opinion that as of November 6, 2009, the Patents offered no improvements upon, or nonconventional ways of using, the laboratory techniques

[1] *See id.* at 7:30-36 ("In all these applications of circulating nucleic acids [referring to publications no later than 2008], the presence of sequences differing from a patient's normal genotype has been used to detect disease. In cancer, mutations of genes are a tell-tale sign of the advance of the disease; in fetal diagnostics, the detection of sequences specific to the fetus compared to maternal DNA allows for analysis of the health of the fetus.").

recited in the claims. These techniques were well known in the art and standard for detecting and quantifying foreign cell-free nucleic acids in a patient's blood by at least November 6, 2009, when the Patents were filed. I base these opinions on the statements in the Patents' specification (written description and claims) describing all of these techniques as conventional, as well as my own experience and the literature analyzed below.

II. BACKGROUND AND QUALIFICATIONS

7. I am the 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. In addition, I direct the MS program in Computational Biology and Quantitative Genetics, the Big Data to Knowledge Ph.D. training grant, and am a co-director for our Cancer Ph.D. Training renewal application. I am also a Professor of Biostatistics and Computational Biology and Cancer Biology at the Dana-Farber Cancer Institute. Until recently, I also served as the Director of the Center for Cancer Computational Biology at Dana-Farber, which performed large-scale data analyses in support of research studies, including many that involved circulating DNA. I now direct the Quantitative Biomedical Research Center at the Harvard Chan School, which performs similar work in support of research.

8. As discussed in more detail below, prior to November 6, 2009, I had significant experience using the standard techniques and standard combinations of

techniques recited in the Patent claims. For example, prior to November 6, 2009, I used standard PCR techniques for genotyping, including by selective amplification of SNPs, in large-scale projects analyzing cancer from diverse biological samples including cell lines, human tissue samples, circulating tumor cells, and circulating cell-free DNA. In all of these projects, I used standard and routine PCR applications for genotyping, including for genotyping involving selective amplification of SNPs. I also have significant personal experience using arrays for genotyping, including SNP genotyping. I also have significant personal experience using what as of November 6, 2009 were conventional multiplex or high-throughput sequencing instruments for genotyping and quantifying SNPs, including in cell-free nucleic acids. I personally used routine and conventional techniques to carry out the combinations of methods recited in the Patents as of November 6, 2009.

9. In addition to my academic appointments and research responsibilities, I serve on the Editorial Board of Cancer Research, the flagship scientific journal of the American Association for Cancer Research. I also serve on the Editorial Boards of BMC Genomics and Biotechniques, and was previously the Editor-in-Chief at Genomics. Over the years, I have also taken on significant regional and national research leadership responsibilities. I served on the Scientific Advisory Board of the St. Jude's Children's Research Hospital, the Lovelace Respiratory Research Institute, the National Institute of Health's TOPMED program, and the Board of

Scientific Counselors of the Environmental Protection Agency. I am currently on the Scientific Advisory Board at Shriner's Hospitals for Children and advise Merck KGaA on precision medicine and immunotherapy. I have served on numerous scientific review panels for the National Institutes of Health, the National Cancer Institute, the National Research Council, the National Science Foundation, the Wellcome Trust, Genome British Columbia, Genome Canada, the European Union, and the National Academy of Sciences.

10. I am a co-founder of Genospace LLC. 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.

11. I have presented more than 300 talks, courses, and workshops since joining Dana-Farber and the Harvard School of Public Health in 2005. I have also taught numerous courses and workshops on the analysis of genomic data. I have particular expertise in molecular biology techniques and gene expression, including DNA sequencing, gene expression analysis, methylation analysis, the mapping of transcription factor binding sites, and the development of methods for data analysis and interpretation in computational and systems biology.

12. 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 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.

13. 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 degree in Physics from the University of California, Los Angeles and a Bachelor of Science degree in Physics from the California Institute of Technology.

14. In addition to the NCI Outstanding Investigator Award, I have received numerous additional 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.

15. I have studied, taught, practiced, and conducted research in genomics and computational analysis of data related to genomic technologies. I have expertise in genomic technologies, computational analysis, and modeling of systems related to human diseases and genetic mutations. I have more than 25 years of continuous grant support from the National Institutes of Health, including extensive support from the National Cancer Institute. In 2018, I was awarded an NCI Outstanding

Investigator Award, which is a research grant "Career Award" that provides long-term research funding support and increased flexibility for investigators with outstanding records of research productivity, allowing them to continue or embark upon a research program of unusual potential in cancer. I have published 301 peer-reviewed scientific papers, more than half of which directly involve some aspect of cancer research, and nearly all of which are relevant to understanding what drives the development and progression of the disease. My articles have been cited over 75,000 times.

16. My curriculum vitae is attached to this declaration as Exhibit 1.

III. MATERIALS RELIED UPON

17. I have based my opinions reflected in this declaration on my education and professional experience, my review of the Patents and file histories, my review of the literature and other patents, and my review of product information as reflected in this declaration and also in Exhibit 2 to this declaration.

IV. FRAMEWORK FOR MY ANALYSIS

A. *Alice/Mayo* Framework

18. I am not a lawyer, but it is my understanding that the Supreme Court has held that laws of nature, natural phenomena, and abstract ideas are not patentable under the Patent Act, 35 U.S.C. § 101 ("Section 101"). It is my understanding that a two-step analysis is used to determine what is and what is not patent eligible. First, a court must decide whether the claims at issue are directed to one of the three patent-

ineligible concepts: abstract ideas, natural laws, or natural phenomena. If so, then the court must analyze whether the claims provide an inventive concept.

19. It is my understanding that a method reciting a law of nature or natural phenomena is not patentable where the steps involve well-understood, routine, conventional activity previously engaged in by researchers in the field. It is further my understanding that appending routine, conventional steps to a natural phenomenon, specified at a high level of generality, is not enough to supply an inventive concept for purposes of making claims patentable under Section 101.

20. I further understand that a court may analyze claims as "representative" of other claims for purposes of determining patent eligibility. I understand that a claim is representative of other claims if the claims are substantially similar and linked to the same patent-ineligible concept (*e.g.*, the same natural phenomenon).

21. I have applied these standards in my analysis of the claims of the Patents.

B. Priority Date

22. I assume, for purposes of this analysis only, that the Stanford Patents are entitled to the benefit of the November 6, 2009 date of the earliest-filed provisional application, U.S. Patent Application No. 61/280,674, to which the patents claim priority. I have not analyzed whether the Patents are entitled to the date of the earliest filed provisional application to which they claim priority. My

conclusions would be the same, however, if I used any of the dates of the later-filed applications to which the Patents claim priority.

C. Person Of Ordinary Skill In The Art

23. I am informed that for purposes of a Section 101 patent eligibility, claims are analyzed from the perspective of a person of ordinary skill in the art ("POSA") at the time of the priority date of the patent. I am further informed that a POSA is a hypothetical person to whom one could assign a routine task in the relevant field with reasonable confidence that the task would be successfully carried out. I have been asked to assume that the relevant timeframe for the Patents is prior to November 6, 2009, which I am assuming, for purposes of this analysis, is the priority date.

24. In my opinion, a POSA, for the Patents, at or prior to November 6, 2009, would have at least a bachelor's degree (with several additional years of laboratory experience) in a life or physical sciences discipline with applications in genomics. Through further education or graduate or postdoctoral experience, such a person would have obtained a thorough understanding of nucleic acid detection, quantification, and genotyping, and would have understood the basic underlying scientific principles of nucleic acid assay and detection and genotyping techniques. A POSA would also have general knowledge of, and be able to read literature on, nucleic acid assay methods including, for example, amplification by polymerase

chain reaction ("PCR"), nucleic acid sequencing including by multiplex or high-throughput sequencing, genotyping, and the detection and quantification of nucleic acids using, e.g., polymorphism-based genotyping. *See* Section VII.B. below for exemplary prior art relating to these laboratory techniques. In determining the level of ordinary skill in the art, more education could compensate for less experience, and vice versa.

25. Moreover, a POSA would have an understanding of methods for detecting and quantifying cell-free DNA ("cfDNA"), including cfDNA that is foreign to or abnormal in a sample, in various contexts including fetal cfDNA in samples from pregnant women, tumor cfDNA in samples from cancer patients, pathogen cfDNA in samples from infectious disease patients, and transplant cfDNA in organ transplant recipients, among others. In determining the level of ordinary skill in the art, more education could compensate for less experience, and vice versa.

26. On and prior to November 6, 2009, I would have been at least a person of ordinary skill in the art as it pertains to the Patents.

V. BACKGROUND OF THE PATENTS

A. The Patent Family

27. I understand that the Patents are members of the same patent family. They all claim priority to the same November 6, 2009 provisional patent application,

U.S. Patent Application Number 61/280,674, and they share the same written description.

28. The Patent claims are likewise directed to similar methods. Claim 1 of the '652 Patent recites "[a] method for ***detecting transplant rejection … or organ failure***" by "***determining a quantity of [] donor cell-free nucleic acids based on the detection of the donor cell-free nucleic acids***…." Claim 1 of the '497 Patent recites "[a] method of ***detecting donor-specific circulating cell-free nucleic acids in a solid organ transplant recipient***." And claim 1 of the '607 Patent recites "[a] method of ***quantifying kidney transplant-derived circulating [cfDNA] in a human kidney transplant recipient***."

B. Stanford's Patent Claim Strategy

29. As I will explain in more detail below, the claims all recite using the same general combination of techniques for the recited detection of quantities of cell-free donor nucleic acids in a transplant recipient's body, and its correlation to transplant status. The claimed methods are a combination of four steps, recited at a high level, as follows:

a. Obtaining/providing a biological sample containing cfDNA from a transplant recipient

- *See* '652 Patent Claim 1(a) ("providing a sample comprising [cfDNA]")

- *See* '497 Patent Claim 1(c) ("obtaining a biological sample")[2]
- *See* '607 Patent Claims 1(a) and (b) ("providing a plasma sample" and "extracting circulating [cfDNA]")

b. Genotyping the transplant donor and/or recipient to establish profiles of genetic polymorphisms (or SNPs or distinguishable markers)

- *See* '652 Patent Claim 1(b) ("obtaining a genotype … to establish a polymorphism profile")[3]
- *See* '497 Patent Claims 1(a) and (b) ("genotyping … to obtain a SNP profile")[4]
- *See* '607 Patent Claim 1(c) ("performing a selective amplification of [SNPs] … by [PCR]"); Claim 1(f) ("using markers distinguishable between said [recipient and donor] [that] comprise [SNPs]")[5]

c. Performing multiplex or high-throughput sequencing of the cfDNA to detect the genotyped polymorphisms (or SNPs)

- *See* '652 Patent Claim 1(c) ("multiplex sequencing of the [cfDNA] in the sample followed by analysis of the sequencing results using the polymorphism profile")[6]
- *See* '497 Patent Claim 1(d) ("determining an amount of donor-specific [cfDNA] by … high-throughput sequencing or

[2] Dependent '497 patent claims 2, 9, 12, and 27 recite similar sampling and transplant-related limitations.
[3] Dependent '652 patent claims 2 and 11 recite similar genotyping and polymorphism-related limitations.
[4] Dependent '497 patent claims 6, 15, 17, and 25 recite similar genotyping and polymorphism-related limitations.
[5] Dependent '607 patent claims 2-5 recite similar genotyping and polymorphism-related limitations.
[6] Dependent '652 patent claim 6 recites similar sequencing-related limitations.

[dPCR]")[7]

- *See* '607 Patent Claims 1(d-e) ("performing a high throughput sequencing reaction … compris[ing] … sequencing-by-synthesis … [and] … providing sequences from said high throughput sequencing reaction")

d. Quantifying the transplant cfDNA in the sample using the genetic differences in the sequences

- *See* '652 Patent Claim 1(d) ("determining a quantity of [transplant cfDNA] based on the detection of [donor and recipient cfDNA] by the multiplexed sequencing")

- *See* '497 Patent Claim 1(d) ("determining an amount of [transplant cfDNA] … by detecting a homozygous or a heterozygous SNP within the [transplant cfDNA]")

- *See* '607 Patent Claim 1(f) ("quantifying an amount of [transplant cfDNA] … using markers distinguishable between … recipient and … donor"

C. The Patent Claims[8]

1. The '652 Patent

30. The '652 Patent has 16 claims. They include one independent claim (Claim 1) and 15 dependent claims (Claims 2-16). I understand that each of

[7] Dependent '497 patent claims 3-5, and 10 recite similar sequencing-related limitations.

[8] I understand that the asserted claims in this action against Natera and/or Eurofins include the following: '652 Patent claims 1-4, 6, 11-12, and 14-15; '497 Patent claims 1-6, 9-10, 12, 15, 17, 19-21, 23, 25, and 27; and '607 Patent claims 1-10. For purposes of this analysis, I opine only on these asserted patent claims. I do not offer opinions on the unasserted dependent claims of any of the Patents.

dependent claims 2-16 incorporate by reference the requirements of claim 1 and add further limitations as recited in each dependent claim, respectively.

31. Representative claim 1 of the ’652 Patent recites (emphasis added):

> 1. A method for detecting transplant rejection, graft dysfunction, or organ failure, the method comprising:
>
> (a) ***providing a sample comprising cell-free nucleic acids*** from a subject who has received a transplant from a donor;
>
> (b) ***obtaining a genotype*** of donor-specific polymorphisms or a genotype of subject-specific polymorphisms, or obtaining both a genotype of donor-specific polymorphisms and subject-specific polymorphisms, ***to establish a polymorphism profile*** for detecting donor cell-free nucleic acids, wherein ***at least one single nucleotide polymorphism (SNP) is homozygous*** for the subject if the genotype comprises subject-specific polymorphisms comprising SNPs;
>
> (c) ***multiplex sequencing*** of the cell-free nucleic acids in the sample followed by analysis of the sequencing results using the polymorphism profile to detect donor cell-free nucleic acids and subject cell-free nucleic acids; and
>
> (d) ***diagnosing, predicting, or monitoring a transplant status*** or outcome of the subject who has received the transplant ***by determining a quantity of the donor cell-free nucleic acids*** based on the detection of the donor cell-free nucleic acids and subject cell-free nucleic acids by the multiplexed sequencing, ***wherein an increase in the quantity of the donor cell-free nucleic acids over time is indicative of transplant rejection, graft dysfunction or organ failure***, and wherein sensitivity of the method is greater than 56% compared to sensitivity of current surveillance methods for cardiac allograft vasculopathy (CAV).

32. The dependent claims of the ’652 Patent are substantially similar to claim 1, and are all linked to the same natural phenomenon of donor-derived cell-

free nucleic acids in a transplant recipient's body, as well as the same natural phenomenon of genetic polymorphisms in those cell-free nucleic acids. These dependent claims are discussed in more detail in section VII.A. below.

2. The '497 Patent

33. The '497 Patent has 33 claims. They include one independent claim (Claim 1) and 32 dependent claims (Claims 2-32). I understand that each of dependent claims 2-32 incorporate by reference the requirements of claim 1, or depend from other claims that incorporate by reference the requirements of claim 1, and add further limitations as recited in each dependent claim, respectively.

34. Representative claim 1 of the '497 Patent recites (emphasis added):

> 1. A method of detecting donor-specific circulating cell-free nucleic acids in a solid organ transplant recipient, the method comprising:
>
> (a) ***genotyping*** a solid organ transplant donor ***to obtain a single nucleotide polymorphism (SNP) profile*** of the solid organ transplant donor;
>
> (b) ***genotyping*** a solid organ transplant recipient ***to obtain a SNP profile*** of the solid organ transplant recipient, wherein the solid organ transplant recipient is selected from the group consisting of: a kidney transplant, a heart transplant, a liver transplant, a pancreas transplant, a lung transplant, a skin transplant, and any combination thereof;
>
> (c) ***obtaining a biological sample*** from the solid organ transplant recipient after the solid organ transplant recipient has received the solid organ transplant from the solid organ transplant donor, wherein the biological sample is selected from the group consisting of blood, serum and plasma, and ***wherein the biological sample comprises***

circulating cell-free nucleic acids from the solid organ transplant; and

(d) ***determining an amount of donor-specific circulating cell-free nucleic acids*** from the solid organ transplant in the biological sample by ***detecting a homozygous or a heterozygous SNP*** within the donor-specific circulating cell-free nucleic acids from the solid organ transplant in at least one assay, wherein the at least one assay comprises ***high-throughput sequencing or digital polymerase chain reaction (dPCR)***, and wherein the at least one assay detects the donor-specific circulating cell-free nucleic acids from the solid organ transplant when the donor-specific circulating cell-free nucleic acids make up at least 0.03% of the total circulating cell-free nucleic acids in the biological sample.

35. The dependent claims of the '497 Patent are substantially similar to claim 1, and are all linked to the same natural phenomenon of donor-derived cell-free nucleic acids in a transplant recipient's body as well as the same natural phenomenon of genetic polymorphisms in those cell-free nucleic acids. These dependent claims are discussed in more detail in section VII.A. below.

3. The '607 Patent

36. The '607 Patent has 10 claims. They include one independent claim (Claim 1) and nine dependent claims (Claims 2-10). I understand that each of dependent claims 2-10 incorporate by reference the requirements of claim 1 (or depend from other claims that incorporate by reference the requirements of claim 1), and add further limitations as recited in each dependent claim, respectively.

37. Representative claim 1 of the '607 Patent recites (emphasis added):

1. A method of quantifying kidney transplant-derived circulating cell-free deoxyribonucleic acids in a human kidney transplant recipient, said method comprising:

(a) ***providing a plasma sample*** from said human kidney transplant recipient, wherein said human kidney transplant recipient has received a kidney transplant from a kidney transplant donor, wherein said plasma sample from said human kidney transplant recipient ***comprises kidney transplant-derived circulating cell-free deoxyribonucleic acid and human kidney transplant recipient-derived circulating cell-free deoxyribonucleic acid***;

(b) ***extracting circulating cell-free deoxyribonucleic acid from said plasma sample*** from said human kidney transplant recipient in order to obtain extracted circulating cell-free deoxyribonucleic acid, wherein said extracted circulating cell-free deoxyribonucleic acid comprises said kidney transplant-derived circulating cell-free deoxyribonucleic acid and human kidney transplant recipient-derived circulating cell-free deoxyribonucleic acid;

(c) ***performing a selective amplification of target deoxyribonucleic acid sequences***, wherein said selective amplification of said target deoxyribonucleic acid sequences is of said extracted circulating cell-free deoxyribonucleic acid, wherein said selective amplification of said target deoxyribonucleic acid sequences amplifies a plurality of genomic regions ***comprising at least 1,000 single nucleotide polymorphisms***, wherein said at least 1,000 single nucleotide polymorphisms comprise ***homozygous*** single nucleotide polymorphisms, ***heterozygous*** single nucleotide polymorphisms, or both homozygous single nucleotide polymorphisms and heterozygous single nucleotide polymorphisms, and wherein said selective amplification of said target deoxyribonucleic acid sequences is ***by polymerase chain reaction (PCR)***;

(d) ***performing a high throughput sequencing reaction***, wherein said high throughput sequencing reaction comprises performing a ***sequencing-by-synthesis*** reaction on said selectively-amplified target deoxyribonucleic acid sequences from said extracted circulating cell-free deoxyribonucleic acid, wherein said

sequencing-by-synthesis reaction has a sequencing error rate of less than 1.5%;

(e) ***providing sequences from said high throughput sequencing reaction***, wherein said provided sequences from said high throughput sequencing reaction comprise said at least 1,000 ***single nucleotide polymorphisms***; and

(f) ***quantifying an amount of said kidney transplant-derived circulating cell-free deoxyribonucleic acid*** in said plasma sample from said human kidney transplant recipient to obtain a quantified amount, wherein said quantifying said amount of said kidney transplant-derived circulating cell-free deoxyribonucleic acid in said plasma sample from said human kidney transplant recipient comprises ***using markers distinguishable between said human kidney transplant recipient and said kidney transplant donor***, wherein said markers distinguishable between said human kidney transplant recipient and said kidney transplant donor ***comprises single nucleotide polymorphisms*** selected from said at least 1,000 single nucleotide polymorphisms identified in said provided sequences from said high throughput sequencing reaction, and wherein said quantified amount of said kidney transplant-derived circulating cell-free deoxyribonucleic acid in said plasma sample from said human kidney transplant recipient comprises at least 0.03% of the total circulating cell-free deoxyribonucleic acid from said plasma sample from said human kidney transplant recipient.

38. The dependent claims of the ’607 Patent are substantially similar to claim 1, and are all linked to the same natural phenomenon of donor-derived cfDNA in a transplant recipient’s body as well as the same natural phenomena of SNPs in the cf DNA. These dependent claims are discussed in more detail in section VII.A. below.

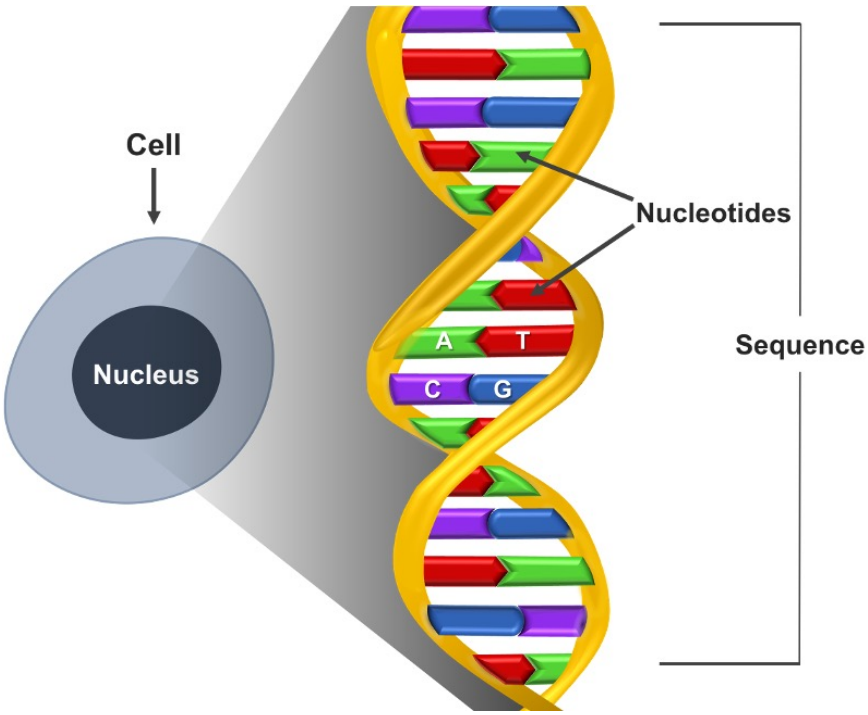
VI. BACKGROUND ON THE TECHNOLOGY

A. Nucleic Acids And Genes

39. Nucleic acids are biomolecules that are present in all living things. They store and encode the information living cells need to grow and function. Although there are many different subtypes of nucleic acids, naturally occurring, biologically active nucleic acids fall into two main classes: deoxyribonucleic acid ("DNA") and ribonucleic acid ("RNA").

40. DNA naturally exists as two long strands in a double-helical structure. Each strand is comprised of a sequence of "nucleotides" connected at regular intervals along the strand. On each nucleotide is one of several different types of "bases." There are four natural types of bases that occur in DNA: adenine ("A"), cytosine ("C"), guanine ("G") and thymine ("T"), which pair together complementarily (A to T, and C to G) to hold the two strands of DNA together. Each base, with its corresponding piece of the backbone of the DNA strand, is referred to as a "nucleotide."

41. An individual's DNA is organized into chromosomes, which reside in the nucleus of eukaryotic cells. This, and the general structure of DNA, is depicted in the figure below:

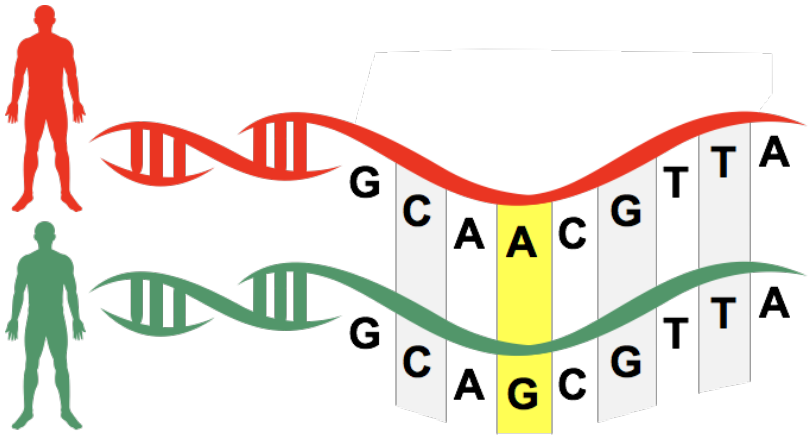


B. Genes, Polymorphisms, Homozygosity, And Heterozygosity

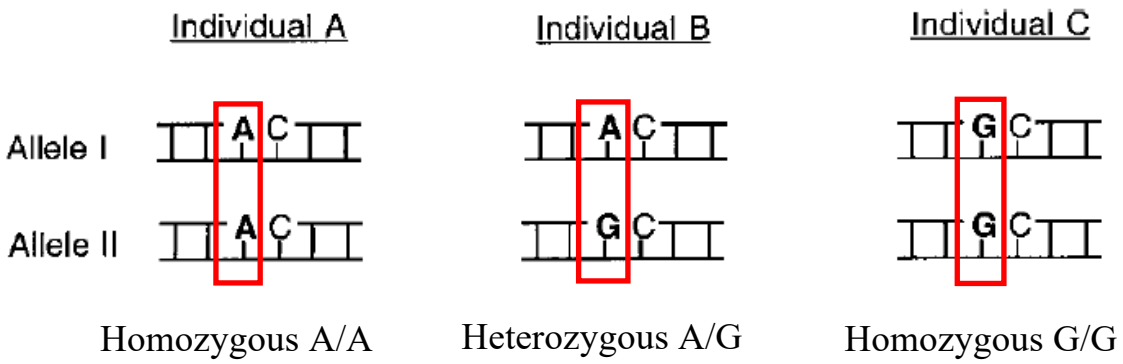
42. It is the unique sequence of bases in DNA that creates a code, or blueprint, for the production of other molecules such as proteins, which form cells' structures and make them function.

43. All individuals (apart from identical twins) have unique DNA sequences that genetically distinguish one person from another. The genetic constitution of an individual is called a "genotype." The differences in genetic sequences that give rise to everyone's unique genotype are called "polymorphisms." A "single nucleotide polymorphism," or "SNP," occurs where the nucleotide at a specific location may differ between different individuals. SNPs occur at about one

base out of 1,000 and there are about 3,000,000 differences between any two individuals. Such variations can be used to distinguish one individual from another. An example of a SNP is depicted in the figure below:



44. Humans have two copies of each of their chromosomes – one inherited from the mother, and the other from the father. Zygosity describes the similarities or differences in polymorphisms between the maternally inherited and paternally inherited copies. An organism is "homozygous" for a particular gene or polymorphism if its maternally inherited and paternally inherited copies have the same sequence. An organism is "heterozygous" for a particular gene or polymorphism if it has different sequences on its maternally inherited and paternally inherited copies, respectively. An example of a homozygous and heterozygous SNPs, reproduced from Figure 1 of B0074-B0079 (Nordstrom 2000) at B0075, is shown below (red boxes and "homozygous" and "heterozygous" labels added):



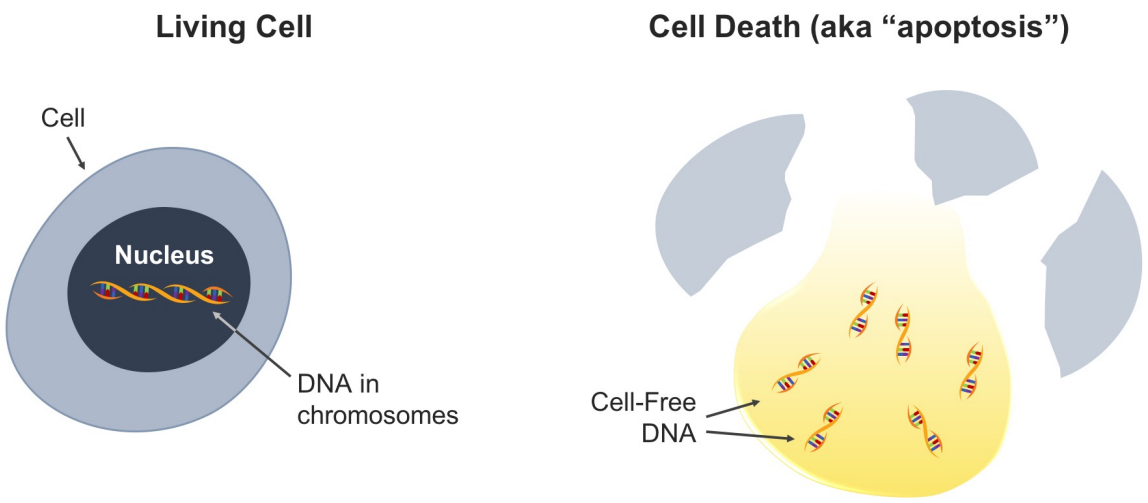
45. There is on average one SNP in every ~1,000 nucleotides in humans, and two individuals on average vary in approximately ~3,000,000 SNPs. B0001-B0024 ('652 Patent) at B0016, 13:41-44; B0481-B0580, (Vogel 2009) at B0574. Given the diversity in SNPs among individuals, scientists have exploited SNPs for many different purposes, including identifying people in forensic analyses, or determining patterns of SNPs associated with specific populations or disease states.

46. In addition to SNPs, there are other types of genetic variations. For example, "VNTRs," or variable numbers of tandem repeats, are short sequence repeats that are polymorphic due to variable numbers of short sequence repeat units as between individuals. While "dinucleotide repeats" are most common VNTRs can also be tri-, tetra-, or penta-repeats (often called microsatellites). VNTRs with longer repeat units of 15-500 nucleotides are called "minisatellites." "Insertion" and "deletion" polymorphisms are due to the presence of additional sequences or

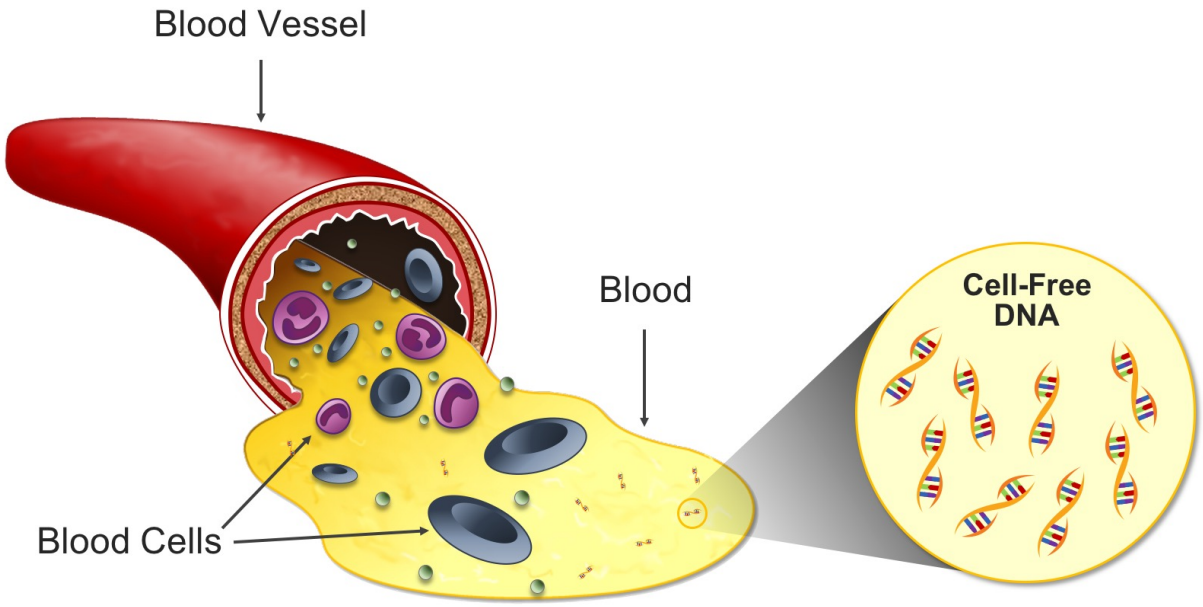
absence of normally present sequences, respectively. Regions of DNA that have many of the polymorphisms described above are commonly known as "hypervariable regions."

C. Cell-Free Nucleic Acids

47. As discussed above, human DNA is most commonly found in the nuclei of cells in chromosomes. But DNA also can naturally exist outside of the cell. What scientists typically refer to as "cell-free" nucleic acids (and in the case of DNA, cell-free DNA or "cfDNA") are small fragments of nucleic acids that are released from cells into the bloodstream. They circulate freely in the blood and other bodily fluids. These cell-free nucleic acids are usually produced when cells die, in a process called "apoptosis." In apoptosis, cells undergo various biochemical processes that cause the nucleic acids, including DNA of the cells' chromosomes, to be cleaved and fragmented before being released into the circulation. They are eventually cleared out of the body with other bodily waste products. The process of release of cfDNA from apoptotic cells is illustrated below:



48. Because cells are constantly dying and being replaced throughout the body, cell-free nucleic acids, including cfDNA, are constantly released into and circulating through the bloodstream or other body fluids, such as urine, before being cleared out of the body, as depicted below:



D. Cell-Free Nucleic Acids Having Sequences Different From A Patient's Genotype Occur Naturally In An Individual's Circulation

49. In certain diseases and conditions, nucleic acids including cfDNA that are foreign or abnormal to a person are present in the person's body. This occurs in a variety of contexts where a source of foreign DNA, or cells with different genotypes, reside in the host. Some of these are discussed below.

1. Cell-Free Nucleic Acids In Organ Transplant

50. For example, a transplanted organ can be a source of a genotype that is different from that of the transplant recipient. If an organ from a donor is transplanted into the body of a recipient, then the donor's DNA, with its unique genotype, is also transplanted into the body of the recipient. Because a transplanted organ's cells naturally die at a certain rate, just like all of the cells in the rest of a recipient's body, cell-free nucleic acids from the donated transplant organ will be released into, and circulate in, the recipient's blood along with the cell-free nucleic acids released from the cells in all of the rest of the recipient's body. This gives rise to two different genotypes being present in the same individual recipient's blood.

51. While many transplanted organs remain healthy, sometimes the body of the transplant recipient will reject the transplanted organ, or the transplanted organ will fail to survive for some other reason in the transplant recipient's body. Where

rejection occurs, the recipient's immune system attacks the transplanted organ and kills its cells, causing the cells to die more rapidly than they normally would. In transplant failure for other reasons, the cells of the transplanted organ die more rapidly than normal for a variety of reasons. Whether the transplant is rejected or otherwise fails, the transplanted organ's cells release more cell-free nucleic acids into the recipient's circulation than would be expected from a healthy transplant. These elevated levels of cell-free nucleic acids originating from the donor organ is the natural result of biological processes occurring in the transplant recipient's body. As of the priority date, it was well known that donor-derived DNA is present in a transplant recipient. *See, e.g.*, B0001-B0024 ('652 Patent) at B0013, 7:53-8:22; B0080-B0081 (Lo 1998) at B0080.

2. Cell-Free Nucleic Acids In Pregnancy

52. Another example occurs in the case of pregnancy, as the fetus growing in the body of its mother has a genotype that is different from its mother. As of the priority date, it was well-known that fetal DNA is naturally found as cfDNA in the pregnant mother's bloodstream, as the result of release from fetal cells that die while the fetus is growing in the mother's body. *See, e.g.*, B0001-B0024 ('652 Patent) at B0013, 7:19-23. This also gives rise to two different genotypes being present in an individual pregnant woman's blood.

53. Cell-free nucleic acids released from fetal cells into the mother's blood will reflect the fetus's genotype. Some fetuses have certain conditions such as disease, or an extra or missing chromosome ("aneuploidy," and for example, Down Syndrome). These are conditions of the fetus that are detectable while the fetus is in the womb through the fetal cell-free nucleic acids being released into the mother's bloodstream. For example, if the fetus is aneuploid, the fetal cfDNA from the extra or missing chromosome will be present in the mother's blood in greater or lesser quantity than would be expected if the fetus were not aneuploid. This observable difference, like the observable difference in a transplant recipient, is the natural result of biological processes occurring in the pregnant woman's body.

3. Cell-Free Nucleic Acids In Cancer

54. Yet another example occurs in the case of cancer, where a tumor with a unique genotype as a result of mutations grows in the body of a cancer patient. The sequences of the DNA in the tumor cells differ from those of the cells in the rest of the cancer patient, which do not have the mutations unique to the tumor. Because these tumor cells die and release their cell-free nucleic acids into the cancer patient, where they circulate together with the cell-free nucleic acids released from the non-cancerous cells in the body, the tumor acts as a source of an additional, different genotype present in the cancer patient. As of the priority date, it was well known that

the DNA sequences of tumor cells differ from the DNA of a cancer patient. *See, e.g.*, B0001-B0024 ('652 Patent) at B0012, 6:67-7:5.

55. Cancer cell-free nucleic acids can serve as a source for monitoring cancer as well. The cancerous tumor cells have genetic mutations not present in the patient's normal cells, and as the tumor cells die (either because the patient's immune system attacks and kills them or they are starved for resources due to the expanding tumor mass) they release cell-free nucleic acids into the cancer patient's circulation. As a result, these genetically unique tumor cells release cell-free nucleic acids with genomes that are not normally present in the cancer patient's circulation. This phenomenon becomes more pronounced as tumors grow, and again, is the natural result of biological processes occurring in the cancer patient's body.

4. Cell-Free Nucleic Acids In Infectious Disease

56. In the case of infectious disease, bacteria or viruses invading a person's body have genotypes that are different than that of the person they are infecting. The bacteria or viruses are destroyed by the person's immune system, causing the release of bacterial or viral cell-free nucleic acids into the person's blood. Those bacterial or viral cell-free nucleic acids will have sequences that differ from those of the infected person, such that they act as a source of additional, different genotypes present in the infected person's body. As of the priority date, this, too, was well-known and understood by POSAs. *See* section VII.C.2. below.

5. Scientists Have Long Used Cell-Free Nucleic Acids To Monitor Disease

57. In each of the above-referenced settings (transplant, pregnancy, cancer, and infectious disease), cells from the source of foreign DNA naturally die while inside of the patient and release their nucleic acids to circulate as cell-free nucleic acids in the patient's body. The resulting presence of cell-free nucleic acids with a different genotype than the patient's normal genotype is a natural result of biological processes that occur in the transplant recipient's, mother's, or cancer or infected patient's bodies, respectively. Scientists knew of these natural phenomena long before the November 6, 2009 priority date of the Patents. As described here in my declaration, and in the Patents themselves, scientists routinely applied methods for detecting and analyzing the pertinent genetic differences in samples from patients containing different genotypes in order to study conditions relating to them.

58. The Patents acknowledge that cell-free nucleic acids, including cfDNA, can be found circulating in an individual's bodily fluids, and that these circulating cell-free nucleic acids can be correlated to diseases involving cell death, or apoptosis. As the written description explains:

> Circulating, or cell-free, DNA was first detected in human blood plasma in 1948. … Since then, its connection to disease has been established in several areas. … Studies reveal that much of the circulating nucleic acids in blood arise from necrotic or apoptotic cells … and greatly elevated levels of nucleic acids from apoptosis is observed in diseases such as cancer. …

B0001-B0024 ('652 Patent) at B0012, 6:57-67 (citations omitted); *see also* Section VII.C.2. below.

VII. THE PATENT CLAIMS ARE DIRECTED TO WELL-UNDERSTOOD, ROUTINE AND CONVENTIONAL TECHNIQUES APPLIED TO OBSERVING A NATURAL PHENOMENON

A. The Claims of the Patents Are Directed to a Natural Law

59. It is my opinion that the claims of the Patents are directed to a natural law—the presence of naturally occurring donor-derived cell-free nucleic acids in the bodily fluids of transplant recipients, and the correlation of those naturally occurring cell-free nucleic acids to transplant status or rejection.

60. Claim 1 of the '652 Patent recites a "method for detecting transplant rejection, graft dysfunction, or organ failure" by "diagnosing, predicting, or monitoring a transplant status or outcome of the subject who has received the transplant by determining a quantity of the donor cell-free nucleic acids based on the detection of the donor cell free nucleic acids and subject cell-free nucleic acids by the multiplexed sequencing, wherein an increase in the quantity of the donor cell-free nucleic acids over time is indicative of transplant rejection, graft dysfunction or organ failure." Thus, claim 1 of the '652 Patent provides that *if* an increase in donor cell-free nucleic acid is detected in a bodily fluid of a transplant recipient over time, *then* organ transplant rejections can be diagnosed, predicted or monitored.

61. Claim 1 of the ’497 Patent recites a “method of detecting donor-specific circulating cell-free nucleic acids in a solid organ transplant recipient” by “determining an amount of donor-specific circulating cell-free nucleic acids from the solid organ transplant in the biological sample by detecting a homozygous or a heterozygous SNP … in at least one assay … wherein the at least one assay detects the donor-specific circulating cell-free nucleic acids from the solid organ transplant.” Thus, claim 1 of the ’497 Patent starts with donor-specific cell-free nucleic acids that are already present in a sample, and as an end result detects them.

62. Claim 1 of the ’607 Patent recites a “method of quantifying kidney transplant-derived circulating cell-free deoxyribonucleic acids in a human kidney transplant recipient” by “quantifying an amount of said kidney transplant-derived circulating cell-free deoxyribonucleic acid in said plasma sample to obtain a quantified amount.” This claim, too, starts with a certain quantity of transplant-derived cfDNA already present in a sample, and as an end result observes that quantity.

63. This presence of naturally occurring cell-free nucleic acids in the bodily fluids of transplant recipients, and its correlation to transplant status or rejection, exists apart from any human intervention. Cell death is a natural consequence of a transplanted organ’s biological response to being rejected or otherwise failing. The natural result of this biological response is an elevated level of donor-specific cell-

free nucleic acids in the transplant recipient's bodily fluids. As explained in the Patents' written description, this is a natural phenomenon that does not require any human intervention. *See, e.g.*, B0001-B0024 ('652 Patent) at B0013, 7:40-46 (stating that "as cell-free DNA or RNA often arises from apoptotic cells, the relative amount of donor-specific sequences in circulating nucleic acids should provide a predictive measure of on-coming organ failure in transplant patients for many types of solid organ transplantation including, but not limited to, heart, lung, liver, and kidney.").

64. The independent claims of the Patents are representative of each of the dependent claims—all of which are directed to a natural law. Each dependent claim relies on the method of the independent claim of each patent. In addition, all recite substantially similar methods and all are linked to observing the same natural phenomenon of donor-specific cell-free nucleic acids in a transplant recipient's circulation, or its correlation to transplant rejection.

65. Beyond the elements of the claims from which they depend, the dependent claims of the Patents recite:

- Using different types or numbers of polymorphisms ('652 Patent Claims 2 and 11; '497 Patent Claims 6, 17, and 25; '607 Patent Claims 2, 3, 4, and 5)
- Using cf-DNA ('652 Patent Claim 3; '497 Patent Claim 12)
- Using certain common modifications, error rates, or quality scores associated with multiplexed or high throughput sequencing ('652 Patent Claims 4 and 6; '497 Patent Claims 3, 4, 5, 19, 21, and 23 '607 Patent Claim 6)
- Using known PCR or amplification methods '497 Patent Claim 10)

- Using different sample types ('652 Patent Claim 12; '497 Patent Claims 2 and 27)
- Further genotyping prior to quantifying cell-free nucleic acids ('497 Patent Claim 15)
- Certain concentrations of cell-free nucleic acids in the sample ('497 Patent Claim 20; '607 Patent Claims 7, 8, 9, and 10)

66. By depending on and incorporating the phenomenon of the presence of naturally occurring donor-derived cell-free nucleic acids in the circulation of transplant recipients and the natural law correlating it to transplant rejection, without adding more, each of the dependent claims of the Patents are also directed to this natural law.

B. Each Recited Laboratory Technique Was Conventional

67. In my experience as a researcher in the genomics field in the 2009 timeframe, all of the techniques recited in the Patent claims were well understood, routine and conventional as of their November 6, 2009 filing date. Indeed, the Patents' written description states: "The practice of the present invention employs, unless otherwise indicated, conventional techniques of immunology, biochemistry, chemistry, molecular biology, microbiology, cell biology, genomics and recombinant DNA, which are within the skill of the art." B0001-B0024 ('652 Patent) at B0012, 5:36-40. I have reviewed the written description, including the portions referenced herein, and I have not seen any disclosure in which the Patents "otherwise indicate[]" that any of the laboratory techniques recited in the claims, or that the particular way they are combined in the claims, is nonconventional. To the contrary,

the written description repeatedly describes all of the recited techniques, and the claimed combinations of them, as conventional and routinely practiced using commercially available products.

68. As summarized above in section V, a POSA would understand the methods of the Patents to recite four common categories of laboratory techniques that were well-established approaches to cell-free nucleic acid detection and quantification by November, 2009:

- **Obtaining / providing a biological sample** containing cell-free nucleic acids from a transplant recipient;[9]
- **Genotyping** the transplant donor and/or recipient **to establish profiles of genetic polymorphisms (or SNPs)**;[10]
- **Performing multiplex or high-throughput sequencing** of the cell-free nucleic acids to detect the genotyped polymorphisms (or SNPs);[11] and

---

[9] For independent claims, see '652 Patent Claim 1(a) ("providing a sample comprising [cfDNA]"); '497 Patent Claim 1(c) ("obtaining a biological sample"); '607 Patent Claims 1(a) and (b) ("providing a plasma sample" and "extracting circulating [cfDNA]"). For dependent claims that recite similar sampling and transplant-related limitations, see '497 Patent Claims 2, 9, 12, and 27.

[10] For independent claims, see '652 Patent Claim 1(b) ("obtaining a genotype … to establish a polymorphism profile"); '497 Patent Claims 1(a) and (b) ("genotyping … to obtain a SNP profile"); and '607 Patent Claim 1(c) ("performing a selective amplification of [SNPs] … by [PCR]") and Claim 1(f) ("using markers distinguishable between said [recipient and donor] [that] comprise [SNPs]"). For dependent claims that recite similar genotyping and polymorphism-related limitations, see '652 Patent Claims 2 and 11; '497 Patent Claims 6, 15, 17, and 25; and '607 Patent Claims 2-5 recite similar genotyping and polymorphism-related limitations.

[11] For independent claims, see '652 Patent Claim 1(c) ("multiplex sequencing of the [cfDNA] in the sample followed by analysis of the sequencing results using the

- **Quantifying the transplant (donor-derived) cell-free nucleic acids in the sample using the genetic differences** in the sequences.[12]

69. The Patent claims identify the recited steps only at a high level of generalization. The written description identifies many different conventional ways each of the laboratory techniques recited in those steps, and the combinations of them, can be performed. However, the claims do not identify any particular approach to performing the steps or describe any unconventional performance of the common techniques already in use in 2009 as described in the written description.

70. As I explain in section VII.B.3. below, there was what those of us in the field have characterized as an explosion of technology driven in part by rapid advances in the sensitivity, accuracy, and throughput of DNA sequencing and analysis methods that had become routine and conventional for cell-free nucleic acid

---

polymorphism profile"); '497 Patent Claim 1(d) ("determining an amount of donor-specific [cfDNA] by … high-throughput sequencing or [dPCR]"); '607 Patent Claims 1(d-e) ("performing a high throughput sequencing reaction … compris[ing] … sequencing-by-synthesis … [and] … providing sequences from said high throughput sequencing reaction"). For dependent claims that recite similar sequencing-related limitations, see '652 Patent Claims 4 and 6; and '497 Patent Claims 3-5, and 10.

[12] For independent claims, see '652 Patent Claim 1(d) ("determining a quantity of [transplant cfDNA] based on the detection of [donor and recipient cfDNA] by the multiplexed sequencing"); '497 Patent Claim 1(d) ("determining an amount of [transplant cfDNA] … by detecting a homozygous or a heterozygous SNP within the [transplant cfDNA]"); and '607 Patent Claim 1(f) ("quantifying an amount of [transplant cfDNA] … using markers distinguishable between … recipient and … donor").

detection as recited in the Patent claims in the years leading up to November 2009. I know this from my own professional experience at the time, and it is corroborated by the literature.

71. For example, during the relevant 2009 time period I was a scientific advisor for Helicos Biosciences, Inc. ("Helicos"), which was co-founded by named Patent inventor Dr. Stephen Quake. A 10-K Annual Report for fiscal year 2007 from Helicos explains the state of the industry at that time, and reinforces my opinion that the laboratory techniques recited in the claims of the Patents were well understood, routine and conventional by November 2009:

> Since the development of genetic engineering techniques in the 1970s, the analysis of genetic material has become a mainstay of biological research. The first automated DNA sequencer was invented in 1986, based on technology developed by Frederick Sanger and his colleagues in 1975, which is commonly referred to as Sanger sequencing. Subsequent versions of commercial DNA sequencers have increased the speed of DNA sequencing by 3,000 fold, making possible the Human Genome Project. In 1996 the first commercial microarray was introduced and enabled a new era of RNA analysis by measuring gene expression across many genes in a single experiment. Subsequent versions of the commercial microarrays including DNA and RNA have significantly increased the amount of information per run and provided selected SNPs of the whole human genome on a single chip, enabled large scale genome-wide SNP association studies and have been commercialized for several diagnostic applications. Today, manufacturers of systems, supplies and reagents for performing genetic analysis, which includes DNA sequencing, genotyping, and gene expression analysis, serve a worldwide market of approximately $5 billion, according to Strategic Directions International. Strategic Directions International estimates that DNA sequencing serves approximately 17% of this demand for genetic analysis.

B0581-B0693 (Helicos 10-K) at B0587. As named inventor Dr. Quake's company acknowledged in its SEC filings, by the time of the Helicos 10-K in 2007, systems and methods for DNA sequencing and genotyping had become well understood, routine and conventional, and served a worldwide market of approximately $5 billion. In its 10-K, Helicos also acknowledged "**established** genomic analysis technologies," including sequencing, "next generation sequencing" (which as explained in section VII.B.3. below is a term that for pertinent purposes is used interchangeably with multiplex sequencing or high-throughput sequencing), and genotyping (B0581-B0693 Helicos 10-K) at B0589:

Comparison of established genomic analysis technologies

Analysis	Description	Technology	Advantages	Disadvantages
Sequencing	Determination of the complete sequences of DNA or RNA molecules	Automated Sanger-based instruments	• Comprehensive sequence information • Industry standard technology	• High cost • Low throughput • Complex sample preparation
Next Generation Sequencing	Determination of the complete sequences of DNA and RNA molecules	Ensemble-on-bead based technologies	• Comprehensive sequence information • High throughput/lower cost per sequence • Seen as "upgrade" to Sanger sequencers	• Complex sample preparation • Limited scalability • High cost of sample preparation • Limited quantitation
Gene Expression Analysis	Detection and quantitation of RNA to determine gene expression levels	DNA arrays on chips or beads	• Can perform genome-wide analysis of expressed genes • Widely available	• Low sensitivity • Relative quantitation • Limited sequence information • Limited to known genomic sequences • Biased based on templates
		RT-PCR	• Absolute quantitation • Highest sensitivity	• Higher cost per gene than arrays • Labor intensive • Not scalable
Genotyping	Analysis of short specific sequences within genomic DNA to look for known variants	DNA arrays on chips or beads	• High throughput/low cost per genotype • Can be applied to large numbers of samples	• Provides only limited genomic information • Only interrogates known sequence variants
		RT-PCR	• Higher sensitivity than arrays	• Provides very limited genomic information • Higher cost per genotype than arrays • Biased based on templates

72. The Patents' written description itself describes the recited methods as conventional. I summarize portions of the Patents' written description identifying the conventional nature of the recited method steps in turn below.

1. Obtaining Or Providing a Biological Sample Containing Cell-Free Nucleic Acids From a Transplant Recipient

73. '652 Patent Claim element 1(a), '497 Patent Claim element 1(c), and '607 Patent Claim element 1(a) recite providing or obtaining a sample comprising cell-free nucleic acids from a transplant recipient. Techniques for providing or

obtaining such samples were routine long before the 2009 filing of the Patents, which the Patents' written description confirms.

74. Obtaining and studying biological samples of cell-free nucleic acids from an individual was routine and conventional by 2009. As the Patents' written description explains, circulating or cell-free DNA was already well characterized in the prior art. B0001-B0024 ('652 Patent) at B0012, 6:57-58 ("[c]irculating, or cell-free, DNA was first detected in human blood plasma in 1948.").

75. Techniques for obtaining biological samples were routine, conventional and well-established by 2009. For example, venipuncture, or the science of making punctures in a patient's vein with a needle in order to draw blood from the vein, was well known in the prior art. *Id*. at B0014, 10:11-12 (emphasis added) ("To obtain a blood sample, ***any technique known in the art may be used***, e.g., a syringe or other vacuum suction device ...").

76. The '607 Patent, at Claim element 1(b), further recites extracting cfDNA from the plasma sample. This also was routine by 2009, with numerous options for commercial kits available to carry it out.

77. The Patents do not purport to claim an improvement over or nonconventional application of the traditional laboratory techniques for obtaining or providing biological samples comprising cell-free nucleic acids. I also have seen no indication in the claims or written description of the Patents that demonstrates an

improvement over or nonconventional application of those conventional techniques or a new or unique way of carrying them out.

2. Genotyping The Transplant Donor And/Or Recipient To Establish Profiles Of Genetic Polymorphisms (Or SNPs), Including By Comprising Selectively Amplifying At Least 1,000 SNPs By PCR And Sequencing

78. The ’652 and ’497 Patent claims also recite “genotyping” to establish “profiles” of “SNPs” or “polymorphisms,” including homozygous and/or heterozygous variants. *See* ’652 Patent claim element 1(b) and ’497 Patent claim elements 1(a) and (b). The ’607 Patent, at elements 1(c)-(e), claims a generalized and conventional method for genotyping, reciting “performing a selective amplification of target [DNA] sequences” that “amplifi[es] … at least 1,000 [homozygous and/or heterozygous] [SNPs]” by [PCR]” (Claim 1(c)), followed by “high throughput sequencing” (Claim 1(d)), and then “providing sequences” of the “[SNPs]” (Claim 1(e)). These claim limitations describe what a POSA in 2009 would recognize as a common approach to genotyping, including based on targeted PCR amplification and sequencing.

79. A POSA in 2009 would recognize that the Patents claim these steps at a high level without purporting to describe any improvements to then-existing genotyping techniques. As the written description explains, numerous techniques were known as of the November 6, 2009 filing date—including numerous

commercially available options—for accomplishing the genotyping and SNP/polymorphism profiles recited in the claims.

80. The written description explains that:

> Genotyping of the transplant donor and/or the transplant recipient ***may be performed by any suitable method known in the art including those described herein such as sequencing, nucleic acid array or PCR***. … In some embodiments, the marker profile is a polymorphic marker profile. Polymorphic marker profile may comprise one or more single nucleotide polymorphisms (SNP's) ….

B0001-B0024 ('652 Patent) at B0019, 20:31-44 (emphasis added); *see also id.* at B0014, 9:8-14 ("Detection, identification and/or quantitation of the donor-specific markers (e.g. polymorphic markers such as SNPs) can be performed using real-time PCR, chips (e.g., SNP chips), ***high-throughput shotgun sequencing*** of circulating nucleic acids (e.g. cell-free DNA), ***as well as other methods known in the art including the methods described herein***.") (emphasis added).

81. Moreover, literature from as early as 2003 confirms that as of that time, approximately 2.8 million SNPs had been identified and were available in public databases. *See, e.g.*, B0086-B0093, (HapMap Project) at B0089; *see also* B0094-B0155 (International Human Genome Sequencing Consortium 2001) at B0094, B0098-B0099, B0101, B0145, B0148; B0156-B0163 (Human Genome Project NIH FAQ) at B0159; B0164-B0183 (1000 Genome Project) at B0164. Thus, the written description acknowledges, and the literature confirms (as also discussed in more detail below), that routine techniques including sequencing, nucleic acid array and

PCR were available and used for genotyping up to thousands (or more) of SNPs by 2009.

(a) PCR, Including Selective Amplification, Was a Well-Understood, Routine and Conventional Method For Genotyping By 2009

82. PCR, including selective amplification, was a well-understood, routine and conventional method for genotyping by 2009. The written description describes PCR-based commercial products that were available at the time of filing to perform the claimed methods under the heading "Genotyping":

> For example, after genotyping a transplant donor and transplant recipient, ***using existing genotyping platforms known in the art including the one described herein***, one could identify approximately 1.2 million total variations between a transplant donor and transplant recipient. Usable SNPs may comprise approximately 500,000 ***heterozygous*** donor SNPs and approximately 160,000 ***homozygous*** donor SNPs. ***Companies*** (such as Applied Biosystems, Inc.) ***currently offer both standard and custom-designed TaqMan probe sets*** for SNP genotyping that can in principle target any desired SNP position for a ***PCR-based assay*** (Livak, K. L., Marmaro, J., Todd, J. A., Nature Genetics, 9, 341-342 (1995); De La Vefa, F. M., Lazaruk, K. D., Rhodes, M. D., Wenz, M. H., Mutation Research, 573, 111-135 (2005)). With such a large pool of potential SNPs to choose from, a ***usable subset of existing or custom probes*** can be selected to serve as the probe set for any donor/recipient pair."

B0001-B0024 ('652 Patent) at B0016, 13:51-67 (emphasis added). As noted by the dates of the references cited in this portion of the written description, these PCR-based tools for genotyping SNPs, including for genotyping up to hundreds of thousands of SNPs, were commercially available as far back as 1995.

83. By 2000, PCR had been universally adopted as a standard tool in applications of molecular biology, and it was already successful in the simultaneous detection of thousands of SNPs. *See e.g.*, B1291-B1300 (Germer 2000) at B1291. In fact, by 2009, PCR was also being widely used to measure SNPs in cell-free DNA. *See, e.g.*, B0191-B0212 (Mei 2005) at B0196; B0213-B0223 (Lichtenstein 2001) at B0214. PCR was a well-established, conventional and indispensable tool for genetic testing that was routinely used to target and amplify specific, pre-selected genes for further study. *See e.g.*, B0001-B0024 ('652 Patent) at B0016, 14:29-67.

84. By 2009, PCR also was commonly used to amplify nucleic acids, including to selectively amplify nucleic acids, in order to obtain sufficient amounts of them for "sequencing." As discussed in more detail below, nucleic acid sequencing is the process of determining the nucleic acid sequence, or the order of nucleotide bases, in a nucleic acid. "High-throughput sequencing" as recited in '607 Patent Claim element 1(d) (and '497 Patent Claim element 1(d)) is a term commonly understood in the art of nucleic acid analysis to refer to sequencing technologies that sequence multiple DNA molecules in parallel. As discussed in section VII.B.3. below, this, too, had been widely accepted and had become a routine tool for nucleic acid analysis by 2009.

85. For example, by 2008, a product called RDT 1000 from the company RainDance Technologies combined "targeted sequencing" by PCR of "hundreds to

thousands of genomic loci." B0224-B0225 (RainDance 2008) at B0224. As the manufacturer indicated, "[t]he RDT 1000 fits seamlessly into any targeted sequencing workflow," and its "'open' design easily integrates with all next-generation DNA sequencing platforms." *Id.*

86. Several other products for selective amplification were on the market by November 6, 2009, including from Agilent Technologies, Roche, and Febit. *See, e.g.*, B0226-B0243 (Volkerding 2009) at B0233. For example, in a Review article in 2009, Volkerding *et al.* described several "targeted genomic resequencing" products available at the time that combine selectively amplifying "polymorphisms" and other genetic regions of interest, and then sequencing them using multiplex or high-throughput sequencing instruments. *Id.* Volkerding *et al.* identified using "[c]apture probes [] immobilized on a solid surface," as provided in the Roche NimbleGen product, Agilent products, and Febit products, or capture probes that are "used in solution," as in other Agilent products. *Id.* These arrays could capture up to "350,000" sequences. *Id.* Using these products, the target DNA for selective amplification would be captured by probes on an array or in solution, then "the enriched DNA amplified by PCR before NGS [*e.g.,* high-throughput sequencing] library preparation." *Id.*

87. I also have significant experience using these standard PCR techniques for genotyping prior to 2009. Starting in 1992, I ran a large program at the Salk

Institute using PCR to assay polymorphic genetic markers to produce a map of human chromosome 11, a project that required tens of thousands of multiplexed PCR reactions. Between 1994 and 1996, while at Stanford University, I was responsible for the design and conduct of hundreds of thousands of PCR reactions for mapping the entire human genome and as part of developing a strategy for sequencing regions of human chromosomes 4 and 21. After joining The Institute for Genomic Research in 1997, I was responsible for projects involving hundreds of thousands of PCR reactions for genome sequencing and microarray analysis. And after joining the faculty at the Dana-Farber Cancer Institute and Harvard School of Public Health, I led large-scale projects analyzing gene expression in cancer and establishing high-throughput genome sequencing that required PCR and other amplification reactions from diverse biological samples, including cell lines, human tissue samples, circulating tumor cells, and circulating cell-free DNA. In all of these projects, I used standard and routine PCR applications for genotyping.

(b) Arrays Were a Well-Understood, Routine and Conventional Method For Genotyping By 2009

88. Use of arrays was also a well-understood, routine and conventional method for genotyping by 2009. The written description states that genotyping SNPs could also be executed using arrays, including commercially available arrays manufactured by Illumina, Inc., ParAllele, and Affymetrix. B0001-B0024 ('652 Patent) at B0013, 8:55-60; B0018, 17:40-18:53.

89. DNA microarrays were developed in the 1990s and essentially allow the detection of many unique fragments of DNA in parallel. DNA microarrays are based on the simple concept of printing specific sequences of DNA fragments onto different spots on a glass and detecting complementary sequences of DNA that bind to each spot. Using DNA fragments containing SNPs was a routine application of this technology. Indeed, Affymetrix developed the first commercial microarray system (the Affymetrix GeneChip®) in 1994, and in 1999, released an array called GeneChip HuSNP, consisting of a panel of 1,200 SNPs printed onto the array.

90. After the human genome was sequenced and the HapMap project was launched in 2002, there was intense interest in developing SNP arrays that could measure enough SNPs to genotype the entire genome of an individual on a single array. The NIH selected ParAllele and Illumina as two of the five US HapMap project participants responsible for genotyping. *See* B0244-B0250 (Ogren 2003) at B0248. In connection with this, ParAllele partnered with Affymetrix to develop a universal array system that could be used with any custom SNP. Illumina similarly adapted its GoldenGate® Assay to be used with a universal BeadChip to analyze large numbers of SNPs. The Illumina and ParAllele/Affymetrix arrays became standard SNP genotyping platforms by the time the HapMap project was launched in 2002.

91. Applied Biosystems TaqMan probe sets were first described in 1995 by Kenneth J. Livak and Jeffrey Marmaro. *See* B0184-B0185 (Livak 1995) at B0184-B0185. This technology uses fluorescently-labeled probes targeting specific SNPs as a simple variation of the primers used in standard PCR. Because the amplification of DNA can be detected and quantified in real time, i.e., as the PCR assay progresses, this methodology is often called "real-time PCR" or "quantitative PCR." Thus, methods of SNP-typing using fluorescent TaqMan probes and real-time PCR were established almost a decade before the filing date of the Patents.

92. I have significant personal experience using these standard and routine arrays for genotyping prior to 2009. For example, I was involved in the early development of DNA microarrays while at Stanford University and was hired by TIGR in 1997 in large part to establish and lead a large-scale effort to establish and use microarrays to understand gene function. My group developed DNA microarray technology for profiling gene expression in humans, mouse, rat, zebrafish, many bacterial species, other mammalian species, plant species including *Arabidopsis*, maize, tomato, potato, rice, and eukaryotic parasites such as *Plasmodium falciparum,* the causative agent of malaria. As part of my recruitment to Dana-Farber in 2005, I negotiated access to commercial microarrays including Affymetrix GeneChips™ and Illumina arrays and I was involved in projects that used both gene

expression and SNP genotyping data. In all of these applications, I used standard and routine commercially available array applications for genotyping.

(c) Sequencing Was a Well-Understood, Routine and Conventional Method For Genotyping By 2009

93. Sequencing was also a well-understood, routine and conventional method for genotyping by 2009. For example, the written description states that:

> Genotyping donor and recipient nucleic acids, and/or detection, identification and/or quantitation of the donor-specific nucleic acids after transplantation (e.g. polymorphic markers such as SNPs) can be performed by sequencing such as whole genome sequencing or exome sequencing. Sequencing can be accomplished through classic Sanger sequencing methods which are well known in the art.

B0001-B0024 ('652 Patent) at B0017, 15:1-8. The Patents also explain that:

> Genotyping of donor and recipient can establish a single nucleotide polymorphism (SNP) profile for detecting donor DNA. Shotgun sequencing of cell-free DNA in plasma, with analysis of observed unique SNPs, allows quantitation of % donor DNA.

Id. at B0018, 17:32-36; *see also id.* at B0019, 20:31-36 ("Genotyping of the transplant donor and/or the transplant recipient may be performed by any suitable method known in the art including those described herein such as sequencing, nucleic acid array or PCR. In some embodiments, genotyping of the transplant donor and/or the transplant recipient is performed by shotgun sequencing.").

94. I have significant personal experience using these standard and routine multiplex / high-throughput sequencing techniques for genotyping prior to November 6, 2009. For example, when I began working in the biological sciences at

the Salk Institute in 1992, the first molecular biology techniques that I learned were PCR and Sanger DNA sequencing. As part of my move to Stanford University in 1994, I was tasked with developing a new large-scale DNA sequencing method based on PCR mapping of transposon insertions. My recruitment to the Dana-Farber Cancer Institute in 2005 was largely based on the expectation that high throughput DNA sequencing would soon generate unprecedented quantities of data that could be used to understand cancer. And that expectation very quickly was realized throughout the genomics field.

95. Due to the rapid explosion in use of sequencing-by-synthesis instruments including the 454 sequencer in 2005, the Illumina Genome Analyzer and Applied Biosystems SOLiD sequencer in 2006, and others shortly thereafter, my group and I quickly became involved in multiplex / high-throughput (or Next Generation Sequencing ("NGS")) DNA sequencing. By 2009, my group and I were routinely analyzing DNA sequence data generated on nearly every multiplex or high-throughput (*e.g.,* NGS) platform available at the time.

96. In addition, in 2008, while a scientific advisor at Helicos Biosciences, I wrote a scientific paper with scientists at Helicos exploring applications of high throughput DNA sequencing entitled "What would you do if you could sequence everything?" that reported on a wide variety of applications, including genotyping.

B0251-B0259 (Kahvejian 2008). By 2009, I was routinely using multiplex / high-throughput sequencing for a wide variety of applications including genotyping.

(d) The Patents Claim No Improvements To Or Nonconventional Uses Of The Conventional Genotyping Methods Available In 2009

97. The claims recite no improvements to these common approaches to "genotyping," nor any novel way of carrying them out; rather the Patents explain that the "existing genotyping platforms known in the art" can be employed, handling up to hundreds of thousands of heterozygous and homozygous "[u]sable SNPs," and note that "[c]ompanies … currently offer both standard and custom-designed TaqMan probe sets for SNP genotyping that can in principle target any desired SNP position for a PCR-based assay." B0001-B0024 ('652 Patent) at B0016, 13:58-64. Consistent with the disclosure of the Patents, and my own experience as a researcher in the genomics field in 2009, a POSA understood "genotyping," including using "selective amplification … by … PCR," and establishing "polymorphism" or "SNP" "profiles" including homozygous or heterozygous variants to be a routine and conventional practice.

98. A declaration submitted to the USPTO during prosecution of the '652 Patent further confirms that the recited genotyping limitations were routine and conventional as of 2009. In order to overcome a rejection by the patent examiner, a scientist who worked in named inventor Stephen Quake's Stanford laboratory

submitted a declaration, dated January 30, 2014, to "demonstrate that genotype information obtained from a transplant recipient can be used to establish a polymorphism profile to detect donor-derived cell-free nucleic acids in a sample from the transplant recipient." B0260-B0263 (Beausang Dec.). In his declaration, the scientist, Mr. John Beausang, describes an analysis he had performed using the claimed methods. *Id.* He states that "[g]enotypes of the transplant recipients were obtained from genomic DNA using the Illumina Omni1-Quad Beadchip following standard protocols." *Id.* I am familiar with the Illumina BeadChip, which as discussed above in paragraph 89,was available by 2002 and was routinely used by researchers for genotyping to obtain polymorphism, including SNP, profiles following standard protocols by 2009.

3. Performing Multiplex Or High-Throughput Sequencing Of (Or Digital PCR On) The cfDNA To Detect The Genotyped Polymorphisms (Or SNPs)

99. Claim element 1(c) of the '652 Patent recites "multiplex sequencing of the cell-free nucleic acids in the sample followed by analysis of the sequencing results using the polymorphism profile." Claim element 1(d) of the '497 Patent recites "determining an amount of donor-specific cell-free nucleic acids by … high-throughput sequencing or [dPCR]." Claim element 1(d) of the '607 Patent recites "performing a high throughput sequencing reaction … compris[ing] … sequencing-by-synthesis." These also were standard techniques for analyzing nucleic acids by

the time of the November 6, 2009 filing date, and they were routinely used by that time to analyze cell-free nucleic acids in samples.

(a) Background On Multiplex / High Throughput Sequencing Technology

100. As explained above, methods for determining the exact sequences of the bases in a stretch of DNA – called sequencing – were well known as of November 6, 2009. *See, e.g.*, B0001-B0024 ('652 Patent) at B0017, 15:6-8. Methods for simultaneously sequencing multiple stretches of DNA in a high-throughput instrument – called multiplex or high-throughput sequencing, also known as Next Generation Sequencing or "NGS" – also were well known and routinely used as of November 6, 2009. *See, e.g.*, *id.* at B0017, 15:8-16:41. The terms "multiplex sequencing" as recited in the '652 Patent, and "high-throughput sequencing" as recited in the '497 and '607 Patents, were used interchangeably by POSAs in the field of nucleic acid analysis in 2009 when the Patents were filed. These terms were understood by POSAs (both in 2009 and now) to include parallel sequencing methods using automated instrumentation.

101. "Sequencing-by-synthesis," as recited in the '607 Patent claims, is a sequencing technique that by 2009 was one of the multiplex or high-throughput sequencing techniques most routinely used in the field of nucleic acid analysis. It involves determining the sequence of a particular nucleic acid by synthesizing a strand complementary to it inside of a sequencing machine. In the sequencing-by-

synthesis reaction, the bases being added are "tagged" in an identifiable manner, such that the sequencing machine can tell what base is being added at each position, and therefore identify the complete sequence of the nucleic acid that is being sequenced. Because nucleic acids, particularly DNA, form double-stranded structures through complementary binding and specific base-pairing, if a POSA knows the sequence of the strand that is newly sequenced, she can deduce the sequence of the complementary strand being interrogated.

102. The specification also describes a "shotgun" sequencing approach for sequencing circulating nucleic acids. *Id*. at B0014, 9:8-14. ("Detection, identification and/or quantitation of the donor-specific markers (e.g. polymorphic markers such as SNPs) can be performed using real-time PCR, chips (e.g., SNP chips), high-throughput shotgun sequencing of circulating nucleic acids (e.g. cell-free DNA), as well as other methods known in the art including the methods described herein."). Shotgun sequencing is a general strategy that was developed in the 1980s. B0264-B0274 (Hutchison) at B0268; B0275-B0285 (Green 2001) at B0276, B0281, B0282. In shotgun sequencing, multiple overlapping "reads" for the target DNA are obtained by performing several rounds of sequencing of DNA fragments. After all the fragments are sequenced, computer programs use the overlapping ends of different reads to assemble them into a continuous sequence.

(b) Commercially Available Multiplex And High-Throughput Sequencing Platforms As Of 2009

103. Multiplex and high-throughput sequencing techniques, including sequencing-by-synthesis techniques, had been in use for more than a decade before the 2009 filing date of the Patents. For example, a method called pyrosequencing, which performs sequencing-by-synthesis, was developed and first published in 1993. B0286-B0290 (Nyren 1993); *see also* B0291-B0300 (Ronaghi 2001). It was the foundation for different companies. PYROSEQUENCING AB, out of Uppsala Sweden, launched its first commercial automated pyrosequencing instrument in 1999. B0694-B0903 (Marsh 2007) at B0713. Another is called 454 Life Sciences, Inc. (later acquired by Roche Diagnostics, Inc.). *See also* B0001-B0024 ('652 Patent) at B0017, 15:38-45 (disclosing using 454 sequencers as embodiments of the claimed methods); B0309-B0312 (Coombs) at B0311. 454 Life Sciences commercially launched its first sequencing instruments in 2005. B0301-B0308 (Heather 2016) at B0303; B0309-B0312 (Coombs) at B0311. Based on the machines made available by Pyrosequencing, by the year 2000, pyrosequencing had established itself as a standard and conventional means for multiplex or high-throughput sequencing, including for genotyping SNPs. *See, e.g.*, B0313-B0320 (Ahmadian 2000 at B0313, B0318); B0074-B0079 (Nordstrom 2000) at B0074, B0078; B0291-B0300 (Ronaghi 2001) at B0292-B0293.

104. Another high throughput sequencing-by-synthesis method developed by a group out of Cambridge University in 1998 became the foundation for a company called Solexa (later acquired by Illumina, Inc.). Solexa commercially launched its first sequencing instrument, the "Genome Analyzer," in 2006. B0321-B0323 (Illumina, History of Sequencing by Synthesis, available at https://www.illumina.com/science/technology/next-generation-sequencing/illumina-sequencing-history.html); *see also* B0001-B0024('652 Patent) at B0017, 15:53-60; B0017-B0018,16:57-17:13 (disclosing using Solexa instruments including the Genome Analyzer as embodiments of the claimed method). As the group that developed this technology reported in Margulies 2005, the instrument branded as the Genome Analyzer could "sequence 25 million bases, at 99% or better accuracy, in one four-hour run," B11305-1309 (Margulies 2005) at B1305 (Abstract), and had an average read length of 108 bases, *id.* at B1307, Table 1.

105. Illumina acquired Solexa in 2007. Illumina provides a comprehensive line of products for large scale analysis of genetic variation and biological function, including genotyping and sequencing. In 2007, in a 10K SEC filing, Illumina reported that its "[i]nstrument revenue increased by $77.6 million over prior year, of which $68.7 million was due to increased sales of our sequencing systems, particularly the Genome Analyzer and cluster stations." By the end of 2008, Illumina

reported revenue of $573 million for its products. B0904-B0994 (Illumina 10K) at B0927. Illumina reported that its "Instrument revenue increased by $64.8 million over prior year, of which $63.0 million was due to increased sales of our sequencing systems. This increase in revenue can be primarily attributed to shipments of our second generation Genome Analyzer, the Genome Analyzer II (GAII)." *Id.* at B0935. Illumina's Genome Analyzer, which was in widespread routine use by 2009, is cited in the written description as an example product capable of performing the claimed laboratory techniques. B0001-B0024 ('652 Patent) at B0017, 16:50-59; B0018, 17:1-12; B0022, 26:47-52; *see also* B0260-B0263 (Beausang Decl.) at B0260-B0261.

106. As another example, a high throughput sequencing method co-developed by the named inventor Dr. Quake became the foundation for his company, Helicos, which was founded in 2003, and on which I was a member of the Scientific Advisory Board. Helicos commercially launched its first sequencing instrument, the HeliScope, in 2008. B0581-B0693 (Helicos 10K 2008); *see also* B0001-B0024 ('652 Patent) at B0017, 15:22-37; B0018, 17:22-28 (disclosing using Helicos instruments as embodiments of the claimed method).

107. The written description also describes the Pico Titer Plate device manufactured by 454 Lifesciences, Inc. as a commercial product for performing the recited sequencing element. *Id*. at B0017, 15:38-46. The Pico Titer Plate device,

available as of 2006, combines conventional PCR amplification with pyrosequencing in a high throughput, automated format. The Pico Titer Plate was one of the first next generation multiplex sequencing instruments developed, and was in common usage as of 2009. B0226-B0243 (Voelkerding 2009) at B0227.

108. The acceleration of nucleic acid sequencing technology before 2009 was incredibly rapid, with authors of a pertinent review article describing:

> The capabilities of DNA sequencers have grown at a rate even faster than that seen in the computing revolution described by Moore's law: the complexity of microchips (measured by number of transistors per unit cost) doubles approximately every two years, while sequencing capabilities between 2004 and 2010 doubled every five months. The various offshoot technologies are diverse in their chemistries, capabilities and specifications, providing researchers with a diverse toolbox with which to design experiments.

B0301-B0308 (Heather 2016) at B0305; *see also* B0324-B0333 (Morozova (2008) at B0324 ("A new generation of sequencing technologies, from Illumina/Solexa, ABI/SOLiD, 454/Roche, and Helicos, has provided unprecedented opportunities for high-throughput functional genomic research. To date, these technologies have been applied in a variety of contexts, including whole-genome sequencing, targeted resequencing, discovery of transcription factor binding sites, and noncoding RNA expression profiling."); *id.* at B0331 (noting that the commercially available high throughput sequencers "have provided genome-scale sequencing capacity to individual laboratories").

109. In an article published in March 2009, named inventor Dr. Quake reported a study on sequencing using the 454 FLX and Solexa DNA platforms, both of which are described as commercially available in the Patents. B0334-B0345 (White 2009) at B0334; *see also* B0001-B0024 ('652 Patent) at B0017, 15:38-60. In that publication, Dr. Quake states that high-throughput sequencing technologies such as 454 and Solexa that are "based on sequencing by synthesis and sequencing by ligation are revolutionizing biology, biotechnology, and medicine," and are a "key advance facilitating higher throughput and lower costs for several of these platforms was migration from the clone-based sample preparation used in Sanger sequencing to the massively parallel clonal PCR amplification of sample molecules on beads (Roche 454 and ABI Solid) or on a surface (Solexa)." B0334-B0345 (White 2009) at B0334.

110. By November 2009, POSAs also routinely used multiplex or high-throughput sequencing to detect SNPs in samples. This is confirmed, for example, in a scientific paper by Shen *et al.*, B0346-B0354 (Shen 2010), published online on December 17, 2009, titled "A SNP Discovery Method To Assess Variant Allele Probability From Next-Generation Resequencing Data." Shen *et al*. explain that "[i]n recent years, next-generation sequencing (NGS) technologies have propelled the rapid progress of genomics studies []. Continuous improvement in NGS technologies are increasing the throughput while lowering costs, thus enabling ultra-

large-scale sequencing efforts." *Id*. at B0346. The authors explain that [c]urrently, there are several methods available for detecting SNPs from NGS data," listing several examples, and further explain that they have provided a "freely available software package, Atlas-SNP2" to reduce error rates in interpreting sequencing data. *Id*. at B0346-B0347.

111. Based on my review of the literature and my own experience as a researcher active in cfDNA analysis in 2009, there were a number of commercially available sequencers from which to choose to carry out what the Patent claims describe at a high level as "multiplex" or "high-throughput" "sequencing," including what the Patents describe as "sequencing-by-synthesis."

112. The Patents' disclosure acknowledges these numerous commercial providers of multiplex or high-throughput sequencing instruments as of the time the Patents were filed. For example, the written description states that:

> Detection, identification and/or quantitation of the donor-specific markers (e.g. polymorphic markers such as SNPs) can be performed using real-time PCR, chips (e.g., SNP chips), ***high-throughput shotgun sequencing*** of circulating nucleic acids (e.g. cell-free DNA), ***as well as other methods known in the art including the methods described herein***.

B0001-B0024 ('652 Patent) at B0014, 9:8-14 (emphasis added). The Patents disclose, under the heading "Sequencing," *id.* at B0017, 15:1, numerous examples of commercial products that can carry out the multiplex or high-throughput

sequencing (including sequencing by synthesis), including those discussed above, such as:

- "technology available by Helicos Biosciences Corporation [] such as the Single Molecule Sequencing by Synthesis (SMSS) method," *id.* at B0017, 15:22-25;
- "technology available by 454 Lifesciences," *id.* at B0017, 15:39;
- "fiber optics detection [as] described in Marguiles, M., et al.," *id.* at B0017, 15:46-52;
- "Clonal Single Molecule Array (Solexa, Inc.) or sequencing-by-synthesis (SBS) utilizing reversible terminator chemistry," *id.* at B0017, 15:53-56;
- "AnyDot-.chips (Gonovoxx, Germany)," *id.* at B0017, 15:61-65; and
- "sequencers such as the Illumina Genome Analyzer," *id.* at B0017, 16:57-59.

113. Furthermore, a declaration submitted by Mr. Beausang to the USPTO during prosecution of the '652 Patent, in order to overcome rejections by the patent examiner, further confirms that the multiplex and high-throughput sequencing limitations were routine and conventional as of 2009. In his declaration, Mr. Beausang states that in carrying out the method recited in the '652 Patent claims, "[s]equencing libraries were constructed from cell-free DNA using commercially available kits and sequenced following standard protocols using an Illumina GAII sequencer." B0260-B0263 (Beausang Decl.) at B0260-B0261. I am familiar with the Illumina GAII sequencer, which was routinely used in 2009 by researchers, including myself, using commercially available kits and standard protocols.

(c) Sensitivities And Error Rates Inherent In The Available Sequencing Instruments

114. The Patent claims further recite certain levels of sensitivity in the claimed methods, for example at '652 Patent claim element 1(d), which recites "wherein sensitivity of the method is greater than 56% compared to sensitivity of current surveillance methods for cardiac allograft vasculopathy (CAV)"[13]; and at '497 Patent claim element 1(d) and '607 Patent claim element 1(f), which recite detection when the donor-specific circulating cell-free nucleic acids make up "at least 0.03% of the total circulating cell-free [nucleic acids/[DNA]]" in the sample. These sensitivity levels are not improvements over the methods already available, nor do the Patents disclose or claim any nonconventional use of the available methods in order to achieve a particular sensitivity. Rather, the recited sensitivites

[13] I do not believe a POSA would understand the meaning of the phrase "wherein sensitivity of the method is greater than 56% compared to sensitivity of current surveillance methods for cardiac allograft vasculopathy (CAV)" of '652 Patent claim element 1(d), as it makes no sense in the context of the claims. For example, "sensitivity" can have different meanings and there are different ways to measure it depending on what it refers to. For example, there is a level of "sensitivity" inherent in the sequencing methods recited in the claims that can be measured in one of several different ways based on the outcome of the sequencing instrument. There is also a clinical "sensitivity" inherent in the method based on how accurate the method is in identifying those patients who have organ transplant rejection or failure. The claim does not convey to a POSA what "sensitivity" it refers to or how it is measured. Notwithstanding, regardless of what the "sensitivity" in the claim refers to or how it is measured, it is an inherent feature of the claimed methods and neither the claims nor the written description refer to any improved or nonconventional way of achieving a specific level of "sensitivity' beyond what inherently results from the use of the conventional methods recited.

are inherent features of standard use of the methods, which themselves are routine and conventional.

115. For example, as to '652 Patent claim element 1(d), the Patents' written description explains that "[t]he invention provides methods that are sensitive and specific," and "[i]n some embodiments, the methods described herein have at least 56% sensitivity." B0001-B0024 ('652 Patent) at B0021, 23:31-36. A POSA would understand this is an inherent feature of the recited method, based on the conventional techniques recited. The Patents' written description states that "[t]he practice of the present invention employs, unless otherwise indicated, conventional techniques … which are within the skill of the art. …" *Id.* at B0012, 5:36-49. The written description does not anywhere indicate that the methods are practiced in a nonconventional way to achieve greater than 56% sensitivity compared to sensitivity of then-current surveillance methods for CAV – or compared to any other metric for that matter.

116. Furthermore, to the extent the element "sensitivity of the method is greater than 56%" ('652 Patent claim element 1(d)) refers to the sensitivity of a clinical diagnosis, this would be an inherent outcome of practicing the routine, conventional method steps. In a clinical diagnosis, "sensitivity" refers to the ability of the test to correctly identify those with the condition – here transplant rejection or failure (*i.e.*, true positive rate). There is nothing inventive about reciting a sensitivity

greater than 56% compared to anything. As far as the claims are concerned, a sensitivity of greater than 56% is slightly better odds than flipping a coin in terms of diagnosing transplant rejection or failure. Indeed, even a test that identifies every transplant recipient as undergoing transplant rejection or failure would have a 100% sensitivity rate (because the group would include all patients undergoing rejection/failure)--but the test would be useless because it would not eliminate false positives (it would also include patients who were not actually undergoing rejection/failure). The written description contains no explanation of a nonconventional way of achieving even 56% sensitivity (which, in my opinion, would be ineffective at diagnosing transplant rejection or failure as the claims purport in any event).

117. The written description only describes at a high level "sensitive and non-invasive methods, devices, compositions and kits for monitoring organ transplant patients, and/or for diagnosing or predicting transplant status or outcome (e.g. transplant rejection)," and the lowest sensitivity described in the patent is 56%. B0001-B0024 ('652 Patent) at B0021, 23:32-36 ("In some embodiments, the methods described herein for diagnosing or predicting transplant status or outcome have at least 56%, 60%, 70%, 80%, 90%, 95% or 100% sensitivity. In some embodiments, the methods described herein have at least 56% sensitivity."). The written description does not indicate any nonconventional technique for arriving at

this (albeit clinically ineffective) sensitivity for diagnosis – to the contrary it lacks any explanation as to how this sensitivity is achieved. Accordingly, a POSA would know that a sensitivity of at least 56% would have to be (and is) an inherent result of using the claimed techniques.

118. As to sensitivity of the sequencing methods, including as recited in '497 Patent claim element 1(d) and '607 Patent claim element 1(f), wherein detection occurs when the donor-specific circulating cell-free nucleic acids make up "at least 0.03% of the total circulating cell-free [nucleic acids/[DNA]]" in the sample, the Patents' written description states that "[h]igher sensitivity can be achieved simply by sequencing more molecules, i.e., using more channels." B0001-B0024 ('652 Patent) at B0018, 17:12-13. The written description also explains that, for example, "[c]urrently, sequencers such as the Illumina Genome Analyzer have read lengths exceeding 36 base pairs," *id.* at B0017, 16:57-59, and that "[o]n the Genome Analyzer … If one wants to establish a lower limit of sensitivity for this method by requiring at least 100 donor molecules to be detected, then it should have a sensitivity capable of detecting donor molecules when the donor fraction is as low as 0.03%." *Id.* at B0018, 17:1-11. This is merely states, again, an inherent feature of the conventional sequencing instruments, in standard use, as recited in the claims.

119. Indeed, it was well-known in the art by November 6, 2009 that "[h]igher sensitivity can be achieved simply by sequencing more molecules," as the

written description states. *See* B0001-B0024 ('652 Patent) at B0018, 17:12-13. The written description acknowledges that "[t]here are two components to sensitivity: (i) the number of molecules analyzed (depth of sequencing) and (ii) the error rate of the sequencing process." *Id.* at B0017, 16:52-55. POSAs prior to November 6, 2009 were well aware that sequencing more molecules, or increasing sequencing depth, improves sensitivity. As Voelkerding *et al*. explained in a 2009 review article reviewing multiplex / high-throughput (NGS) sequencing technology: "Accuracy in NGS is achieved by sequencing a given region multiple times, enabled by the massively parallel process, with each sequence contributing to 'coverage' depth." B0226-B0243 (Voelkerding 2009) at B0236. In another example, Thomas *et al.* in 2006 noted that "[s]equence variants that represent a fraction of a complex sample can be vastly oversampled, thus enabling statistically meaningful quantification of low-abundance variants." B1301-1304.1 (Thomas 2006) at B1301. This is further demonstrated in Thomas 2006 at Supplemental Figure 2, which shows that the detection limit is determined by sequencing error rate and sequencing read depth (*i.e.*, how many molecules are sequenced), reflecting the fact that there was a well-known correlation between the number of molecules sequenced and sensitivity including limit of detection of donor fraction. *Id.* at B1304.1.

120. The '607 Patent, at claim element 1(d), also recites that "said sequencing-by-synthesis reaction has a sequencing error rate of less than 1.5%."

This, too, is a feature inherent in common methods of using the sequencing instruments disclosed in the written description. In the context of the Patents, the sequencing error rate is an estimate of the proportion of all bases (often expressed as a decimal) sequenced by a sequencer that are incorrect. These errors can be caused by a combination of problems, for example, during sample preparation, misidentification of single base additions by the sequencer, background noise, or some combination of these and other factors. The Patents' written description explains that "it is possible to systematically lower the sequencing error rate by resequencing the sample template multiple times, as has been demonstrated by Helicos BioSciences (Harris, T.D. et al., Science, 320, 106-109 (2008))." B0001-B0024 ('652 Patent) at B0018, 17:22-26. The written description further explains that, as to the numerous commercial sequencing platforms disclosed in the patents, "[t]ypical sequencing error rates for base substitutions vary between platforms, but are between 0.5-1.5%." *Id.* at B0018, 17:20-21. Thus, as the Patents' written description recognizes, error rates were inherent in the recited sequencing techniques, and methods for lowering sequencing error rates were already known and established in the field before the Patents were filed.

121. The literature as of 2009 confirms that the recited sensitivity levels were inherent in the conventional use of the sequencing instruments at the time. For example, in a paper co-authored by Patent inventor Stephen Quake, Harris *et al.*,

Science (2008) 320:106-109, B0355-B0359, the authors explain the science behind the inherent sequencing error rates in Helicos sequencing technology. The authors also describe inherent error rates significantly lower than 1.5%. *Id.* at B0357.

(d) Digital PCR Methods Were Routine And Commercially Available In 2009

122. '497 Patent claim element 1(d) recites, as an alternative to high-throughput sequencing, a technique called digital PCR, or "dPCR." This technique also was known and conventional as of 2009. For example, the written description cites to and incorporates by reference a 2006 publication disclosing its use, at B0001-B0024 ('652 Patent) at B0016, 14:58-67.

(e) The Patents Claim No Improvements To Or Nonconventional Uses Of The Conventional Multiplex/High-Throughput Sequencing Or dPCR Methods Available In 2009

123. The "multiplex sequencing" and "high-throughput sequencing" (and "dPCR") techniques recited in the claims of the Patents refer to a broad range of well-established and routine parallel sequencing technologies, including sequencing-by-synthesis, as of 2009. The Patents recite using them only at a high level of generality, and do not purport to improve on them or claim any nonconventional use of them.

4. Quantifying The Transplant cfDNA In The Sample, Using The Polymorphisms / SNPs In The Sequences

124. Claim element 1(d) of the ’652 Patent recites “determining a quantity of the donor cell-free nucleic acids based on the detection of [donor and recipient cell-free nucleic acids] by the multiplexed sequencing.” Claim element 1(d) of the ’497 Patent recites “determining an amount of donor-specific cell-free nucleic acids … by detecting a homozygous or a heterozygous SNP within the [donor cell-free nucleic acids].” Claim element 1(f) of the ’607 Patent recites “quantifying an amount of [transplant cfDNA] … using markers distinguishable between … recipient and … donor … wherein said markers … comprise [SNPs] selected from said at least 1,000 [SNPs].” These quantification steps are broadly recited with no further specificity. A POSA would have understood that they could employ any of a wide range of standard techniques for quantifying nucleic acids, which were already well known and established by the filing date of the Patents.

125. The Patents’ written description explains that “[d]etection, identification, and/or quantification of the donor-specific markers (e.g., polymorphic markers such as SNPs) can be performed using real-time PCR, chips (e.g., SNP chips), high-throughput shotgun sequencing of circulating nucleic acids (e.g. cell-free DNA), as well as other methods known in the art including the methods described herein.” B0001-B0024 (’652 Patent) at B0014, 9:8-14. This is also described elsewhere in the written description, for example *id*. at B0020, 21:5-9,

which states that "[t]he presence or absence of one or more nucleic acids from the transplant donor in the transplant recipient may be determined by any suitable method known in the art including those described herein such as sequencing, nucleic acid arrays or PCR."

126. Quantification of SNPs based on multiplex or high-throughput sequencing data was a well-known and routine practice by the time of the 2009 filing date of the Patents. I, myself, was researching cancer-related polymorphisms, including SNPs, present in cfDNA in cancer patients using high-throughput sequencing data at that time. For example, my position at Dana-Farber involved contributing to developing analytic methods for the processing, including quantification, of DNA sequencing data generated by multiplex or high-throughput sequencing. This included analysis, including quantification, of sequencing data from cancer cell lines, bulk tumor samples, circulating tumor cells, and cell-free nucleic acids. Using multiplex or high-throughput sequencing to generate quantitative sequencing data, including from cfDNA, was a routine practice by November 6, 2009, and I used standard techniques for generating such sequencing data in my own experience.

127. By November 2009, POSAs were routinely publishing data generated by multiplex or high-throughput sequencing to quantify SNPs, including for the purpose of detecting different genotypes in a sample. For example, in a scientific

paper published by Sampson and Zhao, B0360-B0388 (Statistical Applications in Genetics and Molecular Biology (2009) 8(1):1-27, titled "Identifying Individuals in a Complex Mixture of DNA with Unknown Ancestry"), the authors described use of high-throughput sequencing data, including as generated from "Illumina" instruments, to "determine whether a specific individual contributes DNA" to "a mixture of DNA samples from numerous individuals." B0360-B0388 (Sampson 2009) at B0360-61. The authors provide algorithms for determining "allele frequencies," which a POSA would understand describes determining quantities of SNPs, from this data. *Id.* at, *e.g.,* B0362. The authors further examine and compare SNP frequency data generated on high-throughput sequencing platforms, such as Illumina's, to such data generated on other types of platforms. *Id.* at B0368; *see also id.* at B0361 (explaining the platforms being tested are "e.g. Illumina, Affymetrix").

128. The claims of the Patents recite quantifying donor cell-free nucleic acids only at a high level, and do not specify any particular technique for doing so. The claims do not purport to provide any innovative or nonconventional method for quantifying polymorphisms or SNPs in cell-free nucleic acids based on sequencing data. Nor do the claims purport to claim any improvement over the existing means for doing so as of 2009.

C. The Claimed Combinations Of Steps For Detecting Multiple Genomes In A cfDNA Sample Were Conventional

1. The Patents' Written Description States That The Laboratory Techniques In The Claims As A Whole Were Conventional As Of 2009

129. As I note above, the Patents' written description states that "[t]he practice of the present invention employs, ***unless otherwise indicated***, ***conventional techniques*** of immunology, biochemistry, chemistry, molecular biology, microbiology, cell biology, genomics and recombinant DNA, ***which are within the skill of the art***." B0001-B0024 ('652 Patent) at B0012, 5:36-40 (emphasis added). I did not find any indication in the Patents that the claimed combination of laboratory techniques was nonconventional, or carried out in a way that modifies or improves upon the traditional combination of these laboratory techniques.

130. To the contrary, the Patents' written description states that the combination of elements in the claims was routine and carried out using commercial products without modification at the time of the Patents' filing date. For example, the Patents' written description states that "***[g]enotyping***[14] donor and recipient nucleic acids, and/or detection, identification and/or ***quantification of the donor-specific nucleic acids***[15] after transplantation (e.g. ***polymorphic markers such as***

[14] '652 Patent claim element 1(b); '497 Patent claim elements 1(a) and (b); '607 Patent claim element 1(c).
[15] '652 Patent claim element 1(d); '497 Patent claim element 1(d) ; '607 Patent claim element 1(f).

SNPs)[16] ***can be performed by sequencing***[17] such as whole genome sequencing or exome sequencing." B0001-B0024 ('652 Patent) at B0017, 15:2-6 (emphasis added). The written description, in the same section, identifies several commercial multiplex and high-throughput sequencing[18] platforms routinely used to carry out this combination of steps by the 2009 filing date. *Id.* at B0017-B0018, 15:22-17:28 (further disclosing the inherent sensitivities and error rates inherent in the standard use of the sequencing equipment); see also section VI.B above.

131. The written description identifies obtaining or providing a sample of cell-free nucleic acids[19] as routine and conventional, and does not identify any nonconventional way of obtaining a sample that is used with the claimed combination. *See* section VII.B. above.

132. As to the '607 Patent, the written description does not indicate anything nonconventional about "selective amplification of … at least 1,000 [SNPs] … by PCR," as recited in claim element 1(c), or its combination with high throughput sequencing as recited in elements 1(d) and (e) to quantify the cfDNA as recited in

---

[16] '652 Patent claim element 1(b); '497 Patent claim elements 1(a) and (b) ; '607 Patent claim element 1(f).
[17] '652 Patent claim element 1(c); '497 Patent claim elements 1(d) ; '607 Patent claim elements 1(d) and (e).
[18] '652 Patent claim element 1(c); '497 Patent claim elements 1(d) ; '607 Patent claim element 1(d).
[19] '652 Patent claim element 1(a); '497 Patent claim elements 1(c) ; '607 Patent claim elements 1(a) and (b).

element 1(f). Rather, again, the written description explains that "***[g]enotyping*** donor and recipient nucleic acids, and/or detection, identification and/or ***quantitation*** of the donor-specific nucleic acids after transplantation (e.g. ***polymorphic markers such as SNPs***) ***can be performed by PCR***." B0001-B0024 ('652 Patent) at B0016, 14:29-32 (emphasis added). The written description describes commercially available PCR-based products that were routinely used to genotype and quantitate SNPs as of 2009, stating that with these products, "[u]sable SNPs may comprise approximately 500,000 heterozygous donor SNPs and approximately 160,000 homozygous donor SNPs," and that "[c]ompanies … currently offer both standard and custom-designed TaqMan probe sets for SNP genotyping that can in principle target any desired SNP position for a PCR-based assay." *Id.* at B0017, 13:55-64.

133. Moreover, the written description does not indicate anything nonconventional about selectively amplifying 1,000 SNPs (or even 10,000 SNPs as recited in dependent claim 5) and combining that with high-throughput sequencing of the selectively amplified DNA. Indeed, contemporaneous literature indicates that was conventional, and numerous commercial products were available to perform it. *See, e.g.* B0224-B0225 (RainDance 2008) (describing commercial product for amplifying thousands of SNPs and combining with subsequent high-throughput sequencing by synthesis); B0226-B0243 (Volkerding 2009) (review article

describing several different products for "targeted genomic resequencing," including to detect "polymorphisms," that involve up to hundreds of thousands of capture probes to "enrich[]" for the target polymorphisms, "with the enriched DNA amplified by PCR before NGS [*e.g.*, high-throughput sequencing] library preparation").

134. In fact, the written description does not mention sequencing of selectively amplified cfDNA at all – there is no reference to it in the specification. The written description does not indicate that the claimed method performs this combination by modifying the recited laboratory techniques or using them in a nonconventional way, nor does it state this is an improvement over the methods the written description otherwise describes as conventional. *See* B0001-B0024 ('652 Patent) at B0012, 5:36-40 ("The practice of the present invention employs, unless otherwise indicated, conventional techniques … which are within the skill of the art.").

135. Thus, each of the combinations of the claims of the Patents as a whole is described in the written description as conventional and routinely performed, including by commercial instruments, as of the Patents' filing date.

2. POSAs Have Long Recognized That The Same Methods For Detecting Fetal, Cancer, And Infectious cfDNA Could Be Used To Detect Transplant cfDNA

136. By 2009, the combinations of laboratory techniques claimed in the Patents were well understood, routine and conventional methods for quantifying cfDNA having different genotypes in a biological sample. Other scientists including myself had established, and were routinely using, the claimed combination of laboratory techniques for quantifying fetal cfDNA in samples from a pregnant woman, cancer cfDNA in samples from a cancer patient, and bacterial or viral cfDNA in samples from an infected patient or an infected animal, among other applications, as I discuss in more detail below. POSAs prior to 2009 stated that these same methods could be used to detect and quantify cfDNA from any of these sources, and acknowledged that they also applied to detecting transplant donor-derived cfDNA in an organ transplant recipient's sample.

137. Indeed, the Patents' written description recognizes that common approaches were being applied broadly to study transplant rejection/failure, pregnancy, cancer, and other conditions in which cell-free nucleic acids representing different genotypes are present in a patient's circulation. *See, e.g.*, B0001-B0024 ('652 Patent) at B0012, 6:67-7:5; *id.* at B0013, 7:16-19 (explaining through examples that "results collectively establish both circulating DNA, either free in plasma or from circulating cells, as a useful species in cancer detection and

treatment"); *id.* at B0013, 7:19-23 (explaining with examples that "[c]irculating DNA has also been useful in healthy patients for fetal diagnostics, with fetal DNA circulating in maternal blood serving as a marker for gender, rhesus D status, fetal aneuploidy, and sex-linked disorders"); *id.* at B0013, 8:1-21 (explaining with examples that "results establish that for heart transplant patients, donor-derived DNA present in plasma can serve as a potential biomarker for the onset of organ failure"). As the written description explains, and a POSA in the relevant November 2009 period appreciated, the approaches to detecting and quantifying foreign or abnormal cell-free nucleic acids could be broadly applied to many natural phenomena:

> In all these applications of circulating nucleic acids, the presence of sequences differing from a patient's normal genotype has been used to detect disease. In cancer, mutations of genes are a tell-tale sign of the advance of the disease; in fetal diagnostics, the detection of sequences specific to the fetus compared to maternal DNA allows for analysis of the health of the fetus. … [A]s cell-free DNA or RNA often arises from apoptotic cells, the relative amount of donor-specific sequences in circulating nucleic acids should provide a predictive measure of on-coming organ failure in transplant patients…

B0001-B0024 ('652 Patent) at B0013, 7:30-46.

138. Consistent with the Patents' disclosure, the application of methods for detecting fetal and cancer cell-free nucleic acids to detecting analogous transplant-related natural phenomena was well-accepted among POSAs. The named inventors of the Patents have acknowledged this not only in the Patents, but also in other

publications. For example, in 2011 Stanford published a report in one of its newsletters quoting two of the named inventors – Steven Quake and Hannah Valantine. They explained that the application of known methods for detecting fetal cfDNA to the analogous phenomenon of transplant cfDNA was the basis for their idea underlying the Patents. As the Stanford newsletter reported:

> The current study began when Valantine noticed research by Quake in 2008 showing that it is possible to detect fetal chromosomal abnormalities by sequencing cell-free DNA fragments in a maternal blood sample.
>
> "When I saw that, I thought, wow, this technique could probably be used to monitor heart rejection," said Valantine, ***noting that cells damaged during rejection also release DNA into the circulatory system***.
>
> "Hannah sought me out and I realized that an organ transplant can also be thought of as a genome transplant," said Quake. "***Someone else's genome is in your body. So by looking at variations in the DNA sequence, we can identify which DNA segments come from the new heart, and which come from you."***

B0389-B0391 (https://med.stanford.edu/news/all-news/2011/03/to-better-detect-heart-transplant-rejections-scientists-test-for-traces-of-donors-genome.html) (emphasis added).

139. Similarly, a 2014 article in a Stanford Engineering publication described the Patent inventors' adaptation of a method for fetal testing into the foundation for the Patents' methods as follows:

> The cell-free DNA technique hinges on the existence in the genome of naturally occurring regions of variation called single nucleotide polymorphisms, or SNPs. In 2008, Hannah Valantine, then a Stanford professor of cardiology, ***realized that a DNA-sequencing technique developed in Quake's lab to pick out small quantities of fetal DNA from a pregnant woman's blood might also be useful to track the fate of a transplanted organ.***

B0392-B0394 (Stanford Engineering 2014) at B0393 (emphasis added).

140. POSAs up to more than a decade before the 2009 filing date of the Patents recognized the ease of applying routine methods for detecting foreign genomes to the detection of transplant cfDNA. For example, in 1998 Lo *et al.* reported in a publication that: "We have shown that DNA from fetuses is present in the plasma of their mothers, and now suggest that, ***in transplant patients, DNA from the organ donor may also be present in the plasma of the recipient***." B0080-B0081 (Lo 1998) at B0080 (emphasis added). The Lo group further recognized that "[s]ince graft rejection is an important cause of cell death in the transplanted organ, our observations raise the possibility that the concentration of donor DNA in the recipient's plasma may be a marker for rejection." B0080-B0081 (Lo 1998) at B0080.

141. Thereafter, and following the Human Genome Project's complete sequencing of the human genome in April 2003, methods were quickly developed, standardized, commercialized, and applied by POSAs to detect different genotypes in a single patient sample. *See, e.g.*, B0094-B0155 (2001 Human Genome Project)

at B0098, B0101; B0156-B0183# (Human Genome Project NIH FAQ) at B0158, B0162; B0164-B0183 (1000 Genome Project) at B0164, B168. For example, in 2004, Ding *et al.* published a scientific paper describing laboratory techniques for detecting circulating cell-free nucleic acids that could be broadly applied to conditions involving different genotypes in a sample:

> These advances will help in catalyzing the clinical applications of fetal nucleic acids in maternal plasma. This analytical approach also will have ***implications for many other applications of circulating nucleic acids in areas such as oncology and transplantation***.

B0395-B0400 (Ding 2004) at B0395t (emphasis added). Ding 2004 also explained that, beyond detecting fetal cfDNA, the same laboratory techniques "could be extended to other areas of circulating nucleic acid analysis, including circulating tumor-specific DNA, such as Epstein–Barr virus DNA in nasopharyngeal carcinoma patients, KRAS point mutations, and donor-specific DNA in transplant recipients." *Id.* at B0399.

142. Likewise, in a PCT patent application filed on July 9, 2004 and published March 17, 2005, Cantor *et al*. described a method for detecting different genotypes arising from a range of natural phenomena in a sample:

> The method allows accurate detection of nucleic acids that are present in very small amounts in a biological sample. For example, the method of the present invention is preferably used to detect fetal nucleic acid in maternal blood sample; circulating tumor-specific nucleic acids in a blood, urine or stool sample; and ***donor-specific acids in transplant recipients***. In another embodiment, one can detect viral, bacterial, fungal, or other foreign nucleic acids in biological sample.

B0995-B1049 (WO 2005/023091) ("Cantor 2005"), at B0995, Abstract (emphasis added).

143. Similarly, in a PCT patent application filed March 26, 2007 and published March 26, 2008, B1050-B1104 (WO2008/118988 A1) ("Ehrich 2008"), Ehrich *et al*. described a laboratory technique for detecting and analyzing a range of different types of circulating cell-free nucleic acids, noting that "[i]n addition to prenatal applications, the methods find utility in a range of applications, including, but not limited to, detecting rare cancer mutations, ***detecting transplant rejection*** and forensics." B1050-B1104 (WO2008/118988 A1) ("Ehrich 2008") at B1067, 17:20-22(emphasis added).

144. In another example, Lo *et al.* filed a PCT patent application on July 23, 2008, which published on January 29, 2009, B1105-B1219 (PCT Application No. WO2009/013492) ("Lo PCT 2009"). At paragraph [0003], the Lo group described a method for "determining a nucleic acid sequence imbalance," which "generally relates to the diagnostic testing of genotypes and diseases by determining ***an imbalance between two different nucleic acid sequences***…" The Lo group noted that the method relates "particularly to the identification of … mutations and genotypes in a fetus via testing of a sample of maternal blood," and "also relates to the detection of cancer, ***the monitoring of transplantation***, and the monitoring of infectious diseases." *Id*. (emphasis added).

145. In all of the above examples from the literature, scientists describe methods for analyzing different nucleic acid sequences, or genotypes, in the same sample. As these examples demonstrate, by the November 6, 2009 filing date of the Patents, POSAs recognized that the same laboratory techniques for detecting different genotypes in a cfDNA sample would apply to a range of natural phenomena, including fetal conditions, cancer, organ transplant monitoring, and infectious disease.

3. POSAs Had Already Established The Claimed Laboratory Techniques By 2009 To Detect Multiple Genotypes In A Sample, Including In The Transplant Setting

146. Confirming the Patents' disclosure of established methods known in the art, numerous prior and contemporaneous publications describe combining the claimed laboratory techniques to detect different genotypes in different contexts, including transplant, fetal, cancer, infectious disease, and others.

147. I provide other non-exhaustive examples here.

(a) Cantor et al. – 2007

148. Cantor *et al.* filed a U.S. Patent Application on February 28, 2006, which published as B0401-B0427 (U.S. 2007/0207466) ("Cantor 2007") on September 6, 2007. Cantor 2007 describes a method for "***accurate detection of nucleic acids that are present in very small amounts in a biological sample … preferably used to detect*** fetal nucleic acid in maternal blood sample; circulating

tumor-specific nucleic acids in a blood, urine or stool sample; and ***donor-specific acids in transplant recipients***." B0401-B0427 (Cantor 2007) at B0401, Abstract (emphasis added). Cantor 2007 explains the same method also "can detect viral, bacterial, fungal, or other foreign nucleic acids in biological sample." *Id*.

149. Exemplary disclosures of Cantor 2007 describing the combination of Patent claim elements are included in the table below:

Claim Element	**Exemplary Disclosure**
Obtaining a sample	Cantor described a method of detecting "foreign nucleic acids" "in a biological sample" such as a "blood sample," which a POSA would understand requires obtaining or providing a biological sample as recited in '652 Patent claim element 1(a) and '497 Patent claim element 1(c). *See* B0401-B0427 (Cantor 2007) at B0401, Abstract. Cantor also disclosed that certain embodiments are "a method for the detection of … mutations in maternal plasma…," which requires providing a plasma sample as recited in '607 Patent claim element 1(a). *Id.* at B0409, ¶ [0016].
Genotyping to obtain a SNP profile	In an exemplary embodiment in the fetal context, Cantor disclosed that the method comprised "selecting one or more [SNPs]," then "determining the fetal genotype from a sample DNA isolated from the plasma, serum, or whole blood of the pregnant mother … [by] performing SNP[-specific] … primer-extension assay in several replicates…" B0401-B0427 (Cantor 2007) at B0411 ¶ [0023]. A POSA would understand this disclosure of Cantor to refer to genotyping to generate a polymorphic, particularly SNP, profile as disclosed in '652 Patent claim element 1(b) and '497 Patent claim elements 1(a) and (b). A POSA would further understand Cantor's disclosure of a "primer-extension assay" to refer to a form of amplification, such as PCR, and disclosure of a SNP-specific primer-extension assay as a form of selective amplification of SNPs by PCR, as recited

	in ’607 Patent claim element 1(c). *See, e.g., id.* at B0409, ¶ [0016] (describing “using methods such as the primer-extension of polymerase chain reaction (PCR) products”).
Multiplex / high-throughput sequencing-by-synthesis	Cantor disclosed numerous commercially available means for detecting the SNPs, including “pyrosequencing techniques (Pyrosequencing, Inc., Westborough, Mass.).” *Id.* at B0420-B0424 ¶ [0054]; [0063]. A POSA would understand this to be a type of multiplex, or high-throughput sequencing instrument that uses sequencing-by-synthesis, as recited in ’652 Patent claim element 1(c), ’497 Patent claim element 1(d), and ’607 Patent claim elements 1(d) and (e). *See id.* at B0424, ¶ [0063] (“… PYROSEQUENCING™ (Uppsala, Sweden) [] essentially is sequencing by synthesis.”).
Quantitating the foreign cfDNA	Cantor disclosed “quantitating the fetal nucleic acids or alleles using the methods of the present invention” disclosed in its written description. *E.g.*, B0401-B0427 (Cantor 2007) at B0415; B0420, ¶ [0036]; [0053]. A POSA would understand the disclosure of quantitating fetal alleles in Cantor to refer to quantifying or determining an amount of foreign cfDNA in the sample as recited in ’652 Patent claim element 1(d), ’497 Patent claim element 1(d), and ’607 Patent claim element 1(f).

150. As this exemplary disclosure of Cantor 2007 demonstrates, a POSA understood that as of the 2006 filing date of Cantor 2007, laboratory techniques were commercially available for carrying out the method of obtaining a sample—genotyping—sequencing—and quantifying cell-free nucleic acids as claimed in the Patents to detect different genotypes, including in the organ transplant context. The Patents do not claim any unconventional use of this combination of laboratory

techniques described in Cantor 2007, nor do they claim any unique way of combining or using the techniques.

(b) Lo et al. – 2009

151. In another example, the Lo group described a method that "relates to the detection of cancer, ***the monitoring of transplantation***, and the monitoring of infectious diseases" in a U.S. Patent application filed on July 23, 2008 and published as U.S. 2009/0087847 ("Lo U.S. 2009"). B1220-B1290 (Lo U.S. 2009) at BB1252, ¶ [0003] (emphasis added).

152. Exemplary disclosures of Lo U.S. 2009 describing the combination of Patent claim elements are included in the table below:

Claim Element	Exemplary Disclosure
Obtaining a sample	Lo exemplified its disclosed method in an embodiment in which "plasma DNA will be extracted." B1220-B1290 (Lo U.S. 2009) at B1274, ¶ [0289]. That plasma DNA contains "maternal and fetal DNA in the plasma sample." *Id.* A POSA would understand that this disclosure in Lo requires providing or obtaining a biological sample, specifically a plasma sample from the blood, that contains different genotypes as recited in '652 Patent claim element 1(a), '497 Patent claim element 1(c), and '607 Patent claim elements 1(a) and (b), respectively.
Genotyping to obtain a SNP profile	Lo disclosed genotyping generally, at ¶ [0003], stating the disclosed method "generally relates to the diagnostic testing of genotypes and diseases by determining an imbalance between two different nucleic acid sequences…in a fetus…" and "also relates to the detection of cancer, the monitoring of transplantation, and the monitoring of infectious diseases." Lo described embodiments that evaluate a "clinically relevant

	nucleic acid sequence," *id.* at B1254, ¶ [0059], against a "background nucleic acid sequence." *Id.* at B1254, ¶ [0060]. Lo described clinically relevant nucleic acid sequences as (among other things) "sequences which are mutated, deleted, or amplified in a malignant tumor, e.g. sequences in which loss of heterozygosity or gene duplication occur." *Id.* at B1254, ¶ [0059]. It described background nucleic acid sequences as (among other things) an allele from the same chromosome as a clinically relevant sequence, but nonetheless "distinct due to heterozygosity." *Id.* at B1254, ¶ [0060]. Lo also disclosed that in an example of the method, the fractional percentage of fetal material is "determined by measuring the amount of a fetal-specific marker (e.g. … genetic polymorphism markers (e.g. SNPs) … in relation to a non-fetal-specific marker." *Id.* at B1258, ¶ [0100]. Lo disclosed "determining the fractional concentration of fetal DNA … through the quantification of polymorphic differences between the pregnant woman and the fetus," citing references disclosing established techniques, and providing "[a]n example of this method … to target polymorphic sites at which the pregnant woman is homozygous and the fetus is heterozygous." B1220-B1290 (Lo U.S. 2009) at B1272-B1273, ¶ [0278]. A person of skill in the art would understand these disclosures as describing the kind of genotyping used to establish polymorphism or SNP profiles for different genotypes in the sample as recited in '652 Patent claim element 1(b), and '497 Patent claim elements 1(a) and (b).
Multiplex / high-throughput sequencing and digital PCR	Lo also described using commercially available digital PCR and multiplex/high-throughput (a.k.a. "massively parallel") sequencing instruments that were routinely used to carry out the techniques recited in '652 Patent claim element 1(c), '497 Patent claim element 1(d), and '607 Patent claim element 1(d). These Lo disclosures include using "microfluidics digital PCR chips," citing references dating back to 2006, and "massively parallel genomic sequencing," citing references

	dating back to 2005. B1220-B1290 (Lo U.S. 2009) at B1263, ¶ [0176]. Lo U.S. 2009 disclosed several specific examples of the routine use of these commercial products, for example: > Massively parallel sequencing, such as that achievable on the 454 platform (Roche) (Margulies, M. et al. 2005 Nature 437, 376-380), Illumina Genome Analyzer (or Solexa platform) or SOLiD System (Applied Biosystems) or the Helicos True Single Molecule DNA sequencing technology (Harris TD et al. 2008 Science, 320, 106-109), the single molecule, real-time (SMRT™) technology of Pacific Biosciences, and nanopore sequencing (Soni GV and Meller A. 2007 Clin Chem 53: 1996-2001), allow the sequencing of many nucleic acid molecules isolated from a specimen at high orders of multiplexing in a parallel fashion (Dear BriefFunct Genomic Proteomic 2003; 1: 397-416). Each of these platforms sequences clonally expanded or even non-amplified single molecules of nucleic acid fragments. As a high number of sequencing reads, in the order of hundred thousands to millions or even possibly hundreds of millions or billions, are generated from each sample in each run, the resultant sequenced reads form a representative profile of the mix of nucleic acid species in the original specimen. B1220-B1290 (Lo U.S. 2009) at B1271, ¶ [0262].
Quantitating the foreign cfDNA	As recited in ’652 Patent claim element 1(d), ’497 Patent claim element 1(d), and ’607 Patent claim element 1(f), Lo described measuring “e.g., SNPs” to determine percentage of foreign cfDNA in the sample, whereby “[t]he actual measurement could be done by real-time PCR, digital PCR, sequencing reactions (including massively parallel genomic sequencing) or any other quantitative methods.” B1220-B1290 (Lo U.S. 2009) at B1258, ¶ [0100]. In the Lo

	exemplary fetal embodiment, "[t]he amount of maternal and fetal DNA in the plasma sample will be quantified, for example by the real-time PCR assays previously established … or other types of quantifier well known to those of skill in the art," citing references. B1220-B1290 (Lo U.S. 2009) at B1252, ¶ [0003].

153. A POSA would understand the combination of conventional and commercially available laboratory techniques disclosed in Lo U.S. 2009 for detecting and quantifying different genotypes in a sample to be the same combination of laboratory techniques claimed in the Patents. The Patents do not claim any nonconventional use of the combination of laboratory techniques described in Lo U.S. 2009, nor do they claim any unique way of combining or using the techniques. Though, like Cantor 2007 discussed above, Lo U.S. 2009 often applies the combination of techniques for detecting fetal cfDNA in maternal blood, Lo also recognized its applicability to the analogous natural phenomenon of transplant cfDNA in a transplant recipient's blood. B1220-B1290 (Lo U.S. 2009) at B1252, ¶ [0003]. Thus, like Cantor 2007, Lo U.S. 2009 confirms the conventionality and prior existence of the combination claimed in the Patents at least as of July 23, 2008 (the Lo U.S. 2009 filing date).

(c) Dhallan et al. – 2007

154. In another example, Dhallan *et al.* published a scientific paper in the journal Lancet at 369(9560):474-81 on February 10, 2007, B0428-B0435 ("Dhallan

2007"). Like the preceding Cantor 2007 and Lo U.S. 2009 disclosures, Dhallan 2007 establishes the conventionality of the Patents' claimed combination for detecting multiple different genomes in a sample. The Dhallan 2007 authors note that "[t]he approach described here, which uses standard molecular biology equipment, allows for precise analysis of genetic material." *Id.* at B0434.

155. Exemplary disclosures of Dhallan 2007 describing the combination of Patent claim elements are included in the table below:

Claim Element	**Exemplary Disclosure**
Obtaining a sample	As recited in '652 Patent claim element 1(a), '497 Patent claim element 1(c), and '607 Patent claim element 1(a), Dhallan 2007 disclosed obtaining "formaldehyde-treated blood samples from [] pregnant women," and using "free fetal DNA" from the samples "to diagnose fetal chromosomal abnormalities." B0428-B0435 (Dhallan 2007) at B0428. The Dhallan "Methods" section states that plasma samples were prepared and cfDNA was extracted (as further recited in '607 Patent claim element 1(b)) using standard, then-available commercial products: > Plasma and buffy coat samples were isolated in accordance with methods described previously. Genomic DNA was purified from both the plasma fraction and buffy coat of the same maternal blood sample, and the buffy coat fraction of the paternal sample, with the QIAamp DNA Blood Maxi Kit (Qiagen, Valencia, CA, USA). … Plasma DNA was concentrated to 50-70 uL with a 10 kDa nominal molecular weight cutoff filter (Millipore, Bedford, MA, USA).

	B0428-B0435 (Dhallan 2007) at B0429 (internal citations omitted).
Genotyping to obtain a SNP profile	Dhallan further described using the genotyping steps recited in '652 Patent claim element 1(b) and '497 Patent claim elements 1(a) and (b) (reciting genotyping to obtain polymorphism or SNP profiles, respectively), and '607 Patent claim element 1(c) (reciting selective amplification of at least 1,000 SNPs by PCR). Specifically, Dhallan disclosed that "SNPs were amplified from genomic DNA isolated from the maternal plasma … by PCR," *id.* at B0429-B0430, further noting that "the human genome project has identified over 3.7 million SNPs to date," making "[e]ven on the smallest human chromosome … about 54,000 genotyped SNPs [] available for analysis." *Id.* at B0434. Dhalla further disclosed genotyping homozygous and heterozygous SNPs, exemplifying that "[a]t certain SNP sites, the maternal genome will be homozygous for a nucleotide … while at the same site the paternal genome might be homozygous for a different nucleotide" such that "the fetal genome will be heterozygous [] at the SNP site." *Id.* at B0428.
Sequencing	Dhallan described performing sequencing using "sequencing gels." *Id.* at B0430. Though '652 Patent claim element 1(c), '497 Patent claim element 1(d), and '607 Patent claim element 1(d) recite performing sequencing by multiplex or high-throughput sequencing, one skilled in the art as of the 2009 filing date of the Patents would have understood that the same sequencing performed in Dhallan 2007 using sequencing gels could have been accomplished using any of the numerous commercial multiplex/high-throughput sequencers available at the time (*see* section VII.B.3. above). In my own experience, as confirmed by the literature, it was standard and routine as of 2009 to use the same type of sample preparation, cfDNA extraction, and genotyping/selective amplification techniques disclosed in Dhallan in combination with one of the many commercial multiplex/high-throughput sequencing instruments to generate sequencing results for quantification as further recited in the Patent claims. *See* section VII.B.2. above.

Quantitating the foreign cfDNA	As recited in '652 Patent claim element 1(d), '497 Patent claim element 1(d) and '607 Patent claim element 1(f), Dhallan disclosed quantifying foreign cfDNA in the sample, stating that "us[ing] standard molecular biology equipment … [f]ree fetal DNA is directly quantified from the heterogeneous mixture of maternal and fetal DNA in the maternal plasma." *Id.* at B0434.

156. Thus, a POSA would understand from the Dhallan 2007 disclosure, just like from the other literature cited above, that by 2007, "standard" laboratory techniques were available and used to accomplish the techniques claimed in the Patents.

(d) Beck et al. – 2009

157. In a further example, Beck *et al*. published a paper in the journal Clinical Chemistry at 55(4):730-738 in 2009 (B0436-B0444) ("Beck 2009"). Like the above-referenced literature, Beck 2009 demonstrates the conventionality of the Patents' claimed combination for detecting multiple genotypes in a sample. In the case of Beck 2009, the authors used the combination of techniques to detect both pathogens (*e.g.,* viruses) and mutated cfDNA in otherwise apparently healthy human subjects. Beck 2009 describes an overview of the methods used therein, stating: "Serum DNA from 51 apparently healthy humans was extracted, amplified, sequenced via pyrosequencing (454 Life Sciences/Roche Diagnostics), and

categorized by (a) origin (human vs xenogeneic),[20] (b) functionality (repeats, genes, coding or noncoding), and (c) chromosomal localization." *Id.* at B0436, Abstract Methods.

158. Exemplary disclosures of Beck 2009 describing the methods claimed in the Patents are included in the table below:

Claim Element	Exemplary Disclosure
Obtaining a sample	As recited in '652 Patent claim element 1(a), '497 Patent claim element 1(c), and '607 Patent claim elements 1(a) and (b), Beck 2009 explained that "[s]erum samples were collected and stored," and that a commercial kit was used to extract the cell-free nucleic acids, wherein "[t]otal nucleic acids were extracted … with the High Pure Viral Nucleic Acid Kit (Roche Applied Science) according to the manufacturer's instructions." B0436-B0444 (Beck 2009) at B0437. The authors further stated that "[w]e also collected EDTA-anticoagulated samples of whole blood from a subgroup of the volunteers (2 females, 2 males) and extracted genomic DNA with standard protocols." *Id.*
Genotyping to obtain a SNP profile	Beck 2009 reported genotyping as recited in '652 Patent claim element 1(b) and '497 Patent claim elements 1(a) and (b) (reciting genotyping to obtain polymorphism or SNP profiles, respectively). Beck 2009 reported that "we have applied this high-throughput sequencing technology to generate an unbiased profile of the circulating DNA in healthy individuals." B0436-B0444 (Beck 2009) at B0437. Beck 2009 described using "the BLAST program," which is a publicly available database, as well as "a local install of the RepeatMasker software package (Institute for Systems Biology), which makes use of Repbase (version 12.09; Genetic Information Research Institute" and "querying

[20] A POSA would understand the term "xenogenic" to mean nucleic acids that originate outside the subject itself, and are from a foreign substance introduced into the subject, such as a virus or other pathogen in the context of this study.

	databases of bacterial, viral, and fungal genomes, as well as the human genome (reference genome build 36.2)," including genomes "obtained from the National Center for Biotechnology Information (NCBI) (ftp://ftp.ncbi.nih.gov)." *Id*. These are all publicly available databases and software packages that include reference genomes for genotyping. The authors further compared the circulating cfDNA to genomic DNA in the subjects, looking for example, for polymorphic differences in single genes, repetitive elements, and foreign elements such as from non-human sources. *Id.* at B0438-B0439.
Multiplex / high-throughput sequencing	Beck 2009 described multiplex or high-throughput sequencing as described in '652 Patent claim element 1(c), '497 Patent claim element 1(d), and '607 Patent claim element 1(d), as well as sequencing by synthesis as recited in '607 Patent claim element 1(d). Specifically, Beck 2009 reported that "[t]he amplified DNA preparations were sequenced directly with a GS FLX genome sequencer (454 Life Sciences/Roche Diagnostics) according to the manufacturer's instructions." B0436-B0444 (Beck 2009) at B0437. As discussed in section VII.B.3. above, a POSA would understand the 454 sequencing instrument to employ sequencing-by-synthesis.
Quantitating the foreign cfDNA	As recited in '652 Patent claim element 1(d), '497 Patent claim element 1(d), and '607 Patent claim element 1(f), Beck 2009 disclosed quantifying foreign cfDNA in the sample, stating for example that "[t]o quantify the amounts of unidentified nucleotides, we counted and subtracted the masked nucleotides from the total nucleotide counts." *Id.* at B0437. The authors further described determining whether "representation of the different chromosomes in the serum [cfDNA] is correlated with chromosome gene density or GC content," which a POSA would understand to involve quantifying cfDNA to compare its "representation" or frequency in the cfDNA to that of genomic DNA. *Id*. at B0439. Beck 2009 also reported determining "representation of repetitive elements" in the cfDNA, which a POSA would

	understand involves determining its relative fraction in the sample.

159. Thus, a POSA would understand from the Beck 2009 disclosure, just like from the other literature cited above, that by 2009, the methods claimed in the Patents were routinely being used to detect multiple genotypes in a cfDNA sample.

(e) Beck et al. – 2010

160. In an additional example, Beck *et al.* published a scientific paper in the journal Molecular Cancer Research at 8(3):335-342, which was submitted to the journal on July 16, 2009 and published March 9, 2010 ("Beck 2010") that further establishes the conventionality of the Patents' claimed methods for detecting multiple genotypes in a sample, in this case in the context of cancer. B0445-B0453 (Beck 2010). I note that the submission date of July 16, 2009 indicates that the experiments underlying Beck 2010 were carried out before the Patents' earliest filing date of November 6, 2009. Beck 2010 reported that "[c]irculating nucleic acids (CNA) isolated from serum or plasma are increasingly recognized as biomarkers for cancers," and that related to the authors' work, "[i]dentification of specific breast cancer-related CNA sequences provides the basis for the development of a serum-based routine laboratory test for breast cancer screening and monitoring." B0445-B0453 (Beck 2010) at B0445, Abstract.

161. Exemplary disclosures of Beck 2010 describing the claimed methods are included in the table below:

Claim Element	Exemplary Disclosure
Obtaining a sample	As recited in '652 Patent claim element 1(a), '497 Patent claim element 1(c), and '607 Patent claim elements 1(a) and (b), Beck 2010 reported that serum samples were collected from breast cancer patients, and after storage, using standard, then-available commercial products, "serum was centrifuged at 4,000 x *g* for 20 min and 200 uL of the supernatant was used in the High Pure Nucleic Acids Extraction Kit (Roche) according to the instructions of the manufacturer." *Id.* at B0446.
Genotyping to obtain a SNP profile	Beck 2010 further reported the genotyping steps recited in '652 Patent claim element 1(b) and '497 Patent claim elements 1(a) and (b) (reciting genotyping to obtain polymorphism or SNP profiles, respectively). Specifically, Beck 2010 reported that "[c]ancer-specific DNA perturbations such as microsatellite instability, mutations, sequence length, and promoter methylation patterns detected in serum/plasma have been proposed for the diagnosis and clinical assessment of cancer treatment." *Id.* at B0446. Beck 2010 further reported analyzing repetitive elements based on "assigned nucleotides." *Id.*; *see also id.* at B0447. A POSA would understand that repetitive elements are a form of polymorphism that are prevalent in cancer.
Multiplex / high-throughput sequencing	Beck 2010 taught performing sequencing as recited in '652 Patent claim element 1(c), '497 Patent claim element 1(d), and '607 Patent claim element 1(d), which recite performing sequencing by multiplex or high-throughput sequencing, and by sequencing-by-synthesis, respectively. Specifically, Beck 2010 described sequencing "using a Roche/454 GS-FLX high-throughput sequencer," which as discussed in section VII.B.3. above, is a high-throughput sequencer that performs sequencing by synthesis.

Quantitating the foreign cfDNA	As recited in '652 Patent claim element 1(d), '497 Patent claim element 1(d), and '607 Patent claim element 1(f), Beck 2010 described quantitating the abnormal cfDNA, in that case cancer-specific cfDNA. Specifically, they stated their results as "[h]igh-throughput sequencing of total serum DNA shows differential representation of certain repetitive elements in the CAN of patients with breast cancer compared with healthy controls." *Id.* at B0449. A POSA would understand that determining the "representation" of the analyzed polymorphisms involves quantifying them. A POSA would further understand that the statistical methods the authors used in Beck 2010 were standard and well-accepted means for quantifying data as of 2009.

162. Thus, a POSA would understand from the Beck 2010 disclosure, just like from the other literature cited above, that by 2009, the methods claimed in the Patents were routinely being used to detect multiple genotypes in a cfDNA sample.

(f) Gordon et al. – 2009

163. In an additional example, Gordon *et al.* published a scientific paper in the journal Nucleic Acids Research at 37(2):550-556, which was published online December 5, 2008 and published in print in 2009. This paper also establishes the conventionality of the Patents' claimed methods for detecting multiple genotypes in a sample. B0454-B0460 ("Gordon 2009"). The authors explained that "[n]ext-generation sequencing provided broad elucidation of sample CNAs [circulating nucleic acids]," further explaining that "we detected infection-specific sequences…" in order to identify pathogens that cause chronic wasting disease in cattle and elk cfDNA. *Id.* at B0454, Abstract.

164. Exemplary disclosures of Gordon 2009 describing the claimed methods are included in the table below:

Claim Element	Exemplary Disclosure
Obtaining a sample	As recited in '652 Patent claim element 1(a), '497 Patent claim element 1(c), and '607 Patent claim element 1(a), the Gordon 2009 authors explained that "[s]erum samples were obtained" from cattle and elk, and cfDNA was extracted using a commercial kit: "The WGA4 GenomePlex® Single Cell Whole Genome Amplification Kit (Sigma) was used for the nucleic-acid extraction … according to the manufacturer's protocol." *Id.* at B0455.
Genotyping to obtain a SNP profile	Gordon 2009 also reported using the genotyping steps recited in '652 Patent claim element 1(b) and '497 Patent claim elements 1(a) and (b) (reciting genotyping to obtain polymorphism or SNP profiles, respectively), and '607 Patent claim element 1(c) (reciting selective amplification of at least 1,000 SNPs by PCR). Gordon 2009 explained that "[a] total of 657,431 elk sequences were elucidated, including 401,733 from animals post-infection" and that they were "searched against the public protein and EST datasets" and matched against "public databanks" to genotype the subjects. *Id.* at B0456. The authors also reported that "[a] total of 595,037 quality bovine sequences were elucidated, including 311,786 from animals post-infection," which also were genotyped against public databases. *Id.* Gordon 2009 further described using selective amplification to amplify "3261 control-only motifs and 2896 motifs present in infected-only cows" using PCR under conditions a POSA would understand to be standard and conventional. *Id.* at B0455.
Multiplex / high-throughput sequencing	Gordon 2009 described using multiplex or high-throughput sequencing, including sequencing-by-synthesis, as recited in '652 Patent claim element 1(c), '497 Patent claim element 1(d), and '607 Patent claim element 1(d), respectively. Specifically, Gordon 2009 reported using "a Roche/454 genome sequencer (GS-FLx) according to the manufacturer's

	instructions." *Id.* at B0455. A POSA would understand the 454 sequencer to be a high-throughput sequencer that uses sequencing by synthesis, as explained in section VII.B.3. above.
Quantitating the foreign cfDNA	In addition, as recited in '652 Patent claim element 1(d), '497 Patent claim element 1(d), and '607 Patent claim element 1(f), Gordon 2009 reported that the data generated in the study may be used for "targeted characterization and quantification of disease-specific DNAs." *Id.* at B0458. The authors further reported that "[t]he information from the elk study demonstrates that a time course analysis of the blood CNAs can greatly improve our ability to determine the diagnostic signals present" in such diseases. *Id.* at B0459. Accordingly, a POSA would understand that the combination reported in Gordon 2009 can readily include quantifying foreign cfDNA in the sample.

165. Thus, a POSA would understand from the Gordon 2009 disclosure, just like from the other literature cited above, that by 2009, the methods claimed in the Patents were routinely being used to detect multiple genotypes in a cfDNA sample.

(g) Holt et al. – 2009

166. Holt *et al.* published a scientific paper in the journal Bioinformatics, at 25(16):2074-2075, with advanced access publication on June 3, 2009. B0461-B0462 ("Holt 2009"). Holt 2009 also establishes the conventionality of the Patents' methods for detecting multiple genotypes in a sample. The authors explained that they "present a method for estimating the frequencies of SNP alleles present within pooled samples of DNA using high-throughput short-read sequencing." *Id.* at

B0461, Abstract. They further note that "[t]he method was implemented in Perl and relies on the open source software Maq for read mapping and SNP calling." *Id.*

167. Exemplary disclosures of Holt 2009 describing the claimed methods are included in the table below:

Claim Element	Exemplary Disclosure
Obtaining a sample	'652 Patent claim element 1(a), '497 Patent claim element 1(c), and '607 Patent claim element 1(a) recite obtaining a sample of cell-free nucleic acids, and '607 Patent claim element 1(b) further recites extracting cfDNA from the sample. Holt 2009 analyzed a pool of DNA from six strains of the bacteria *Salmonella Patapyphi* A. *Id.* at B0461, Abstract. Though Holt 2009 did not use cfDNA obtained from blood as a sample, a POSA would appreciate that the same further combination of elements could be equally applied to the DNA extracted from a pool of bacteria as to the DNA extracted from a blood sample containing cfDNA.
Genotyping to obtain a SNP profile	Holt 2009 described using the genotyping steps recited in '652 Patent claim element 1(b) and '497 Patent claim elements 1(a) and (b) (reciting genotyping to obtain polymorphism or SNP profiles, respectively). Specifically, Holt 2009 explained that "[t]his approach facilitates genome-wide SNP detection among closely related isolates." *Id.* at B0461. For this study, Holt 2009 analyzed "a set of 403 SNPs" to generate profiles of allele frequencies that differed as between the strains in the pooled sample. *Id.* at B0461-B0462. The authors stated that their method "relies on opensource software Maq for read mapping and SNP calling," and that it is "freely available from ftp://ftp.sanger.ac.uk/pub/pathogens/pools/." *Id.* at B0461, Abstract. A POSA would understand this software for genotyping was standard and openly available as of 2009.
Multiplex / high-	As recited in '652 Patent claim element 1(c), '497 Patent claim element 1(d), and '607 Patent claim element 1(d), which recite multiplex or high throughput sequencing, and sequencing by

throughput sequencing	synthesis, respectively, Holt 2009 disclosed "us[ing] the Illumina GAI to sequence pools of DNA." *Id.* at B0461. As discussed above in section VII.B.3., the Illumina Genome Analyzer (Illumina GAI) is based on Solexa high throughput sequencing technology that uses sequencing by synthesis.
Quantitating the foreign cfDNA	As recited in '652 Patent claim element 1(d), '497 Patent claim element 1(d), and '607 Patent claim element 1(f), Holt 2009 reported that the data generated in the study was used to quantify the SNPs being analyzed. Specifically, Holt 2009 described determining "[t]he frequency of each SNP [] in [the pools] … using information (read from Maq's pileup output) on each read [] of [] reads mapped to the SNP locus…." *Id.* at B0461. The authors described using standard technology available through the open software, wherein "[f]requencies were calculated according to the following formulae, implemented in a Perl script which calls Maq to do the initial read mapping and SNP calling…" *Id.*

168. Thus, a POSA would understand from the Holt 2009 disclosure, just like from the other literature cited above, that by 2009, conventional laboratory techniques were available for obtaining a sample, genotyping, sequencing, and quantifying different genotypes in the sample.

D. The Dependent Claims Also Recite Well-Understood, Routine and Conventional Steps

169. None of the dependent claims of the Patents add a nonconventional application or inventive concept to the claims. In my opinion, each dependent claim recites well-understood, routine and conventional steps to observe a natural phenomenon. All of the dependent claims of the Patents depend from, and

incorporate by reference the limitations of claim 1 of the respective patents. I address the dependent claims, grouped by related claim elements, in turn below.

1. Dependent Claims Reciting Different Types Or Numbers Of Polymorphisms

170. Regarding the ’652 Patent, dependent claim 2 recites different types of genetic polymorphisms of which the polymorphism profile of claim 1 may be comprised. Specifically, claim 2 recites use of a polymorphism profile that “comprises one or more genetic variations selected from single nucleotide polymorphisms (SNPs), variable number of tandem repeats (VNTRs), hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, simple sequence repeats, insertion elements, insertions, repeats, or deletions.” Dependent claim 11 recites wherein the polymorphism profile of claim 1 comprises at least one SNP.

171. Regarding the ’497 Patent, dependent claim 6 recites “at least ten different homozygous or heterozygous SNPs are detected.” Dependent claim 17 recites that “the SNP profile comprises informative homozygous and heterozygous SNPs.” Dependent claim 25 recites the homozygous or heterozygous SNP comprises a marker having at least two alleles, each occurring at a frequency greater than 1% of the population.”

172. Regarding the ’607 Patent, dependent claim 2 recites the “at least 1,000 [SNPs] comprise [SNPs] that occur at an allele frequency greater than 1% of a

population. Dependent claim 3 recites the "at least 1,000 [SNPs] comprise homozygous [SNPs]." Dependent claim 4 recites the "selective amplification of target [DNA] sequences amplifies a plurality of genomic regions comprising at least 5,000 [SNPs]." Dependent claim 5 recites the "selective amplification of target [DNA] sequences amplifies a plurality of genomic regions comprising at least 10,000 [SNPs]."

173. As explained above in section VII.B.2., genotyping to obtain a polymorphism profile comprising SNPs was well-understood, routine and conventional as of November 2009. As explained in section VII.B.4., it was also well-understood, routine, and conventional as of November 2009 to genotype homozygous and heterozygous SNPs, and SNPs that arise from various mutations including insertions, repeats, or deletions. In addition, the written description makes clear that it was conventional to genotype SNPs having alleles that each occur at a frequency greater than 1% of a population. B0001-B0024 ('652 Patent) at B0015, 11:22-31. Moreover, as explained above in section VII.C.1., it was conventional, including using commercially available kits and products, to genotype 5,000 SNPs, 10,000 SNPs, and many more. Accordingly, there is nothing nonconventional or innovative in the limitations recited in '652 Patent claims 2 and 11; '497 Patent claims 6, 17, 18, 24, and 25; '607 Patent claims 2, 3, 4, and 5.

2. Dependent Claims Reciting Different Types Of cfDNA

174. Regarding the '652 Patent, dependent claim 3 recites wherein the cell-free nucleic acids of claim 1 are DNA.

175. Regarding the '497 Patent, dependent claim 12 recites "the donor-specific circulating cell-free nucleic acids … in the biological sample are DNA, RNA, mRNA, miRNA, double-stranded DNA, single-stranded DNA, single-stranded DNA hairpins, DNA/RNA hybrids, RNA hairpins, or a combination thereof."

176. It was routine and conventional as of Nov. 2009 to assay these recited types of nucleic acids using the claimed methods. The Patents' written description, again, states only at a high level of generality that "[n]ucleic acids from samples that can be analyzed by the methods herein include: double-stranded DNA, single-stranded DNA, single-stranded DNA hairpins, DNA/RNA hybrids, RNA (e.g. mRNA or miRNA) and RNA hairpins." B0001-B0024 ('652 Patent) at B0014, 10:45-51; *see id.* at B0016, 13:17-21. The Patents do not disclose or claim any nonconventional laboratory techniques for analyzing any of these types of nucleic acids.

177. Neither the Patents' written description nor their claims indicate any nonconventional approach to assaying these particular types of nucleic acids, or any innovation based on analyzing any particular form of nucleic acid over another.

Accordingly, there is nothing nonconventional or innovative in the limitations recited in ’652 Patent claim 3 and ’497 Patent claim 12.

3. Dependent Claims Reciting Certain Common Modifications, Error Rates, Or Quality Scores Associated With Multiplexed Or High Throughput Sequencing

178. Regarding the ’652 Patent, dependent claims 4 and 6 recite that the multiplexed sequencing of claim 1 is shotgun sequencing (claim 4) and sequencing at least ten different nucleic acids (claim 6).

179. Regarding the ’497 Patent, dependent claim 3 recites a “high-throughput sequencing assay that generates at least 1,000 sequence reads per hour.” Dependent claim 4 recites “a high-throughput sequencing assay” that comprises “a next-generation sequencing assay.” Dependent claim 5 recites “the high-throughput sequencing assay generates sequencing reads of at least 36 bases.” Dependent claim 19 recites that the “sensitivity of the method is greater than 56%.” Dependent claim 21 recites a sequencing error rate of “less than 1.5%.” Dependent claim 23 recites using a “quality score.”

180. ’607 Patent dependent claim 6 recites the “high throughput sequencing reaction comprises assigning a quality score to bases of said provided sequences.”

181. As explained above in section VII.B.3., and is also evident from the literature and product-related publications, all of these limitations are either themselves conventional methods, or are features inherent in the standard use of the

conventional methods recited in the independent claims from which they depend. For example, shotgun sequencing and next-generation sequencing are types of multiplex/high throughput sequencing that were well understood, routine and conventional as of Nov. 2009. In addition, high-throughput sequencing to generate at least 1,000 reads per hour ('497 Patent claim 3) and read lengths at least 36 bases long ('497 Patent claim 5) was inherent in all high-throughput sequencers available as of November 6, 2009, as demonstrated for example by Margulies *et al*., which published that the instrument which became branded as the Illumina Genome Analyzer "is able to sequence 25 million bases … in one four-hour run," and generates an "[a]verage read length (bases)" of "108." B1305-1309 (Margulies 2005) at B1305 (Abstract); B1307, Table 1.

182. Sequencing at least ten different nucleic acids using multiplex sequencing also was a standard application of the commercial multiplex sequencers at the time. For example, product literature from 2007 for the Illumina Genome Analyzer system states that an "excess of 2.6 billion bases of high-quality filtered data per paired-end run on a single flow cell (as of October, 2007)" could be achieved" (B0463-B0466) (Illumina 2007) at B0463 and the Illumina Genome Analyzer product literature from 2008 states that multiplexed sequencing of 96 different samples on a single flow cell could be achieved using their instrument. B0467-B0470, (Illumina 2008) at B0467. Generation of at least 1,000 sequence

reads per hour, and reads of at least 36 bases, also was a standard feature of the conventional operation of commercial high throughput sequencers in 2009. The Patents' written description states only at a high level of generality, in the "Sequencing" section generally discussing the multiple sequencing instrumentalities available in the market at the time, that:

> In some cases, high throughput sequencing generates at least 1,000, at least 5,000, at least 10,000, at least 20,000, at least 30,000, at least 40,000, at least 50,000, at least 100,000, or at least 500,000 sequence reads per hour, with each read being at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 120, or at least 150 bases per read.

B0001-B0024 ('652 Patent) at B0017, 15:12-18. The Patents do not disclose or claim any nonconventional use of the multiplex/high-throughput sequencing methods elsewhere (including in that section) described in the Patents. A POSA in 2009 would understand this is an inherent feature in the standard use of the conventional multiplex/high-throughput sequencing machines referred to in the Patents.

183. The sequencing error rates and sensitivities also are inherent features of conventional use of the sequencing equipment, as discussed in section VII.B.3.c above.

184. Using a quality score likewise was conventional practice when using the standard high throughput sequencers available in 2009. B0467-B0470 (Illumina 2008) at B0467, B0469; B0471-B0480 (Li 2009b) at B0473. Moreover, I note that

the Patents' written description does not include any disclosure of nonconventional use of a quality score; rather, the only mention of "quality score" in the written description generically states that "[t]he use of quality scores for improved filtering of SNP calls, or the use of resequencing, should reduce error rate and increase sensitivity." B0001-B0024 ('652 Patent) at B0022, 26:51-53. The prior literature further described software that had already been developed and made commercially available before November 2009 to provide quality scores as recited in the Patent claims. *See, e.g.,* B0226-B0243 (Voelkerding 2009) at B0235 ("Image parameters such as intensity, background, and noise are then used in a platform-dependant algorithm to generate read sequences and error probability–related quality scores for each base. Although many researchers use the base calls generated by the platform-specific data-pipeline software, alternative base-calling programs that use more advanced software and statistical techniques have been developed.").

185. Neither the Patents' written description nor the claims identify anything nonconventional or innovative about the limitations recited in these dependent claims. To the contrary, they are routine applications, or inherent features, of the multiplex / high throughput sequencers when used in conventional ways, including in the 2009 timeframe. Accordingly, there is nothing nonconventional or innovative in the limitations recited in '652 Patent claims 4 and 6; '497 Patent claims 3, 4, 5, 19, 21, and 23; and '607 Patent claim 6.

4. Dependent Claims Reciting Organ or Species Limitations

186. Regarding the ’652 Patent, dependent claims 14 and 15 recite the transplant is selected from the group consisting of kidney, pancreas, liver, heart, lung, intestine, pancreas after kidney, and simultaneous pancreas-kidney (claim 14); or is heart or kidney (claim 15), respectively.

187. Regarding the ’497 Patent, dependent claim 9 recites wherein the transplant recipient is human.

188. Each of these additional steps was well-understood, routine and conventional as of Nov. 2009, particularly when used in conjunction with monitoring of transplant status or rejection. The Patents’ written description does not purport to attribute invention of any type of organ transplantation to the named inventors. Nor do the inventors contend they came up with the notion of analyzing the bodily fluid samples of human transplant recipients as opposed to those of some other species. *See* B0001-B0024 (’652 Patent) at B0012, 5:36-40. Accordingly, there is nothing nonconventional or innovative in the limitations recited in ’652 Patent claims 7, 8, 9, 14, and 15; and ’497 Patent claims 7, 8, 9, and 29.

5. Dependent Claims Reciting Conventional PCR Or Amplification Methods

189. Regarding the ’497 Patent, dependent claim 10 recites that “an amplification reaction is performed on the donor-specific circulating cell-free

nucleic acids in the biological sample prior to determining the amount of donor-specific circulating cell-free nucleic acids … in the biological sample."

190. As explained in section VII.B.2. above, performing an amplification reaction on donor cfDNA prior to quantifying it also was routine and conventional as of November 2009. Neither the Patents' written description nor the claims identify anything nonconventional or innovative about the limitation recited in this dependent claim. To the contrary, it was a standard application that was routinely used in the 2009 timeframe. Accordingly, there is nothing nonconventional or innovative in the limitation recited in '497 Patent claim 10.

6. Dependent Claims Reciting Different Sample Types

191. Regarding the '652 Patent, dependent claim 12 recites that the sample of claim 1 is blood or serum.

192. Regarding the '497 Patent, dependent claims 2 and 27 recite that the sample of claim 1 is blood (claim 2) or plasma (claim 27), respectively.

193. As explained above in section VII.B.1., obtaining a biological sample from a patient, including from blood, plasma, or serum, and including for use with the methods claimed in the Patent claims, was well-understood, routine and conventional as of Nov. 2009. Neither the Patents' written description nor the claims identify anything nonconventional or innovative about the limitations recited in these dependent claims. To the contrary, they are standard applications that were

routinely used in the 2009 timeframe. Accordingly, there is nothing nonconventional or innovative in the limitations recited in '652 Patent claim 12 and'497 Patent claims 2 and 27.

7. Dependent Claims Reciting Further Genotyping Prior To Quantifying Cell-Free Nucleic Acids

194. '497 Patent dependent claim 15 recites "further comprising genotyping [donor, recipient, or both donor and recipient] prior to the determining in step (d)." The written description and claims do not disclose any unconventional application of further genotyping prior to determining an amount of donor-specific cell-free nucleic acids in the biological sample (as recited in claim step 1(d)), or any way in which the prior genotyping improves upon the recited conventional methods. The only relevant disclosure in the written description states that "[i]n some embodiments, both the donor and recipient will be genotyped prior to transplantation." B0001-B0024 ('652 Patent) at B0013, 8:55-56; *see also id.* at B0016, 13:2-3 (same). This, however, is only a high-level generalization that does not provide any improvement over the routine and conventional techniques as recited in independent claim 1 from which the dependent claim depends.

8. Dependent Claims Reciting Certain Concentrations Of Cell-Free Nucleic Acids In The Sample

195. Regarding the '497 Patent, dependent claim 20 recites "the determining comprises detecting donor-specific circulating cell-free nucleic acids from the solid

organ transplant wherein the donor-specific circulating cell-free nucleic acids from the solid organ transplant make up between 0.03% and 8.0% of the total circulating cell-free nucleic acids in the biological sample."

196. Regarding the '607 Patent, dependent claims 7 and 8 recite the "quantified amount of said kidney transplant-derived [cfDNA] …" comprises certain percentages "of said total circulating [cfDNA]" in the samples, which are "between 1.5% and 8%" (claim 7) and "at least 0.05%" (claim 8), respectively.

197. As explained in section VII.B.4. above, the ability to detect any of these concentrations of cell-free nucleic acids in a sample is an inherent feature of the standard and conventional methods and instruments recited in the Patent claims. Neither the Patents' written description nor claims identify any nonconventional or innovative use of these conventional techniques that renders them able to detect or quantify cfDNA at the recited concentrations when they would not otherwise be able to do so. Accordingly, there is nothing nonconventional or innovative in the limitations recited in '497 Patent claim 20 and '607 Patent claims 7, 8, 9, and 10.

E. Concluding Opinions Regarding Conventionality of the Recited Combination of Laboratory Techniques

198. In my opinion, based on my review of the Patents and understanding of the state of the art at the time the Patents were filed, the claimed methods of the Patents employ routine and conventional techniques already in use to detect the natural phenomenon of transplant cfDNA. The same techniques recited in the

Patents were already in use for detecting fetal, cancer, and pathogen cfDNA, and those POSAs using the techniques for those applications recognized they would be equally applied in the transplant space.

199. The Patents' claims do not purport to improve upon or describe any nonconventional use of the recited existing combination of laboratory techniques. From the perspective of a POSA in the 2009 timeframe, the combination of laboratory techniques recited in the Patents was well understood, routine and conventional, and already in use to detect a range of natural phenomena.

VIII. THE PATENT METHODS DO NOT CREATE ANY NEW PREPARATION

200. As explained above, the Standard Patents are directed to the steps of (1) obtaining biological samples containing cell-free nucleic acids from a transplant recipient; (2) genotyping the transplant donor and/or recipient to establish profiles of genetic polymorphisms (or SNPs)[21]; (3) performing multiplex or high-throughput sequencing of the cell-free nucleic acids to detect the genotyped polymorphisms (or SNPs); and (4) quantifying the transplant (donor-derived) cell-free nucleic acids in the sample using the genetic differences in the sequences. The method claims are not directed to a novel method of preparing a new composition. The claimed method steps do nothing more than observe that cell-free DNA is present in a transplant

[21] The '607 Patent further recites selectively amplifying at least 1000 SNPs by PCR.

patient, and correlate the presence of the cell-free DNA with transplant rejection, using a conventional combination of techniques.

201. The claims recite the step of obtaining a biological sample containing cell-free nucleic acids from a transplant recipient. This step does not result in a novel composition, as it merely involves obtaining a naturally-occurring sample from a patient.

202. The claims also recite steps of genotyping the transplant donor and/or recipient to establish profiles of genetic polymorphisms and performing multiplex or high-throughput sequencing of the cell-free nucleic acids to detect the genotyped polymorphisms (or SNPs). The '607 Patent claims further recite selectively amplifying at least 1,000 SNPs by PCR in order to detect them. These steps use well-understood, routine and conventional—indeed commercially available—techniques for analyzing the naturally-occurring sample. They do not change the relative amounts of the transplant donor-specific and transplant recipient cfDNA being quantified in the claimed method.

203. Finally, the claims require quantifying the transplant (donor-specific) cell-free nucleic acids in the sample using the genetic differences in the sequences. This step also does not result in a new composition or preparation. As explained above, this quantification step is broadly recited with no further specificity and involves standard techniques for quantifying nucleic acids. This step also does not

change the relative amounts of the transplant donor-specific and transplant recipient cfDNA being quantified in the claimed method.

* * *

I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct.

Dated: __June 11, 2020_____ _____

Professor John Quackenbush

EXHIBIT 1

CURRICULUM VITAE

Date: June 10, 2020

Name: John Quackenbush, Ph.D.

Office Address: Department of Biostatistics, 655 Huntington Avenue, Building II, Room 419, Boston, MA 02115

Home Address: 86 Walpole Street, Dover, MA 02030, USA

Work e-mail: johnq@hsph.harvard.edu

Work Phone: +1.617.432.9028

Work FAX +1.617.432.5619

Mobile Phone: +1.301.526.1740

Date & Place of Birth: January 4, 1962 Kingston, PA, USA

Education:

1983	Physics	B.S.	California Institute of Technology, Pasadena, CA
1984	Physics	M.S.	University of California, Los Angeles, CA
1990	Theoretical Particle Physics	Ph.D.	University of California, Los Angeles, CA

Postdoctoral Training:

1990-1992 Postdoctoral Fellow, Experimental Particle Physics, University of California, Los Angeles, CA

1992-1994 Staff Scientist, Molecular Genetics and Genomics, The Salk Institute for Biological Studies, La Jolla, CA

1994-1996 Research Associate, Genomics and Bioinformatics, Stanford Human Genome Center, Palo Alto, CA

Academic Appointments:

2018- present Henry Pickering Walcott Professor of Computational Biology and Bioinformatics; Chair, Department of Biostatistics; Director, Harvard Health Data Science Center, Harvard TH Chan School of Public Health
Research Activities: Development of methods in computational and systems biology for applications in network medicine to understand the nature of complex disease and to identify new therapeutic interventions;

development of open-source software tools supporting research and clinical applications.
Activities: Director, MS Program in Computational Biology and Quantitative Genetics (CBQG); Director, BD2K T32 Training Grant, Former Director and Executive Committee, Joint Biostatistics and Bioinformatics T32 Training Grant

2005- present Professor of Biostatistics and Computational Biology,
Department of Data Science, Division of Computational Biology, Dana-Farber Cancer Institute;
Professor of Cancer Biology, Department of Cancer Biology, Dana-Farber Cancer Institute
Research Activities: Application of genomic technologies and integrative computational analysis and modeling of cellular systems to the understanding of human cancers and other diseases; development of open-source software tools supporting research and clinical applications.

2005- 2018 Professor of Computational Biology and Bioinformatics,
Department of Biostatistics, Harvard TH Chan School of Public Health
Research Activities: Application of genomic technologies and integrative computational analysis and modeling of cellular systems to the understanding of human cancers and other diseases.
Activities: Director, MS Program in Computational Biology and Quantitative Genetics (CBQG); Director, BD2K T32 Training Grant, Former Director and Executive Committee, Joint Biostatistics and Bioinformatics T32 Training Grant

2008- 2018 Director, Canter for Cancer Computational Biology,
Dana-Farber Cancer Institute
Research Activities: Integrative approaches to data analysis focused on clinical and translational applications and linking clinical and laboratory data in support of basic research; creation of software tools and systems to support research.

2007 – present Adjunct Professor of Bioinformatics, Boston University
Activities: Mentoring PhD student thesis research.

2003- 2005 Professor, Chemical Engineering, University of Maryland
Activities: Mentoring PhD student thesis research.

2002 - 2006 Investigator, The Institute for Genomic Research
Research Activities: Development and implementation of technology and strategies, including computational methods and software tools, necessary for functional analysis of the human and other genomes. Analysis of human gene expression in colon tumors using microarrays, rodent models of heart, lung, blood and sleep disorders and gene expression in Arabidopsis. Director of the TIGR Gene Index Project, providing freely-available databases and software to estimate gene content in more than 100 organisms.

2000- 2005 Professor, Department of Biochemistry,
The George Washington University
Activities: Instructor in various courses, curriculum committee for Genomics.

1998 – 2005 Lecturer, The Department of Biostatistics,
The Johns Hopkins University
Activities: Instructor in the Masters In Biotechnology Program

1997 - 2001 Associate Investigator, The Institute for Genomic Research Research
Activities: Development and implementation of technology and strategies, including computational methods and software tools, necessary for functional analysis of the human and other genomes. Analysis of human gene expression in colon tumors using microarrays, rodent models of heart, lung, blood and sleep disorders and gene expression in Arabidopsis. Director of the TIGR Gene Index Project, providing freely-available databases and software to estimate gene content in more than 50 organisms.

1997 Assistant Investigator, The Institute for Genomic Research
Research Activities: Development and implementation of technology and strategies, including analytical methods and software tools, necessary for functional analysis of the human and other genomes.

1994 - 1997 Research Associate, Stanford Human Genome Center
Stanford University
Research Activities: Project leader for development and implementation of a transposon-mediated strategy for large-scale genomic DNA sequencing including development and implementation of laboratory protocols, analytical methods, computer software, and instrumentation. Mapping, sequencing, and annotating regions of human chromosomes 4 and 21.

1992 - 1994 Staff Scientist, The Salk Institute for Biological Studies
Research Activities: Development of improved methods for DNA sequencing, combinatoric strategies and devices for screening large libraries, simulation and optimization of single pass sequencing strategies including genome sequence sampling. Project director for the STS content mapping of chromosome 11. Development of methods and software tools to support genome mapping.

1990 - 1992 Postdoctoral Fellow, Department of Physics
University of California, Los Angeles
Research Activities: Research in particle physics, field theory and phenomenology. Founding member of the Antiproton Experiment (APEX) collaboration that set the world's best experimental limit on the lifetime of the antiproton.

1987 - 1992 Visiting Lecturer, Department of Physics
University of California, Los Angeles

1984 - 1996 Physics Instructor, Southern California Science Institute
New College of California

1983 – 1990 Teaching Fellow, Department of Physics
University of California, Los Angeles

1983 - 1990 Research Fellow, Department of Physics
University of California, Los Angeles
Thesis Supervisor: E.T. Tomboulis
Research Activities: Research in elementary particle physics, mathematical physics, field theory focusing on the development of two-dimensional gauge field theory models and the construction of associated string theory models.

Nonacademic Employment

2011-2014 Co-Founder and CEO, Genospace LLC
Activities: Genospace provides 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.

2014-2017 Co-Founder and Board Chairman, Genospace LLC
Activities: Genospace provides 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. In January 2017, Genospace was sold to the Hospital Corporation of America as they were interested in acquiring our clinical trials matching capabilities.

Hospital or Affiliated Institution Appointments:

2003 – 2005 Adjunct Professor, Department of Biostatistics
Bloomberg School of Public Health, The Johns Hopkins University

2005- present Professor, Department of Biostatistics and Computational Biology
Dana-Farber Cancer Institute

2014- present Professor, Channing Division of Network Medicine, Brigham and Women's Hospital

Hospital and Health Care Organization Service Responsibilities:

2005- present Primary research areas are computational and system biology focused on the integration of diverse data types to provide insight into biological processes that drive human disease phenotypes.

Department of Data Science, Division of Computational Biology,
Dana-Farber Cancer Institute

Major Committee Assignments:

1994	National Laboratory Genome Project Visit Review Panel, DOE
1995	Caltech/BAC Library Construction Site Visit Review Panel, DOE
1996-1998	Bioinformatics Grant Review Panel, DOE
1997-2001	Grant Review Panel, DOE
1997	Hollander Fellowship Review Panel, DOE
1997	Five-Year Program Advisory Committee, NCRR
1997	Cancer Chromosome Anatomy Project program Steering Committee, NCI
1998	National Laboratory Functional Genomic Review Panel, DOE
1998	Full-Length cDNA Library Construction and Sequencing Advisory, Committee, NCI
1998	Functional Genomics Panel, Welcome Trust
1998	Genomics Grant Review Panel, NSF SBIR
1998	Special Program in Tropical Disease Research Review, UNDP/World Bank/WHO
1999-2000	Low-Dosage radiation Grant Review Panel, DOE
1999	Bioinformatics Review Panel, NSF
1999	Special Emphasis Review Panel, NHLBI
1999	Microarray Working Group Advisory Panel, NIDCR
1999	Special Emphasis Review Panel, NIMH
1999	Plant Biology Review Panel, NSF
1999-2007	Board of Directors, MGED
2000	Grant Review Panel, NIDA SBIR
2000	Plant Biology Review Panel, USDA
2000	Plant Biology Review Panel, NSF
2000	Exceptional Chromosome Regions Working Group, DOE
2000	Working Group on US Scientific Interactions, NSF
2000-2006	Genome/GCAT study section, NIH
2000-2003	PGA Coordinating Committee, HHLBI
2000-2003	PGA Bioinformatics Committee, NHLBI
2000-2003	PGA Microarray Committee, NHLBI
2000	Working Group on Australia Scientific Interactions, NSF
2001	Review Meeting RFP ES-01-01 National Center for Toxicogenomics (NCT) Microarray Resource, HIEHS
2001-2003	Plant Genome Scientific Advisory Board, University of Georgia
2001-2003	Innovative Technology Review Panel, NCI
2001-2002	Genome Study Section, NIH
2002-2006	GCAT Study Section, NIH
2001	Bioinformatics in Neuroscience and Sleep Research Workshop, NIDDK
2001	Plant Genome Site Visit Panel, NSF
2001-2003	BISTI Review Panel, NIH
2001-2005	Grant Review Panel, NCI
2002	Microarray Supplement Review Panel, NHLBI
2002-2006	Panel on Emerging Issues in Toxicogenomics, NRC

2004-2006	Panel on Applications of Toxicogenomics, NRC
2003-2010	Genome British Columbia Scientific Advisory Board
2005-2014	EPA Board of Scientific Counselors
2005-2011	St. Jude's Children's Research Hospital Scientific Advisory Board
2006-2007	Harvard School of Public Health, Lefkopoulou Award Committee Member
2006	Dana Farber Cancer Institute, Computational Biology Search Chair
2005	EU FP6 Grant Review Panel
2006-2007	Harvard School of Public Health, Seminar Committee Member
2006-2019	Harvard School of Public Health, Curriculum Committee Member
2006-2009	Harvard School of Public Health, Computational Biology Search Chair
2007-present	EU FP7 Genomics Grant Review Panel
2008-present	Lovelace Respiratory Research Institute, Scientific Advisory Board
2008-2010	NAS Panel on Collecting, Storing, and Distributing Biodata linked to Social Science Surveys
2009	RC1 Grant Review "Editorial Review" Panel
2009	RC2 Genomic Science Grant Review
2009	Epigenomics Roadmap Grant Review Panel
2009-2014	Epigenomics Roadmap Advisory Board
2010	NCI EUREKA Grant Review Panel
2010	Chair, NCRR S10 Grant Review Panel
2016	NASA ISS Genomic Applications Study Section
2012-present	Research Oversight Committed, Genome British Columbia
2005-present	Various NIH grant review panels, either as an ad hoc member or as an ad hoc reviewer. I typically serve on at least six per year and can provide a more detailed list on request.
2005-present	Various EU and Canadian grant review panels, either as an ad hoc member or as an ad hoc reviewer. I can assemble a detailed list on request.

Professional Societies:

1988	American Physical Society	Member
1988	APS Biophysics Division	Member
1988	APS International Physics Group	Member
1988	APS Division of Particles and Fields	Member
1988	World Federation of Scientists	Member
1990	Microarray Gene Expression Data Society (MGED)/ Functional Genomics Data Society (FGED)	Member Board Member President
	(2011-2014)	
1995	The Human Genome Organization (HUGO)	Member
1996	Association of Science Professionals (ASP)	Member
1998	American Association for the Advancement of Science (AAAS)	Member
1999	Center for the Study of the Evolution of Life (CSEOL)	Member
2005	American Association for Cancer Research	Member

Community Service Related to Professional Work:

I have presented more than 300 talks, courses, and workshops since joining Dana-Farber and the Harvard School of Public Health in 2005; detailed lists will be provided upon request. I have also taught numerous courses and workshops on the analysis of genomic data and have served as an organizer of the EMBO course on DNA microarray analysis for more than 5 years.

In addition, I have served on the Scientific Advisory Board of the St. Jude's Children's Research Hospital, the Scientific Advisory Board for the Lovelace Respiratory Research Institute, Research Oversight Committees for various Genome Canada programs, and the Board of Scientific Counselors of the Environmental Protection Agency. I have also been a member of advisory boards for a number of NIH initiatives, including the Roadmap Epigenome Project, the Extracellular RNA Communication Consortium, the Atopic Dermatitis Research Network (ADRN), and the Trans-Omics for Precision Medicine (TOPMed) program.

(Peer Review for Scientific Journals)

1994-2016	*Genomics*
1994-2009	*BioTechniques*
1996-2010	*Genome Research*
2000-2010	*Genome Biology*
1997-1998	*Analytical Biochemistry*
1997-present	*Nature*
1997-2010	*Nucleic Acids Research*
1999-2010	*Nature Genetics*
1999-2009	*EMBO Reports*
1999-2017	*Nature Biotechnology*
1999-present	*Science*
2000-2016	*Bioinformatics*
2000-2009	*Journal of Theoretical Biology*
2000-2009	*Mammalian Genomics*
2000-2006	*Gene*
2000-2016	*BMC Genomics, BMC Bioinformatics*
2001-2009	*Structural and Functional Genomics*
2001-2010	*Nature Reviews Genetics*
2001-2006	*American Journal of Applied Statistics*
2009-present	*PLoS Computational Biology*
2015-present	*Proceedings of the National Academy of Sciences*
2017-present	Molecular Systems Biology

Editorial Boards:

2000-2009	Editorial Board	*BioTechniques*
2000-2012	Associate Editor	*Bioinformatics*
2000-present	Editorial Board	*BMC Genomics*
2000-2006	Editorial Board	*Genomics*

2006-2017	Editor-in-Chief	*Genomics*
2015-present	Editorial Board	*Cancer Research*

Awards and Honors:

1985 - 1991	**Outstanding Teaching Award (7)**, UCLA Physics Department
1986	**Jun John Sakurai Scholarship**, 24th International School of Subnuclear Physics, Ettore Majorana Centre for Scientific Culture, Erice, Italy
1987	**UCLA Distinguished Teaching Assistant Award**, UCLA Academic Senate
1987	**Prize for Best Student** and **Prize for Best Scientific Secretary**, 25th International School of Subnuclear Physics, Ettore Majorana Centre for Scientific Culture, Erice, Italy
1988	**Graduate Distinguished Scholar Award**, UCLA Alumni Association
1992 - 1997	**SERCA Fellow**, National Human Genome Research Institute
2006	**George D. Wilbanks Lectureship in Gynecological Oncology**, University of South Florida College of Medicine
2007	**President's Distinguished Lectureship**, American Society of Reproductive Medicine
2007	**Leopold G. Koss Lectureship**, Universität Bern, Switzerland
2007	**Distinguished International Advisors** at IEE BIBE, the original and the oldest IEEE flagship international conference in Bioinformatics and Bioengineering.
2008	**Distinguished Lecture in Computer Science**, Wayne State University
2010	**IPM Distinguished Lecture**, Mt. Sinai Medical School, New York
2010	**The Ian Lawson Van Toch Memorial Seminar Series in Computational Biology**, Ontario Cancer Institute
2010	**Harvard-Australia Foundation Fellowship**
2010	**Bancroft Fellow-in-Residence**, Queensland Institute for Medical Research, Queensland, Australia
2010	**Jackson Memorial Fellowship**, Griffith University, Queensland, Australia
2011	**Australia Fellow**, National Health and Medical Research Council, Australia
2013	**White House Open Science Champion of Change**, Office of the President of the United States of America
2016	**Mentor of the Year Award**, Postdoc and Graduate Student Organization, Dana-Farber Cancer Institute
2018	**National Cancer Institute Outstanding Investigator Award** (R35)

Part II: Research, Teaching, and Clinical Contributions

A. Narrative report (*700 words or less*) of Research, Teaching, and Clinical Contributions. Please focus primarily on the areas in which most of your time and effort is spent.

My current research involves development of integrative analytical methods for the analysis of large-scale biomedical data, including genomic data, to better understand the nature of human diseases, including cancer and pulmonary diseases, with the goal of developing better methods for diagnosis and treatment. Two defining characteristics of my work are a commitment to reproducible research through development of open-source software implementations of analytical methods, and a commitment to studying sexual dimorphism in health and disease.

A longstanding component of my work is methods development for robust biomarkers. We have applied this to identifying reproducible predictive biomarkers in cancer and have explored reproducibility in large, published data sets, demonstrating that many are inconsistent. In collaboration with Hugo Aerts, I have been working to extend our methods to the analysis of quantitative feature data extracted from radiographic images (an emerging field called "radiomics"), to use these to identify predictive biomarkers, and to link these features to genetic mutations in cancer.

I have also been developing methods for inferring gene regulatory network models using gene expression and other sources of data. We published a method, PANDA, that builds on "message passing" (an idea developed communications theory) to optimize information flow through a gene regulatory network model, allowing us to infer condition-specific networks and to compare these between phenotypes. We have developed many extensions of PANDA that incorporate other data types and have applied these to the study of diseases, including ovarian cancer, colorectal cancer, chronic obstructive pulmonary disease, asthma, and Alzheimer's disease, as well as to the investigation of sexual dimorphism.

We have also been exploring the connection between genotype and phenotype, using gene expression as a surrogate for phenotype. We have developed a new approach to the analysis of expression quantitative trait loci that uses a bi-partite graph to represent associations between gene expression levels and both local (*cis-acting*) and distant (*trans-acting*) significant genetic variants. We find that the overall structure of the graph provides insight not only into how genetic variants work together to influence phenotype, but also into evolutionary constraints on the genetic variants themselves. This has given us new insight into how many genetic variants of small effect-size can collectively exert strong influence on phenotype.

Our group also builds and maintains a variety of software tools to support genomic analysis. the TM4 suite of open-source software tools for the analysis of DNA microarray data. The most widely used of these, MeV, is a powerful data mining software tool, that places sophisticated analytical tools behind an intuitive graphical user interface. Developed to be accessible to laboratory biologists, MeV has more than 260,000 users worldwide. We recently transitioned MeV to a cloud-based platform, WebMeV

(http://mev.tm4.org) to take advantage of cloud scalability to better handle large sequence-based data. WebMeV now has 3000 registered users and provides more than 100 analysis sessions per day.

I am also very involved in teaching and mentoring. I developed, launched, and direct the Computational Biology and Quantitative Genetics (CBQG) master's degree program, I was co-PI and co-director of the first joint Biostatistics and Bioinformatics training grant in the US and recently stepped down to develop and direct a training grant program as part of the NIH's Big Data to Knowledge (BD2K) program. I also developed and teach an introductory genomics tools and methods course and have co-taught an online course on reproducible research through HarvardX (https://www.edx.org/course/principles-statistical-computational-harvardx-ph527x).

I have trained more than 50 postdoctoral fellows and graduate students (including MS and PhD students) and has mentored a number of junior faculty members since joining the Harvard TH Chan School of Public Health and Dana-Farber Cancer Institute in 2005. I have emphasized diversity in hiring postdocs; more than half of my trainees have been female and three of the four postdocs from my group who have moved onto faculty positions in the past year have been female. I also maintains an extensive network of former trainees with whom I meet regularly as we continue to work collaboratively.

B. Funding Information

I have had more than twenty-five years of continuous grant support from the NIH and other organizations. My current support includes:

Active Research Support

5R01CA205406-03 (Ng, K.) 3/1/2017 – 2/28/2022

NIH/NCI
Novel Randomized Controlled Trials of Vitamin D Supplementation in Patients with Colorectal Cancer: Impact on Survival and Biology
This project is exploring the biology underlying the effects of vitamin D on survival of colorectal patients.
Role: Co-Investigator

5R01HL122805-05 (Ferrari, G.) 8/1/2017-6/30/2020

NIH/NHLBI (Subcontract: Columbia University)
Role of RAGE in Bicuspid Aortic Valve Syndrome
Dr. Quackenbush is working with Dr. Ferrari to model the effects of RAGE in perturbing molecular pathways in cardiovascular disorders associated with bicuspid aortic valve syndrome. Role: Co-Investigator

1R35CA220523-02 (Quackenbush, J.) 9/1/2018 – 8/31/2025

NIH/NCI
Unraveling the Complexities of Risk and Mechanism in Cancer
This NCI Career Award supports Dr. Quackenbush's ongoing work to model genetic processes in cancer and to elucidate the link between genotype and phenotype so as to better understand the drivers of cancer and to identify potential targets for therapeutic

intervention, with a particular emphasis on understanding differences between men and women. Role: Principal Investigator

5R01HL135142-03 (Hunninghake, G. / Cho, M.) 8/1/2017 – 7/30/2021

NIH/NHLBI

Genetic and Genomic Characterization of the Occurrence and Progression of Interstitial Lung Abnormalities

The goal of this project is to use genomic and clinical data to better understand the factors contributing to the development and progression of pulmonary fibrosis. Dr. Quackenbush will use computational and systems biology methods to model and analyze genetic and gene regulatory networks using data from patients suffering from interstitial lung abnormalities and pulmonary fibrosis. Role: Co-Investigator

1U24CA231846-01A1 (Quackenbush, J.) 9/1/2019 – 8/31/2024

NIH/NCI

MeV: A Robust Platform for Intuitive Genomic Data Analysis

As part of the Informatics Technology for Cancer Research (ITCR) program, we will develop and support WebMeV, a robust, scalable data analysis software tool that uses intuitive visual interfaces to provide users with access to advanced data analysis methods for large-scale genomic data in a system that helps ensure their research is reproducible. Role: Principal Investigator

2P50CA127003-11A1 (Bass, A.) 9/1/2019 – 5/31/2024

NIH (Subcontract: Dana Farber Cancer Institute

Dana Farber/ Harvard Cancer Center SPORE in Gastrointestinal Cancer

The GI SPORE supports state-of-the-art investigator-initiated translational research that will contribute to improved prevention, early detection, diagnosis, and treatment of gastrointestinal cancers. Dr. Quackenbush will serve as Co-Director of the Biostatistics and Bioinformatics Core, which will support the quantitative needs of the investigators in the design of experiments and the analysis of the resulting data.
Role: Co-Investigator/Core Co-Director

1R35CA220523-01A1 (Quackenbush, J.) 9/1/2018 – 8/31/2025 6.00 cal
NIH/NCI $628,715
Unraveling the Complexities of Risk and Mechanism in Cancer
This NCI Career Award supports ongoing work to model genetic processes in cancer and to elucidate the link between genotype and phenotype, with an emphasis on understanding differences between men and women.
Role: Principal Investigator

C. Report of Current Research Activities other than those mentioned above (bench research, clinical trials, outcome studies, efficacy studies as applicable)

None to report.

D. Report of Teaching (use only those categories that are applicable)

1983 - 1990	Teaching Fellow, Department of Physics, University of California, Los Angeles
1984 - 1996	Physics Instructor, Southern California Science Institute, New College of California
1987 - 1992	Visiting Lecturer, Department of Physics, University of California, Los Angeles
1998 - 2005	Lecturer, The Johns Hopkins University Instructor in the Masters in Biotechnology Program
2000 - 2005	Professor, Department of Biochemistry, The George Washington University
2003 - 2005	Adjunct Professor, Department of Biostatistics, Bloomberg School of Public Health, The Johns Hopkins University Developed program in genomics and bioinformatics; instructor in various courses.
2005-present	Professor, Department of Biostatistics, Harvard TH Chan School of Public Health
2013-present	Director, MS in Computational Biology and Quantitative Genetics

E. Report of Clinical Activities

1. Description of clinical practice (field, areas of major focus, site(s) of practice [private office, HMO, teaching hospital etc.])

None

2. Patient load (indicate complexity of cases, as appropriate)

None

3. Clinical contributions (e.g., introduction of new methods of clinical diagnosis, prevention, treatment, care delivery)

None

4. Other relevant information about clinical role (receipt of clinical awards, locally or nationally, invitation to participate in clinical activities at other sites, special recognition by peers or professional organizations as a leader in a clinical field)

None

Part III: Bibliography

Total citations: 71,575

h-index: 93

i10-index: 274

(Statistics retrieved November 28, 2019)

Original Articles (296)

1. **Quackenbush J**. Chiral anomalies in two-dimensional quantum field theory. Physical review D: Particles and fields. 1989;40(10):3408-14. Epub 1989/11/15. PubMed PMID: 10011709.
2. **Quackenbush J**. Twisted world-sheet gauge fields and string models Phys Lett B. 1990;234(3):285-91.
3. Buchanan CD, Cousins R, Dib C, Peccei RD, **Quackenbush J**. Testing CP and CPT violation in the neutral kaon system at a phi factory. Physical review D: Particles and fields. 1992;45(11):4088-107. Epub 1992/06/01. PubMed PMID: 10014315.
4. Geer S, Marriner J, Ray R, Streets J, Lindgren M, Muller T, **Quackenbush J**, Armstrong T. Search for antiproton decay at the Fermilab antiproton accumulator. Physical review letters. 1994;72(11):1596-9. Epub 1994/03/14. PubMed PMID: 10055652.
5. Khristich JV, Bailis J, Diggle K, Rodkins A, Romo A, **Quackenbush J**, Evans GA. Large-scale screening of yeast artificial chromosome libraries using PCR. BioTechniques. 1994;17(3):498-501. Epub 1994/09/01. PubMed PMID: 7818903.
6. **Quackenbush J**. Career issues. Science. 1994;266(5189):1306. Epub 1994/11/25. doi: 10.1126/science.266.5189.1306. PubMed PMID: 17772824.
7. Selleri L, Giovannini M, Hermanson GG, Romo A, **Quackenbush J**, Penny L, Khristich JV, Evans GA. Yeast artificial chromosome cloning of 3.2 megabases within chromosomal band 11q24 closely linking c-ets 1 and Fli-1 and encompassing the Ewing sarcoma breakpoint. Genomics. 1994;22(1):137-47. Epub 1994/07/01. doi: 10.1006/geno.1994.1354. PubMed PMID: 7959760.
8. Kupfer K, Smith MW, **Quackenbush J**, Evans GA. Physical mapping of complex genomes by sampled sequencing: a theoretical analysis. Genomics. 1995;27(1):90-100. Epub 1995/05/01. doi: 10.1006/geno.1995.1010. PubMed PMID: 7665188.

9. **Quackenbush J**, Davies C, Bailis JM, Khristich JV, Diggle K, Marchuck Y, Tobin J, Clark SP, Rodkins A, Marcano S, et al. An STS content map of human chromosome 11: localization of 910 YAC clones and 109 islands. Genomics. 1995;29(2):512-25. Epub 1995/09/20. PubMed PMID: 8666402.
10. Schuler GD, Boguski MS, Stewart EA, Stein LD, Gyapay G, Rice K, White RE, Rodriguez-Tome P, Aggarwal A, Bajorek E, Bentolila S, Birren BB, Butler A, Castle AB, Chiannilkulchai N, Chu A, Clee C, Cowles S, Day PJ, Dibling T, Drouot N, Dunham I, Duprat S, East C, Edwards C, Fan JB, Fang N, Fizames C, Garrett C, Green L, Hadley D, Harris M, Harrison P, Brady S, Hicks A, Holloway E, Hui L, Hussain S, Louis-Dit-Sully C, Ma J, MacGilvery A, Mader C, Maratukulam A, Matise TC, McKusick KB, Morissette J, Mungall A, Muselet D, Nusbaum HC, Page DC, Peck A, Perkins S, Piercy M, Qin F, **Quackenbush J**, Ranby S, Reif T, Rozen S, Sanders C, She X, Silva J, Slonim DK, Soderlund C, Sun WL, Tabar P, Thangarajah T, Vega-Czarny N, Vollrath D, Voyticky S, Wilmer T, Wu X, Adams MD, Auffray C, Walter NA, Brandon R, Dehejia A, Goodfellow PN, Houlgatte R, Hudson JR, Jr., Ide SE, Iorio KR, Lee WY, Seki N, Nagase T, Ishikawa K, Nomura N, Phillips C, Polymeropoulos MH, Sandusky M, Schmitt K, Berry R, Swanson K, Torres R, Venter JC, Sikela JM, Beckmann JS, Weissenbach J, Myers RM, Cox DR, James MR, Bentley D, Deloukas P, Lander ES, Hudson TJ. A gene map of the human genome. Science. 1996;274(5287):540-6. Epub 1996/10/25. PubMed PMID: 8849440.
11. Doyle DJ, **Quackenbush J**. Symposium on Genomic Medicine, University of Maryland, Shady Grove Campus, Rockville, Maryland, March 17-18, 1997. Microbial & comparative genomics. 1997;2(2):99-102. Epub 1997/01/01. PubMed PMID: 9689218.
12. Fraser CM, Casjens S, Huang WM, Sutton GG, Clayton R, Lathigra R, White O, Ketchum KA, Dodson R, Hickey EK, Gwinn M, Dougherty B, Tomb JF, Fleischmann RD, Richardson D, Peterson J, Kerlavage AR, **Quackenbush J**, Salzberg S, Hanson M, van Vugt R, Palmer N, Adams MD, Gocayne J, Weidman J, Utterback T, Watthey L, McDonald L, Artiach P, Bowman C, Garland S, Fuji C, Cotton MD, Horst K, Roberts K, Hatch B, Smith HO, Venter JC. Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi. Nature. 1997;390(6660):580-6. Epub 1997/12/24. doi: 10.1038/37551. PubMed PMID: 9403685.
13. Klenk HP, Clayton RA, Tomb JF, White O, Nelson KE, Ketchum KA, Dodson RJ, Gwinn M, Hickey EK, Peterson JD, Richardson DL, Kerlavage AR, Graham DE, Kyrpides NC, Fleischmann RD, **Quackenbush J**, Lee NH, Sutton GG, Gill S, Kirkness EF, Dougherty BA, McKenney K, Adams MD, Loftus B, Peterson S, Reich CI, McNeil LK, Badger JH, Glodek A, Zhou L, Overbeek R, Gocayne JD, Weidman JF, McDonald L, Utterback T, Cotton MD, Spriggs T, Artiach P, Kaine BP, Sykes SM, Sadow PW, D'Andrea KP, Bowman C, Fujii C, Garland SA, Mason TM, Olsen GJ, Fraser CM, Smith HO, Woese CR, Venter JC. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon Archaeoglobus fulgidus. Nature. 1997;390(6658):364-70. Epub 1997/12/06. doi: 10.1038/37052. PubMed PMID: 9389475.
14. Korenberg JR, Aaltonen J, Brahe C, Cabin D, Creau N, Delabar JM, Doering J, Gardiner K, Hubert RS, Ives J, Kessling A, Kudoh J, Lafreniere R, Murakami Y, Ohira M, Ohki M, Patterson D, Potier MC, **Quackenbush J**, Reeves RH, Sakaki Y,

Shimizu N, Soeda E, Van Broeckhoven C, Yaspo ML. Report and abstracts of the Sixth International Workshop on Human Chromosome 21 Mapping 1996. Cold Spring Harbor, New York, USA. May 6-8,1996. Cytogenetics and cell genetics. 1997;79(1-2):21-52. Epub 1997/01/01. PubMed PMID: 9533011.

15. Stewart EA, McKusick KB, Aggarwal A, Bajorek E, Brady S, Chu A, Fang N, Hadley D, Harris M, Hussain S, Lee R, Maratukulam A, O'Connor K, Perkins S, Piercy M, Qin F, Reif T, Sanders C, She X, Sun WL, Tabar P, Voyticky S, Cowles S, Fan JB, Mader C, **Quackenbush J**, Myers RM, Cox DR. An STS-based radiation hybrid map of the human genome. Genome research. 1997;7(5):422-33. Epub 1997/05/01. PubMed PMID: 9149939.
16. Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, Nelson K, **Quackenbush J**, Zhou L, Kirkness EF, Peterson S, Loftus B, Richardson D, Dodson R, Khalak HG, Glodek A, McKenney K, Fitzegerald LM, Lee N, Adams MD, Hickey EK, Berg DE, Gocayne JD, Utterback TR, Peterson JD, Kelley JM, Cotton MD, Weidman JM, Fujii C, Bowman C, Watthey L, Wallin E, Hayes WS, Borodovsky M, Karp PD, Smith HO, Fraser CM, Venter JC. The complete genome sequence of the gastric pathogen Helicobacter pylori. Nature. 1997;388(6642):539-47. Epub 1997/08/07. doi: 10.1038/41483. PubMed PMID: 9252185.
17. **Quackenbush J**, Adams MD, Fraser CM, Venter JC. Genome Projects:From Microbes to Man, From Sequence to Function. In: Dixon GK, Copping LG, Livingstone D, editors. Genomics: Commercial Opportunities from a Scientific Revolution: BIOS Scientific Publishers; 1998.
18. Kelly JM, **Quackenbush J**. Sequencing PCR Products. In: Innis M, Gelfand D, Sninsky J, editors. PCR Methods Manual: Academic PRess; 1999.
19. Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature. 2000;408:796-815.
20. El-Sayed NM, Hegde P, **Quackenbush J**, Melville SE, Donelson JE. The African trypanosome genome. International journal for parasitology. 2000;30(4):329-45. Epub 2000/03/25. PubMed PMID: 10731558.
21. Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, Hughes JE, Snesrud E, Lee N, **Quackenbush J**. A concise guide to cDNA microarray analysis. BioTechniques. 2000;29(3):548-50, 52-4, 56 passim. Epub 2000/09/21. PubMed PMID: 10997270.
22. Liang F, Holt I, Pertea G, Karamycheva S, Salzberg SL, **Quackenbush J**. Gene index analysis of the human genome estimates approximately 120,000 genes. Nature genetics. 2000;25(2):239-40. Epub 2000/06/03. doi: 10.1038/76126. PubMed PMID: 10835646.
23. Liang F, Holt I, Pertea G, Karamycheva S, Salzberg SL, **Quackenbush J**. An optimized protocol for analysis of EST sequences. Nucleic acids research. 2000;28(18):3657-65. Epub 2000/09/13. PubMed PMID: 10982889; PubMed Central PMCID: PMC110731.
24. **Quackenbush J**. Viva la revolution! A report from the FANTOM meeting. Nature genetics. 2000;26(3):255-6. Epub 2000/11/04. doi: 10.1038/81525. PubMed PMID: 11062453.
25. **Quackenbush J**, Liang F, Holt I, Pertea G, Upton J. The TIGR gene indices: reconstruction and representation of expressed gene sequences. Nucleic acids

research. 2000;28(1):141-5. Epub 1999/12/11. PubMed PMID: 10592205; PubMed Central PMCID: PMC102391.

26. Yuan Q, Liang F, Hsiao J, Zismann V, Benito MI, **Quackenbush J**, Wing R, Buell R. Anchoring of rice BAC clones to the rice genetic map in silico. Nucleic acids research. 2000;28(18):3636-41. Epub 2000/09/13. PubMed PMID: 10982886; PubMed Central PMCID: PMC110739.
27. Brazma A, Hingamp P, **Quackenbush J**, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nature genetics. 2001;29(4):365-71. Epub 2001/12/01. doi: 10.1038/ng1201-365. PubMed PMID: 11726920.
28. Flores-Morales A, Stahlberg N, Tollet-Egnell P, Lundeberg J, Malek RL, **Quackenbush J**, Lee NH, Norstedt G. Microarray analysis of the in vivo effects of hypophysectomy and growth hormone treatment on gene expression in the rat. Endocrinology. 2001;142(7):3163-76. Epub 2001/06/21. doi: 10.1210/endo.142.7.8235. PubMed PMID: 11416039.
29. Gaspard R, Dharap S, Malek J, Qi R, **Quackenbush J**. Optimized growth conditions for direct amplification of cDNA clone inserts from culture. BioTechniques. 2001;31(1):35-6. Epub 2001/07/24. PubMed PMID: 11464517.
30. Hegde P, Qi R, Gaspard R, Abernathy K, Dharap S, Earle-Hughes J, Gay C, Nwokekeh NU, Chen T, Saeed AI, Sharov V, Lee NH, Yeatman TJ, **Quackenbush J**. Identification of tumor markers in models of human colorectal cancer using a 19,200-element complementary DNA microarray. Cancer research. 2001;61(21):7792-7. Epub 2001/11/03. PubMed PMID: 11691794.
31. Huang J, Qi R, **Quackenbush J**, Dauway E, Lazaridis E, Yeatman T. Effects of ischemia on gene expression. The Journal of surgical research. 2001;99(2):222-7. Epub 2001/07/27. doi: 10.1006/jsre.2001.6195. PubMed PMID: 11469890.
32. Kappe SH, Gardner MJ, Brown SM, Ross J, Matuschewski K, Ribeiro JM, Adams JH, **Quackenbush J**, Cho J, Carucci DJ, Hoffman SL, Nussenzweig V. Exploring the transcriptome of the malaria sporozoite stage. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(17):9895-900. Epub 2001/08/09. doi: 10.1073/pnas.171185198. PubMed PMID: 11493695; PubMed Central PMCID: PMC55549.
33. Kawai J, Shinagawa A, Shibata K, Yoshino M, Itoh M, Ishii Y, Arakawa T, Hara A, Fukunishi Y, Konno H, Adachi J, Fukuda S, Aizawa K, Izawa M, Nishi K, Kiyosawa H, Kondo S, Yamanaka I, Saito T, Okazaki Y, Gojobori T, Bono H, Kasukawa T, Saito R, Kadota K, Matsuda H, Ashburner M, Batalov S, Casavant T, Fleischmann W, Gaasterland T, Gissi C, King B, Kochiwa H, Kuehl P, Lewis S, Matsuo Y, Nikaido I, Pesole G, **Quackenbush J**, Schriml LM, Staubli F, Suzuki R, Tomita M, Wagner L, Washio T, Sakai K, Okido T, Furuno M, Aono H, Baldarelli R, Barsh G, Blake J, Boffelli D, Bojunga N, Carninci P, de Bonaldo MF, Brownstein MJ, Bult C, Fletcher C, Fujita M, Gariboldi M, Gustincich S, Hill D, Hofmann M, Hume DA, Kamiya M, Lee NH, Lyons P, Marchionni L, Mashima J, Mazzarelli J, Mombaerts P, Nordone P, Ring B, Ringwald M, Rodriguez I, Sakamoto N, Sasaki H, Sato K, Schonbach C, Seya T, Shibata Y, Storch KF, Suzuki H, Toyo-oka K, Wang KH, Weitz C, Whittaker C, Wilming L, Wynshaw-

Boris A, Yoshida K, Hasegawa Y, Kawaji H, Kohtsuki S, Hayashizaki Y, Team RGERGPI, the FC. Functional annotation of a full-length mouse cDNA collection. Nature. 2001;409(6821):685-90. Epub 2001/02/24. doi: 10.1038/35055500. PubMed PMID: 11217851.

34. **Quackenbush J**. Crop Plant Genome Yield Huge Prospects. In: Butler D, editor. The Nature Yearbook of Science and Technology 2001 London, UK: Palgrave Macmillan; 2001.
35. **Quackenbush J**. Computational analysis of microarray data. Nature reviews Genetics. 2001;2(6):418-27. Epub 2001/06/05. doi: 10.1038/35076576. PubMed PMID: 11389458.
36. **Quackenbush J**. The power of public access: the human genome project and the scientific process. Nature genetics. 2001;29(1):4-6. Epub 2001/08/31. doi: 10.1038/ng0901-4. PubMed PMID: 11528377.
37. **Quackenbush J**, Cho J, Lee D, Liang F, Holt I, Karamycheva S, Parvizi B, Pertea G, Sultana R, White J. The TIGR Gene Indices: analysis of gene transcript sequences in highly sampled eukaryotic species. Nucleic acids research. 2001;29(1):159-64. Epub 2000/01/11. PubMed PMID: 11125077; PubMed Central PMCID: PMC29813.
38. Smith TP, Grosse WM, Freking BA, Roberts AJ, Stone RT, Casas E, Wray JE, White J, Cho J, Fahrenkrug SC, Bennett GL, Heaton MP, Laegreid WW, Rohrer GA, Chitko-McKown CG, Pertea G, Holt I, Karamycheva S, Liang F, **Quackenbush J**, Keele JW. Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle. Genome research. 2001;11(4):626-30. Epub 2001/04/03. doi: 10.1101/gr.170101. PubMed PMID: 11282978; PubMed Central PMCID: PMC311058.
39. Tsai J, Sultana R, Lee Y, Pertea G, Karamycheva S, Antonescu V, Cho J, Parvizi B, Cheung F, **Quackenbush J**. RESOURCERER: a database for annotating and linking microarray resources within and across species. Genome biology. 2001;2(11):SOFTWARE0002. Epub 2005/09/22. PubMed PMID: 16173164; PubMed Central PMCID: PMC138985.
40. Yuan Q, **Quackenbush J**, Sultana R, Pertea M, Salzberg SL, Buell CR. Rice bioinformatics. analysis of rice sequence data and leveraging the data to other plant species. Plant physiology. 2001;125(3):1166-74. Epub 2001/03/13. PubMed PMID: 11244096; PubMed Central PMCID: PMC1539370.
41. Agrawal D, Chen T, Irby R, **Quackenbush J**, Chambers AF, Szabo M, Cantor A, Coppola D, Yeatman TJ. Osteopontin identified as lead marker of colon cancer progression, using pooled sample expression profiling. Journal of the National Cancer Institute. 2002;94(7):513-21. Epub 2002/04/04. PubMed PMID: 11929952.
42. Andersson T, Unneberg P, Nilsson P, Odeberg J, **Quackenbush J**, Lundeberg J. Monitoring of representational difference analysis subtraction procedures by global microarrays. BioTechniques. 2002;32(6):1348-50, 52, 54-6, 58. Epub 2002/06/21. PubMed PMID: 12074166.
43. Ball CA, Sherlock G, Parkinson H, Rocca-Sera P, Brooksbank C, Causton HC, Cavalieri D, Gaasterland T, Hingamp P, Holstege F, Ringwald M, Spellman P, Stoeckert CJ, Jr., Stewart JE, Taylor R, Brazma A, **Quackenbush J**, Microarray Gene Expression D. The underlying principles of scientific publication. Bioinformatics. 2002;18(11):1409. Epub 2002/11/09. PubMed PMID: 12424109.

44. Ball CA, Sherlock G, Parkinson H, Rocca-Sera P, Brooksbank C, Causton HC, Cavalieri D, Gaasterland T, Hingamp P, Holstege F, Ringwald M, Spellman P, Stoeckert CJ, Jr., Stewart JE, Taylor R, Brazma A, **Quackenbush J**, Microarray Gene Expression Data S. Standards for microarray data. Science. 2002;298(5593):539. Epub 2002/10/22. PubMed PMID: 12387284.
45. Carlton JM, Angiuoli SV, Suh BB, Kooij TW, Pertea M, Silva JC, Ermolaeva MD, Allen JE, Selengut JD, Koo HL, Peterson JD, Pop M, Kosack DS, Shumway MF, Bidwell SL, Shallom SJ, van Aken SE, Riedmuller SB, Feldblyum TV, Cho JK, **Quackenbush J**, Sedegah M, Shoaibi A, Cummings LM, Florens L, Yates JR, Raine JD, Sinden RE, Harris MA, Cunningham DA, Preiser PR, Bergman LW, Vaidya AB, van Lin LH, Janse CJ, Waters AP, Smith HO, White OR, Salzberg SL, Venter JC, Fraser CM, Hoffman SL, Gardner MJ, Carucci DJ. Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii. Nature. 2002;419(6906):512-9. Epub 2002/10/09. doi: 10.1038/nature01099. PubMed PMID: 12368865.
46. Fahrenkrug SC, Smith TP, Freking BA, Cho J, White J, Vallet J, Wise T, Rohrer G, Pertea G, Sultana R, **Quackenbush J**, Keele JW. Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly. Mammalian genome : official journal of the International Mammalian Genome Society. 2002;13(8):475-8. Epub 2002/09/13. doi: 10.1007/s00335-001-2072-4. PubMed PMID: 12226715.
47. Kim H, Zhao B, Snesrud EC, Haas BJ, Town CD, **Quackenbush J**. Use of RNA and genomic DNA references for inferred comparisons in DNA microarray analyses. BioTechniques. 2002;33(4):924-30. Epub 2002/10/26. PubMed PMID: 12398202.
48. Lee Y, Sultana R, Pertea G, Cho J, Karamycheva S, Tsai J, Parvizi B, Cheung F, Antonescu V, White J, Holt I, Liang F, **Quackenbush J**. Cross-referencing eukaryotic genomes: TIGR Orthologous Gene Alignments (TOGA). Genome research. 2002;12(3):493-502. Epub 2002/03/05. doi: 10.1101/gr.212002. PubMed PMID: 11875039; PubMed Central PMCID: PMC155294.
49. Malek RL, Irby RB, Guo QM, Lee K, Wong S, He M, Tsai J, Frank B, Liu ET, **Quackenbush J**, Jove R, Yeatman TJ, Lee NH. Identification of Src transformation fingerprint in human colon cancer. Oncogene. 2002;21(47):7256-65. Epub 2002/10/09. doi: 10.1038/sj.onc.1205900. PubMed PMID: 12370817.
50. Nene V, Lee D, **Quackenbush J**, Skilton R, Mwaura S, Gardner MJ, Bishop R. AvGI, an index of genes transcribed in the salivary glands of the ixodid tick Amblyomma variegatum. International journal for parasitology. 2002;32(12):1447-56. Epub 2002/10/24. PubMed PMID: 12392910.
51. Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, Nikaido I, Osato N, Saito R, Suzuki H, Yamanaka I, Kiyosawa H, Yagi K, Tomaru Y, Hasegawa Y, Nogami A, Schonbach C, Gojobori T, Baldarelli R, Hill DP, Bult C, Hume DA, **Quackenbush J**, Schriml LM, Kanapin A, Matsuda H, Batalov S, Beisel KW, Blake JA, Bradt D, Brusic V, Chothia C, Corbani LE, Cousins S, Dalla E, Dragani TA, Fletcher CF, Forrest A, Frazer KS, Gaasterland T, Gariboldi M, Gissi C, Godzik A, Gough J, Grimmond S, Gustincich S, Hirokawa N, Jackson IJ, Jarvis ED, Kanai A, Kawaji H, Kawasawa Y, Kedzierski RM, King BL, Konagaya A, Kurochkin IV, Lee Y, Lenhard B, Lyons PA, Maglott DR, Maltais L, Marchionni L, McKenzie L, Miki H, Nagashima T, Numata K, Okido T, Pavan WJ, Pertea G, Pesole G, Petrovsky N, Pillai R, Pontius JU, Qi D, Ramachandran S, Ravasi T,

Reed JC, Reed DJ, Reid J, Ring BZ, Ringwald M, Sandelin A, Schneider C, Semple CA, Setou M, Shimada K, Sultana R, Takenaka Y, Taylor MS, Teasdale RD, Tomita M, Verardo R, Wagner L, Wahlestedt C, Wang Y, Watanabe Y, Wells C, Wilming LG, Wynshaw-Boris A, Yanagisawa M, Yang I, Yang L, Yuan Z, Zavolan M, Zhu Y, Zimmer A, Carninci P, Hayatsu N, Hirozane-Kishikawa T, Konno H, Nakamura M, Sakazume N, Sato K, Shiraki T, Waki K, Kawai J, Aizawa K, Arakawa T, Fukuda S, Hara A, Hashizume W, Imotani K, Ishii Y, Itoh M, Kagawa I, Miyazaki A, Sakai K, Sasaki D, Shibata K, Shinagawa A, Yasunishi A, Yoshino M, Waterston R, Lander ES, Rogers J, Birney E, Hayashizaki Y, Consortium F, I RGERGP, Team II. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature. 2002;420(6915):563-73. Epub 2002/12/06. doi: 10.1038/nature01266. PubMed PMID: 12466851.

52. **Quackenbush J**. Microarray data normalization and transformation. Nature genetics. 2002;32 Suppl:496-501. Epub 2002/11/28. doi: 10.1038/ng1032. PubMed PMID: 12454644.
53. Sonstegard TS, Capuco AV, White J, Van Tassell CP, Connor EE, Cho J, Sultana R, Shade L, Wray JE, Wells KD, **Quackenbush J**. Analysis of bovine mammary gland EST and functional annotation of the Bos taurus gene index. Mammalian genome : official journal of the International Mammalian Genome Society. 2002;13(7):373-9. Epub 2002/07/26. doi: 10.1007/s00335-001-2145-4. PubMed PMID: 12140684.
54. Sturn A, **Quackenbush J**, Trajanoski Z. Genesis: cluster analysis of microarray data. Bioinformatics. 2002;18(1):207-8. Epub 2002/02/12. PubMed PMID: 11836235.
55. Yang IV, Chen E, Hasseman JP, Liang W, Frank BC, Wang S, Sharov V, Saeed AI, White J, Li J, Lee NH, Yeatman TJ, **Quackenbush J**. Within the fold: assessing differential expression measures and reproducibility in microarray assays. Genome biology. 2002;3(11):research0062. Epub 2002/11/14. PubMed PMID: 12429061; PubMed Central PMCID: PMC133446.
56. Abulencia JP, Gaspard R, Healy ZR, Gaarde WA, **Quackenbush J**, Konstantopoulos K. Shear-induced cyclooxygenase-2 via a JNK2/c-Jun-dependent pathway regulates prostaglandin receptor expression in chondrocytic cells. The Journal of biological chemistry. 2003;278(31):28388-94. Epub 2003/05/14. doi: 10.1074/jbc.M301378200. PubMed PMID: 12743126.
57. Agrawal D, Chen T, Irby R, **Quackenbush J**, Chambers AF, Szabo M, Cantor A, Coppola D, Yeatman TJ. Osteopontin identified as colon cancer tumor progression marker. Comptes rendus biologies. 2003;326(10-11):1041-3. Epub 2004/01/28. PubMed PMID: 14744111.
58. Brentani H, Caballero OL, Camargo AA, da Silva AM, da Silva WA, Jr., Dias Neto E, Grivet M, Gruber A, Guimaraes PE, Hide W, Iseli C, Jongeneel CV, Kelso J, Nagai MA, Ojopi EP, Osorio EC, Reis EM, Riggins GJ, Simpson AJ, de Souza S, Stevenson BJ, Strausberg RL, Tajara EH, Verjovski-Almeida S, Acencio ML, Bengtson MH, Bettoni F, Bodmer WF, Briones MR, Camargo LP, Cavenee W, Cerutti JM, Coelho Andrade LE, Costa dos Santos PC, Ramos Costa MC, da Silva IT, Estecio MR, Sa Ferreira K, Furnari FB, Faria M, Jr., Galante PA, Guimaraes GS, Holanda AJ, Kimura ET, Leerkes MR, Lu X, Maciel RM, Martins EA, Massirer KB, Melo AS, Mestriner CA, Miracca EC, Miranda LL, Nobrega FG, Oliveira PS, Paquola AC, Pandolfi JR, Campos Pardini MI, Passetti F, **Quackenbush J**, Schnabel B, Sogayar MC, Souza JE, Valentini SR, Zaiats AC,

Amaral EJ, Arnaldi LA, de Araujo AG, de Bessa SA, Bicknell DC, Ribeiro de Camaro ME, Carraro DM, Carrer H, Carvalho AF, Colin C, Costa F, Curcio C, Guerreiro da Silva ID, Pereira da Silva N, Dellamano M, El-Dorry H, Espreafico EM, Scattone Ferreira AJ, Ayres Ferreira C, Fortes MA, Gama AH, Giannella-Neto D, Giannella ML, Giorgi RR, Goldman GH, Goldman MH, Hackel C, Ho PL, Kimura EM, Kowalski LP, Krieger JE, Leite LC, Lopes A, Luna AM, Mackay A, Mari SK, Marques AA, Martins WK, Montagnini A, Mourao Neto M, Nascimento AL, Neville AM, Nobrega MP, O'Hare MJ, Otsuka AY, Ruas de Melo AI, Paco-Larson ML, Guimaraes Pereira G, Pereira da Silva N, Pesquero JB, Pessoa JG, Rahal P, Rainho CA, Rodrigues V, Rogatto SR, Romano CM, Romeiro JG, Rossi BM, Rusticci M, Guerra de Sa R, Sant' Anna SC, Sarmazo ML, Silva TC, Soares FA, Sonati Mde F, de Freitas Sousa J, Queiroz D, Valente V, Vettore AL, Villanova FE, Zago MA, Zalcberg H, Human Cancer Genome Project/Cancer Genome Anatomy Project Annotation C, Human Cancer Genome Project Sequencing C. The generation and utilization of a cancer-oriented representation of the human transcriptome by using expressed sequence tags. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(23):13418-23. Epub 2003/11/01. doi: 10.1073/pnas.1233632100. PubMed PMID: 14593198; PubMed Central PMCID: PMC263829.

59. Causton HC, **Quackenbush J**, Brazma A. Microarray gene expressions data analysis: a beginner's guide. Malden, MA: Blackwell Pub.; 2003. viii, 160 p. p.
60. Chen T, Yang I, Irby R, Shain KH, Wang HG, **Quackenbush J**, Coppola D, Cheng JQ, Yeatman TJ. Regulation of caspase expression and apoptosis by adenomatous polyposis coli. Cancer research. 2003;63(15):4368-74. Epub 2003/08/09. PubMed PMID: 12907606.
61. Cook DN, Wang S, Howles GP, Speer M, Churchhill G, **Quackenbush J**, Schwartz DA. The genetics of innate immunity in the lung. Chest. 2003;123(3 Suppl):369S. Epub 2003/03/12. PubMed PMID: 12628980.
62. Dudoit S, Gentleman RC, **Quackenbush J**. Open source software for the analysis of microarray data. BioTechniques. 2003;Suppl:45-51. Epub 2003/04/01. PubMed PMID: 12664684.
63. Kasukawa T, Furuno M, Nikaido I, Bono H, Hume DA, Bult C, Hill DP, Baldarelli R, Gough J, Kanapin A, Matsuda H, Schriml LM, Hayashizaki Y, Okazaki Y, **Quackenbush J**. Development and evaluation of an automated annotation pipeline and cDNA annotation system. Genome research. 2003;13(6B):1542-51. Epub 2003/06/24. doi: 10.1101/gr.992803. PubMed PMID: 12819153; PubMed Central PMCID: PMC403710.
64. Kim H, Snesrud EC, Haas B, Cheung F, Town CD, **Quackenbush J**. Gene expression analyses of Arabidopsis chromosome 2 using a genomic DNA amplicon microarray. Genome research. 2003;13(3):327-40. Epub 2003/03/06. doi: 10.1101/gr.552003. PubMed PMID: 12618363; PubMed Central PMCID: PMC430289.
65. Klapa MI, **Quackenbush J**. The quest for the mechanisms of life. Biotechnology and bioengineering. 2003;84(7):739-42. Epub 2004/01/07. doi: 10.1002/bit.10858. PubMed PMID: 14708113.
66. Lee Y, **Quackenbush J**. Using the TIGR gene index databases for biological discovery. Current protocols in bioinformatics / editoral board, Andreas D

Baxevanis [et al]. 2003;Chapter 1:Unit 1 6. Epub 2008/04/23. doi: 10.1002/0471250953.bi0106s03. PubMed PMID: 18428690.

67. Merrick JM, Osman A, Tsai J, **Quackenbush J**, LoVerde PT, Lee NH. The Schistosoma mansoni gene index: gene discovery and biology by reconstruction and analysis of expressed gene sequences. The Journal of parasitology. 2003;89(2):261-9. Epub 2003/05/23. doi: 10.1645/0022-3395(2003)089[0261:TSMGIG]2.0.CO;2. PubMed PMID: 12760639.
68. Pertea G, Huang X, Liang F, Antonescu V, Sultana R, Karamycheva S, Lee Y, White J, Cheung F, Parvizi B, Tsai J, **Quackenbush J**. TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. Bioinformatics. 2003;19(5):651-2. Epub 2003/03/26. PubMed PMID: 12651724.
69. **Quackenbush J**. John Quackenbush talks about the clinical promise of genetic microarrays. Interviewed by Brian Vastag. JAMA : the journal of the American Medical Association. 2003;289(2):159-60, 63. Epub 2003/01/09. PubMed PMID: 12517207.
70. **Quackenbush J**. Open-source software accelerates bioinformatics. Genome biology. 2003;4(9):336. Epub 2003/09/04. doi: 10.1186/gb-2003-4-9-336. PubMed PMID: 12952530; PubMed Central PMCID: PMC193651.
71. **Quackenbush J**. Genomics. Microarrays--guilt by association. Science. 2003;302(5643):240-1. Epub 2003/10/11. doi: 10.1126/science.1090887. PubMed PMID: 14551426.
72. Rexroad CE, 3rd, Lee Y, Keele JW, Karamycheva S, Brown G, Koop B, Gahr SA, Palti Y, **Quackenbush J**. Sequence analysis of a rainbow trout cDNA library and creation of a gene index. Cytogenetic and genome research. 2003;102(1-4):347-54. Epub 2004/02/19. doi: 75773. PubMed PMID: 14970727.
73. Ronning CM, Stegalkina SS, Ascenzi RA, Bougri O, Hart AL, Utterbach TR, Vanaken SE, Riedmuller SB, White JA, Cho J, Pertea GM, Lee Y, Karamycheva S, Sultana R, Tsai J, **Quackenbush J**, Griffiths HM, Restrepo S, Smart CD, Fry WE, Van Der Hoeven R, Tanksley S, Zhang P, Jin H, Yamamoto ML, Baker BJ, Buell CR. Comparative analyses of potato expressed sequence tag libraries. Plant physiology. 2003;131(2):419-29. Epub 2003/02/15. doi: 10.1104/pp.013581. PubMed PMID: 12586867; PubMed Central PMCID: PMC166819.
74. Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, **Quackenbush J**. TM4: a free, open-source system for microarray data management and analysis. BioTechniques. 2003;34(2):374-8. Epub 2003/03/05. PubMed PMID: 12613259.
75. Wang HY, Malek RL, Kwitek AE, Greene AS, Luu TV, Behbahani B, Frank B, **Quackenbush J**, Lee NH. Assessing unmodified 70-mer oligonucleotide probe performance on glass-slide microarrays. Genome biology. 2003;4(1):R5. Epub 2003/01/24. PubMed PMID: 12540297; PubMed Central PMCID: PMC151289.
76. Wells CA, Ravasi T, Sultana R, Yagi K, Carninci P, Bono H, Faulkner G, Okazaki Y, **Quackenbush J**, Hume DA, Lyons PA, Group RG, Members GSL. Continued discovery of transcriptional units expressed in cells of the mouse mononuclear phagocyte lineage. Genome research. 2003;13(6B):1360-5. Epub 2003/06/24. doi: 10.1101/gr.1056103. PubMed PMID: 12819134; PubMed Central PMCID: PMC403663.

77. Whitelaw CA, Barbazuk WB, Pertea G, Chan AP, Cheung F, Lee Y, Zheng L, van Heeringen S, Karamycheva S, Bennetzen JL, SanMiguel P, Lakey N, Bedell J, Yuan Y, Budiman MA, Resnick A, Van Aken S, Utterback T, Riedmuller S, Williams M, Feldblyum T, Schubert K, Beachy R, Fraser CM, **Quackenbush J**. Enrichment of gene-coding sequences in maize by genome filtration. Science. 2003;302(5653):2118-20. Epub 2003/12/20. doi: 10.1126/science.1090047. PubMed PMID: 14684821.
78. Yuan Q, Ouyang S, Liu J, Suh B, Cheung F, Sultana R, Lee D, **Quackenbush J**, Buell CR. The TIGR rice genome annotation resource: annotating the rice genome and creating resources for plant biologists. Nucleic acids research. 2003;31(1):229-33. Epub 2003/01/10. PubMed PMID: 12519988; PubMed Central PMCID: PMC165506.
79. Zhu Y, King BL, Parvizi B, Brunk BP, Stoeckert CJ, Jr., **Quackenbush J**, Richardson J, Bult CJ. Integrating computationally assembled mouse transcript sequences with the Mouse Genome Informatics (MGI) database. Genome biology. 2003;4(2):R16. Epub 2003/03/07. PubMed PMID: 12620126; PubMed Central PMCID: PMC151306.
80. Ball C, Brazma A, Causton H, Chervitz S, Edgar R, Hingamp P, Matese JC, Icahn C, Parkinson H, **Quackenbush J**, Ringwald M, Sansone SA, Sherlock G, Spellman P, Stoeckert C, Tateno Y, Taylor R, White J, Winegarden N, Microarray Gene Expression Data S. An open letter on microarray data from the MGED Society. Microbiology. 2004;150(Pt 11):3522-4. Epub 2004/11/06. doi: 10.1099/mic.0.27637-0. PubMed PMID: 15528642.
81. Ball C, Brazma A, Causton H, Chervitz S, Edgar R, Hingamp P, Matese JC, Parkinson H, **Quackenbush J**, Ringwald M, Sansone SA, Sherlock G, Spellman P, Stoeckert C, Tateno Y, Taylor R, White J, Winegarden N, Society M. Standards for microarray data: an open letter. Environmental health perspectives. 2004;112(12):A666-7. Epub 2004/09/04. PubMed PMID: 15345376; PubMed Central PMCID: PMC1277123.
82. Ball CA, Brazma A, Causton H, Chervitz S, Edgar R, Hingamp P, Matese JC, Parkinson H, **Quackenbush J**, Ringwald M, Sansone SA, Sherlock G, Spellman P, Stoeckert C, Tateno Y, Taylor R, White J, Winegarden N. Submission of microarray data to public repositories. PLoS biology. 2004;2(9):E317. Epub 2004/09/02. doi: 10.1371/journal.pbio.0020317. PubMed PMID: 15340489; PubMed Central PMCID: PMC514887.
83. Bloom G, Yang IV, Boulware D, Kwong KY, Coppola D, Eschrich S, **Quackenbush J**, Yeatman TJ. Multi-platform, multi-site, microarray-based human tumor classification. The American journal of pathology. 2004;164(1):9-16. Epub 2003/12/26. doi: 10.1016/S0002-9440(10)63090-8. PubMed PMID: 14695313; PubMed Central PMCID: PMC1602228.
84. Cook DN, Wang S, Wang Y, Howles GP, Whitehead GS, Berman KG, Church TD, Frank BC, Gaspard RM, Yu Y, **Quackenbush J**, Schwartz DA. Genetic regulation of endotoxin-induced airway disease. Genomics. 2004;83(6):961-9. Epub 2004/06/05. doi: 10.1016/j.ygeno.2003.12.008. PubMed PMID: 15177550.
85. Cronin M, Ghosh K, Sistare F, **Quackenbush J**, Vilker V, O'Connell C. Universal RNA reference materials for gene expression. Clinical chemistry. 2004;50(8):1464-71. Epub 2004/05/25. doi: 10.1373/clinchem.2004.035675. PubMed PMID: 15155546.

86. Ghedin E, Pumfery A, de la Fuente C, Yao K, Miller N, Lacoste V, **Quackenbush J**, Jacobson S, Kashanchi F. Use of a multi-virus array for the study of human viral and retroviral pathogens: gene expression studies and ChIP-chip analysis. Retrovirology. 2004;1:10. Epub 2004/06/01. doi: 10.1186/1742-4690-1-10. PubMed PMID: 15169557; PubMed Central PMCID: PMC442135.
87. Grigoryev DN, Ma SF, Irizarry RA, Ye SQ, **Quackenbush J**, Garcia JG. Orthologous gene-expression profiling in multi-species models: search for candidate genes. Genome biology. 2004;5(5):R34. Epub 2004/05/07. doi: 10.1186/gb-2004-5-5-r34. PubMed PMID: 15128448; PubMed Central PMCID: PMC416470.
88. Imanishi T, Itoh T, Suzuki Y, O'Donovan C, Fukuchi S, Koyanagi KO, Barrero RA, Tamura T, Yamaguchi-Kabata Y, Tanino M, Yura K, Miyazaki S, Ikeo K, Homma K, Kasprzyk A, Nishikawa T, Hirakawa M, Thierry-Mieg J, Thierry-Mieg D, Ashurst J, Jia L, Nakao M, Thomas MA, Mulder N, Karavidopoulou Y, Jin L, Kim S, Yasuda T, Lenhard B, Eveno E, Suzuki Y, Yamasaki C, Takeda J, Gough C, Hilton P, Fujii Y, Sakai H, Tanaka S, Amid C, Bellgard M, Bonaldo Mde F, Bono H, Bromberg SK, Brookes AJ, Bruford E, Carninci P, Chelala C, Couillault C, de Souza SJ, Debily MA, Devignes MD, Dubchak I, Endo T, Estreicher A, Eyras E, Fukami-Kobayashi K, Gopinath GR, Graudens E, Hahn Y, Han M, Han ZG, Hanada K, Hanaoka H, Harada E, Hashimoto K, Hinz U, Hirai M, Hishiki T, Hopkinson I, Imbeaud S, Inoko H, Kanapin A, Kaneko Y, Kasukawa T, Kelso J, Kersey P, Kikuno R, Kimura K, Korn B, Kuryshev V, Makalowska I, Makino T, Mano S, Mariage-Samson R, Mashima J, Matsuda H, Mewes HW, Minoshima S, Nagai K, Nagasaki H, Nagata N, Nigam R, Ogasawara O, Ohara O, Ohtsubo M, Okada N, Okido T, Oota S, Ota M, Ota T, Otsuki T, Piatier-Tonneau D, Poustka A, Ren SX, Saitou N, Sakai K, Sakamoto S, Sakate R, Schupp I, Servant F, Sherry S, Shiba R, Shimizu N, Shimoyama M, Simpson AJ, Soares B, Steward C, Suwa M, Suzuki M, Takahashi A, Tamiya G, Tanaka H, Taylor T, Terwilliger JD, Unneberg P, Veeramachaneni V, Watanabe S, Wilming L, Yasuda N, Yoo HS, Stodolsky M, Makalowski W, Go M, Nakai K, Takagi T, Kanehisa M, Sakaki Y, **Quackenbush J**, Okazaki Y, Hayashizaki Y, Hide W, Chakraborty R, Nishikawa K, Sugawara H, Tateno Y, Chen Z, Oishi M, Tonellato P, Apweiler R, Okubo K, Wagner L, Wiemann S, Strausberg RL, Isogai T, Auffray C, Nomura N, Gojobori T, Sugano S. Integrative annotation of 21,037 human genes validated by full-length cDNA clones. PLoS biology. 2004;2(6):e162. Epub 2004/04/23. doi: 10.1371/journal.pbio.0020162. PubMed PMID: 15103394; PubMed Central PMCID: PMC393292.
89. Larkin JE, Frank BC, Gaspard RM, Duka I, Gavras H, **Quackenbush J**. Cardiac transcriptional response to acute and chronic angiotensin II treatments. Physiological genomics. 2004;18(2):152-66. Epub 2004/05/06. doi: 10.1152/physiolgenomics.00057.2004. PubMed PMID: 15126644.
90. Nene V, Lee D, Kang'a S, Skilton R, Shah T, de Villiers E, Mwaura S, Taylor D, **Quackenbush J**, Bishop R. Genes transcribed in the salivary glands of female Rhipicephalus appendiculatus ticks infected with Theileria parva. Insect biochemistry and molecular biology. 2004;34(10):1117-28. Epub 2004/10/12. doi: 10.1016/j.ibmb.2004.07.002. PubMed PMID: 15475305.
91. **Quackenbush J**. Data standards for 'omic' science. Nature biotechnology. 2004;22(5):613-4. Epub 2004/05/04. doi: 10.1038/nbt0504-613. PubMed PMID: 15122299.

92. Sharov V, Kwong KY, Frank B, Chen E, Hasseman J, Gaspard R, Yu Y, Yang I, **Quackenbush J**. The limits of log-ratios. BMC biotechnology. 2004;4:3. Epub 2004/04/29. doi: 10.1186/1472-6750-4-3. PubMed PMID: 15113428; PubMed Central PMCID: PMC400743.
93. Bergman NH, Passalacqua KD, Gaspard R, Shetron-Rama LM, **Quackenbush J**, Hanna PC. Murine macrophage transcriptional responses to Bacillus anthracis infection and intoxication. Infection and immunity. 2005;73(2):1069-80. Epub 2005/01/25. doi: 10.1128/IAI.73.2.1069-1080.2005. PubMed PMID: 15664951; PubMed Central PMCID: PMC547061.
94. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C, Kodzius R, Shimokawa K, Bajic VB, Brenner SE, Batalov S, Forrest AR, Zavolan M, Davis MJ, Wilming LG, Aidinis V, Allen JE, Ambesi-Impiombato A, Apweiler R, Aturaliya RN, Bailey TL, Bansal M, Baxter L, Beisel KW, Bersano T, Bono H, Chalk AM, Chiu KP, Choudhary V, Christoffels A, Clutterbuck DR, Crowe ML, Dalla E, Dalrymple BP, de Bono B, Della Gatta G, di Bernardo D, Down T, Engstrom P, Fagiolini M, Faulkner G, Fletcher CF, Fukushima T, Furuno M, Futaki S, Gariboldi M, Georgii-Hemming P, Gingeras TR, Gojobori T, Green RE, Gustincich S, Harbers M, Hayashi Y, Hensch TK, Hirokawa N, Hill D, Huminiecki L, Iacono M, Ikeo K, Iwama A, Ishikawa T, Jakt M, Kanapin A, Katoh M, Kawasawa Y, Kelso J, Kitamura H, Kitano H, Kollias G, Krishnan SP, Kruger A, Kummerfeld SK, Kurochkin IV, Lareau LF, Lazarevic D, Lipovich L, Liu J, Liuni S, McWilliam S, Madan Babu M, Madera M, Marchionni L, Matsuda H, Matsuzawa S, Miki H, Mignone F, Miyake S, Morris K, Mottagui-Tabar S, Mulder N, Nakano N, Nakauchi H, Ng P, Nilsson R, Nishiguchi S, Nishikawa S, Nori F, Ohara O, Okazaki Y, Orlando V, Pang KC, Pavan WJ, Pavesi G, Pesole G, Petrovsky N, Piazza S, Reed J, Reid JF, Ring BZ, Ringwald M, Rost B, Ruan Y, Salzberg SL, Sandelin A, Schneider C, Schonbach C, Sekiguchi K, Semple CA, Seno S, Sessa L, Sheng Y, Shibata Y, Shimada H, Shimada K, Silva D, Sinclair B, Sperling S, Stupka E, Sugiura K, Sultana R, Takenaka Y, Taki K, Tammoja K, Tan SL, Tang S, Taylor MS, Tegner J, Teichmann SA, Ueda HR, van Nimwegen E, Verardo R, Wei CL, Yagi K, Yamanishi H, Zabarovsky E, Zhu S, Zimmer A, Hide W, Bult C, Grimmond SM, Teasdale RD, Liu ET, Brusic V, **Quackenbush J**, Wahlestedt C, Mattick JS, Hume DA, Kai C, Sasaki D, Tomaru Y, Fukuda S, Kanamori-Katayama M, Suzuki M, Aoki J, Arakawa T, Iida J, Imamura K, Itoh M, Kato T, Kawaji H, Kawagashira N, Kawashima T, Kojima M, Kondo S, Konno H, Nakano K, Ninomiya N, Nishio T, Okada M, Plessy C, Shibata K, Shiraki T, Suzuki S, Tagami M, Waki K, Watahiki A, Okamura-Oho Y, Suzuki H, Kawai J, Hayashizaki Y, Consortium F, Group RGER, Genome Science G. The transcriptional landscape of the mammalian genome. Science. 2005;309(5740):1559-63. Epub 2005/09/06. doi: 10.1126/science.1112014. PubMed PMID: 16141072.
95. Cui L, Fan Q, Hu Y, Karamycheva SA, **Quackenbush J**, Khuntirat B, Sattabongkot J, Carlton JM. Gene discovery in Plasmodium vivax through sequencing of ESTs from mixed blood stages. Molecular and biochemical parasitology. 2005;144(1):1-9. Epub 2005/08/09. doi: 10.1016/j.molbiopara.2005.05.016. PubMed PMID: 16085323.
96. Djebbari A, Karamycheva S, Howe E, **Quackenbush J**. MeSHer: identifying biological concepts in microarray assays based on PubMed references and MeSH

terms. Bioinformatics. 2005;21(15):3324-6. Epub 2005/05/28. doi: 10.1093/bioinformatics/bti503. PubMed PMID: 15919728.

97. Eschrich S, Yang I, Bloom G, Kwong KY, Boulware D, Cantor A, Coppola D, Kruhoffer M, Aaltonen L, Orntoft TF, **Quackenbush J**, Yeatman TJ. Molecular staging for survival prediction of colorectal cancer patients. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2005;23(15):3526-35. Epub 2005/05/24. doi: 10.1200/JCO.2005.00.695. PubMed PMID: 15908663.
98. Guerrero FD, Miller RJ, Rousseau ME, Sunkara S, **Quackenbush J**, Lee Y, Nene V. BmiGI: a database of cDNAs expressed in Boophilus microplus, the tropical/southern cattle tick. Insect biochemistry and molecular biology. 2005;35(6):585-95. Epub 2005/04/29. doi: 10.1016/j.ibmb.2005.01.020. PubMed PMID: 15857764.
99. Hackl H, Burkard TR, Sturn A, Rubio R, Schleiffer A, Tian S, **Quackenbush J**, Eisenhaber F, Trajanoski Z. Molecular processes during fat cell development revealed by gene expression profiling and functional annotation. Genome biology. 2005;6(13):R108. Epub 2006/01/20. doi: 10.1186/gb-2005-6-13-r108. PubMed PMID: 16420668; PubMed Central PMCID: PMC1414107.
100. Irizarry RA, Warren D, Spencer F, Kim IF, Biswal S, Frank BC, Gabrielson E, Garcia JG, Geoghegan J, Germino G, Griffin C, Hilmer SC, Hoffman E, Jedlicka AE, Kawasaki E, Martinez-Murillo F, Morsberger L, Lee H, Petersen D, **Quackenbush J**, Scott A, Wilson M, Yang Y, Ye SQ, Yu W. Multiple-laboratory comparison of microarray platforms. Nature methods. 2005;2(5):345-50. Epub 2005/04/23. doi: 10.1038/nmeth756. PubMed PMID: 15846361.
101. Kwong KY, Bloom GC, Yang I, Boulware D, Coppola D, Haseman J, Chen E, McGrath A, Makusky AJ, Taylor J, Steiner S, Zhou J, Yeatman TJ, **Quackenbush J**. Synchronous global assessment of gene and protein expression in colorectal cancer progression. Genomics. 2005;86(2):142-58. Epub 2005/06/14. doi: 10.1016/j.ygeno.2005.03.012. PubMed PMID: 15951154.
102. Larkin JE, Frank BC, Gavras H, Sultana R, **Quackenbush J**. Independence and reproducibility across microarray platforms. Nature methods. 2005;2(5):337-44. Epub 2005/04/23. doi: 10.1038/nmeth757. PubMed PMID: 15846360.
103. Lee Y, Tsai J, Sunkara S, Karamycheva S, Pertea G, Sultana R, Antonescu V, Chan A, Cheung F, **Quackenbush J**. The TIGR Gene Indices: clustering and assembling EST and known genes and integration with eukaryotic genomes. Nucleic acids research. 2005;33(Database issue):D71-4. Epub 2004/12/21. doi: 10.1093/nar/gki064. PubMed PMID: 15608288; PubMed Central PMCID: PMC540018.
104. Liu F, Vantoai T, Moy LP, Bock G, Linford LD, **Quackenbush J**. Global transcription profiling reveals comprehensive insights into hypoxic response in Arabidopsis. Plant physiology. 2005;137(3):1115-29. Epub 2005/03/01. doi: 10.1104/pp.104.055475. PubMed PMID: 15734912; PubMed Central PMCID: PMC1065411.
105. Margolin AA, Greshock J, Naylor TL, Mosse Y, Maris JM, Bignell G, Saeed AI, **Quackenbush J**, Weber BL. CGHAnalyzer: a stand-alone software package for cancer genome analysis using array-based DNA copy number data. Bioinformatics. 2005;21(15):3308-11. Epub 2005/05/21. doi: 10.1093/bioinformatics/bti500. PubMed PMID: 15905276.

106. Marko NF, Frank B, **Quackenbush J**, Lee NH. A robust method for the amplification of RNA in the sense orientation. BMC genomics. 2005;6:27. Epub 2005/03/03. doi: 10.1186/1471-2164-6-27. PubMed PMID: 15740627; PubMed Central PMCID: PMC554769.
107. Pinent M, Blade MC, Salvado MJ, Arola L, Hackl H, **Quackenbush J**, Trajanoski Z, Ardevol A. Grape-seed derived procyanidins interfere with adipogenesis of 3T3-L1 cells at the onset of differentiation. Int J Obes (Lond). 2005;29(8):934-41. Epub 2005/05/27. doi: 10.1038/sj.ijo.0802988. PubMed PMID: 15917849.
108. **Quackenbush J**. Extracting meaning from functional genomics experiments. Toxicology and applied pharmacology. 2005;207(2 Suppl):195-9. Epub 2005/07/09. doi: 10.1016/j.taap.2005.04.029. PubMed PMID: 16002114.
109. Stanley Kim H, Yu Y, Snesrud EC, Moy LP, Linford LD, Haas BJ, Nierman WC, **Quackenbush J**. Transcriptional divergence of the duplicated oxidative stress-responsive genes in the Arabidopsis genome. The Plant journal : for cell and molecular biology. 2005;41(2):212-20. Epub 2005/01/07. doi: 10.1111/j.1365-313X.2004.02295.x. PubMed PMID: 15634198.
110. Brooksbank C, **Quackenbush J**. Data standards: a call to action. Omics : a journal of integrative biology. 2006;10(2):94-9. Epub 2006/08/12. doi: 10.1089/omi.2006.10.94. PubMed PMID: 16901212.
111. Bruggmann R, Bharti AK, Gundlach H, Lai J, Young S, Pontaroli AC, Wei F, Haberer G, Fuks G, Du C, Raymond C, Estep MC, Liu R, Bennetzen JL, Chan AP, Rabinowicz PD, **Quackenbush J**, Barbazuk WB, Wing RA, Birren B, Nusbaum C, Rounsley S, Mayer KF, Messing J. Uneven chromosome contraction and expansion in the maize genome. Genome research. 2006;16(10):1241-51. Epub 2006/08/12. doi: 10.1101/gr.5338906. PubMed PMID: 16902087; PubMed Central PMCID: PMC1581433.
112. Chan AP, Pertea G, Cheung F, Lee D, Zheng L, Whitelaw C, Pontaroli AC, SanMiguel P, Yuan Y, Bennetzen J, Barbazuk WB, **Quackenbush J**, Rabinowicz PD. The TIGR Maize Database. Nucleic acids research. 2006;34(Database issue):D771-6. Epub 2005/12/31. doi: 10.1093/nar/gkj072. PubMed PMID: 16381977; PubMed Central PMCID: PMC1347435.
113. Chan VS, Chau SY, Tian L, Chen Y, Kwong SK, **Quackenbush J**, Dallman M, Lamb J, Tam PK. Sonic hedgehog promotes CD4+ T lymphocyte proliferation and modulates the expression of a subset of CD28-targeted genes. International immunology. 2006;18(12):1627-36. Epub 2006/09/29. doi: 10.1093/intimm/dxl096. PubMed PMID: 17005629.
114. Deutsch EW, Ball CA, Bova GS, Brazma A, Bumgarner RE, Campbell D, Causton HC, Christiansen J, Davidson D, Eichner LJ, Goo YA, Grimmond S, Henrich T, Johnson MH, Korb M, Mills JC, Oudes A, Parkinson HE, Pascal LE, **Quackenbush J**, Ramialison M, Ringwald M, Sansone SA, Sherlock G, Stoeckert CJ, Jr., Swedlow J, Taylor RC, Walashek L, Zhou Y, Liu AY, True LD. Development of the Minimum Information Specification for In Situ Hybridization and Immunohistochemistry Experiments (MISFISHIE). Omics : a journal of integrative biology. 2006;10(2):205-8. Epub 2006/08/12. doi: 10.1089/omi.2006.10.205. PubMed PMID: 16901227.
115. Frith MC, Bailey TL, Kasukawa T, Mignone F, Kummerfeld SK, Madera M, Sunkara S, Furuno M, Bult CJ, **Quackenbush J**, Kai C, Kawai J, Carninci P, Hayashizaki Y, Pesole G, Mattick JS. Discrimination of non-protein-coding

transcripts from protein-coding mRNA. RNA biology. 2006;3(1):40-8. Epub 2006/11/23. PubMed PMID: 17114936.

116. Hu VW, Frank BC, Heine S, Lee NH, **Quackenbush J**. Gene expression profiling of lymphoblastoid cell lines from monozygotic twins discordant in severity of autism reveals differential regulation of neurologically relevant genes. BMC genomics. 2006;7:118. Epub 2006/05/20. doi: 10.1186/1471-2164-7-118. PubMed PMID: 16709250; PubMed Central PMCID: PMC1525191.
117. Maeda N, Kasukawa T, Oyama R, Gough J, Frith M, Engstrom PG, Lenhard B, Aturaliya RN, Batalov S, Beisel KW, Bult CJ, Fletcher CF, Forrest AR, Furuno M, Hill D, Itoh M, Kanamori-Katayama M, Katayama S, Katoh M, Kawashima T, **Quackenbush J**, Ravasi T, Ring BZ, Shibata K, Sugiura K, Takenaka Y, Teasdale RD, Wells CA, Zhu Y, Kai C, Kawai J, Hume DA, Carninci P, Hayashizaki Y. Transcript annotation in FANTOM3: mouse gene catalog based on physical cDNAs. PLoS genetics. 2006;2(4):e62. Epub 2006/05/10. doi: 10.1371/journal.pgen.0020062. PubMed PMID: 16683036; PubMed Central PMCID: PMC1449903.
118. Maher EA, Brennan C, Wen PY, Durso L, Ligon KL, Richardson A, Khatry D, Feng B, Sinha R, Louis DN, **Quackenbush J**, Black PM, Chin L, DePinho RA. Marked genomic differences characterize primary and secondary glioblastoma subtypes and identify two distinct molecular and clinical secondary glioblastoma entities. Cancer research. 2006;66(23):11502-13. Epub 2006/11/23. doi: 10.1158/0008-5472.CAN-06-2072. PubMed PMID: 17114236.
119. Malek RL, Wang HY, Kwitek AE, Greene AS, Bhagabati N, Borchardt G, Cahill L, Currier T, Frank B, Fu X, Hasinoff M, Howe E, Letwin N, Luu TV, Saeed A, Sajadi H, Salzberg SL, Sultana R, Thiagarajan M, Tsai J, Veratti K, White J, **Quackenbush J**, Jacob HJ, Lee NH. Physiogenomic resources for rat models of heart, lung and blood disorders. Nature genetics. 2006;38(2):234-9. Epub 2006/01/18. doi: 10.1038/ng1693. PubMed PMID: 16415889.
120. Mar JC, Rubio R, **Quackenbush J**. Inferring steady state single-cell gene expression distributions from analysis of mesoscopic samples. Genome biology. 2006;7(12):R119. Epub 2006/12/16. doi: 10.1186/gb-2006-7-12-r119. PubMed PMID: 17169148; PubMed Central PMCID: PMC1794432.
121. **Quackenbush J**. Standardizing the standards. Molecular systems biology. 2006;2:2006 0010. Epub 2006/06/02. doi: 10.1038/msb4100052. PubMed PMID: 16738556; PubMed Central PMCID: PMC1681484.
122. **Quackenbush J**. Microarray analysis and tumor classification. The New England journal of medicine. 2006;354(23):2463-72. Epub 2006/06/09. doi: 10.1056/NEJMra042342. PubMed PMID: 16760446.
123. **Quackenbush J**. From 'omes to biology. Animal genetics. 2006;37 Suppl 1:48-56. Epub 2006/08/05. doi: 10.1111/j.1365-2052.2006.01476.x. PubMed PMID: 16887002.
124. **Quackenbush J**. Computational approaches to analysis of DNA microarray data. Yearbook of medical informatics. 2006:91-103. Epub 2006/10/20. PubMed PMID: 17051302.
125. **Quackenbush J**. Weighing our measures of gene expression. Molecular systems biology. 2006;2:63. Epub 2006/11/15. doi: 10.1038/msb4100096. PubMed PMID: 17102808; PubMed Central PMCID: PMC1682020.

126. **Quackenbush J**, Irizarry RA. Response to Shields: 'MIAME, we have a problem'. Trends in genetics : TIG. 2006;22(9):471-2. Epub 2006/07/25. doi: 10.1016/j.tig.2006.07.007. PubMed PMID: 16859802.
127. **Quackenbush J**, Salzberg SL. It is time to end the patenting of software. Bioinformatics. 2006;22(12):1416-7. Epub 2006/06/13. doi: 10.1093/bioinformatics/btl167. PubMed PMID: 16766564; PubMed Central PMCID: PMC2836512.
128. **Quackenbush J**, Stoeckert C, Ball C, Brazma A, Gentleman R, Huber W, Irizarry R, Salit M, Sherlock G, Spellman P, Winegarden N. Top-down standards will not serve systems biology. Nature. 2006;440(7080):24. Epub 2006/03/03. doi: 10.1038/440024a. PubMed PMID: 16511469.
129. Rayner TF, Rocca-Serra P, Spellman PT, Causton HC, Farne A, Holloway E, Irizarry RA, Liu J, Maier DS, Miller M, Petersen K, **Quackenbush J**, Sherlock G, Stoeckert CJ, Jr., White J, Whetzel PL, Wymore F, Parkinson H, Sarkans U, Ball CA, Brazma A. A simple spreadsheet-based, MIAME-supportive format for microarray data: MAGE-TAB. BMC bioinformatics. 2006;7:489. Epub 2006/11/08. doi: 10.1186/1471-2105-7-489. PubMed PMID: 17087822; PubMed Central PMCID: PMC1687205.
130. Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, Li J, Thiagarajan M, White JA, **Quackenbush J**. TM4 microarray software suite. Methods in enzymology. 2006;411:134-93. Epub 2006/08/31. doi: 10.1016/S0076-6879(06)11009-5. PubMed PMID: 16939790.
131. Stoeckert C, Ball C, Brazma A, Brinkman R, Causton H, Fan L, Fostel J, Fragoso G, Heiskanen M, Holstege F, Morrison N, Parkinson H, **Quackenbush J**, Rocca-Serra P, Sansone SA, Sarkans U, Sherlock G, Stevens R, Taylor C, Taylor R, Whetzel P, White J. Wrestling with SUMO and bio-ontologies. Nature biotechnology. 2006;24(1):21-2; author reply 3. Epub 2006/01/13. doi: 10.1038/nbt0106-21a. PubMed PMID: 16404382.
132. Walker JK, Ahumada A, Frank B, Gaspard R, Berman K, **Quackenbush J**, Schwartz DA. Multistrain genetic comparisons reveal CCR5 as a receptor involved in airway hyperresponsiveness. American journal of respiratory cell and molecular biology. 2006;34(6):711-8. Epub 2006/02/14. doi: 10.1165/rcmb.2005-0314OC. PubMed PMID: 16474097; PubMed Central PMCID: PMC2644233.
133. Chan AP, Rabinowicz PD, **Quackenbush J**, Buell CR, Town CD. Plant database resources at The Institute for Genomic Research. Methods Mol Biol. 2007;406:113-36. Epub 2008/02/22. PubMed PMID: 18287690.
134. Danley PD, Mullen SP, Liu F, Nene V, **Quackenbush J**, Shaw KL. A cricket Gene Index: a genomic resource for studying neurobiology, speciation, and molecular evolution. BMC genomics. 2007;8:109. Epub 2007/04/27. doi: 10.1186/1471-2164-8-109. PubMed PMID: 17459168; PubMed Central PMCID: PMC1878485.
135. Hollingsworth JW, Whitehead G, Berman KG, Tekippe EM, Gilmour MI, Larkin JE, **Quackenbush J**, Schwartz DA. Genetic basis of murine antibacterial defense to streptococcal lung infection. Immunogenetics. 2007;59(9):713-24. Epub 2007/08/19. doi: 10.1007/s00251-007-0242-6. PubMed PMID: 17701033.
136. **Quackenbush J**. Extracting biology from high-dimensional biological data. The Journal of experimental biology. 2007;210(Pt 9):1507-17. Epub 2007/04/24. doi: 10.1242/jeb.004432. PubMed PMID: 17449816.

137. Raj JU, Aliferis C, Caprioli RM, Cowley AW, Jr., Davies PF, Duncan MW, Erle DJ, Erzurum SC, Finn PW, Ischiropoulos H, Kaminski N, Kleeberger SR, Leikauf GD, Loyd JE, Martin TR, Matalon S, Moore JH, **Quackenbush J**, Sabo-Attwood T, Shapiro SD, Schnitzer JE, Schwartz DA, Schwiebert LM, Sheppard D, Ware LB, Weiss ST, Whitsett JA, Wurfel MM, Matthay MA. Genomics and proteomics of lung disease: conference summary. American journal of physiology Lung cellular and molecular physiology. 2007;293(1):L45-51. Epub 2007/05/01. doi: 10.1152/ajplung.00139.2007. PubMed PMID: 17468134; PubMed Central PMCID: PMC4212816.
138. Song JS, Maghsoudi K, Li W, Fox E, **Quackenbush J**, Shirley Liu X. Microarray blob-defect removal improves array analysis. Bioinformatics. 2007;23(8):966-71. Epub 2007/03/03. doi: 10.1093/bioinformatics/btm043. PubMed PMID: 17332024.
139. Aryee MJ, **Quackenbush J**. An optimized predictive strategy for interactome mapping. Journal of proteome research. 2008;7(9):4089-94. Epub 2008/07/23. doi: 10.1021/pr700858e. PubMed PMID: 18642945.
140. Boedigheimer MJ, Wolfinger RD, Bass MB, Bushel PR, Chou JW, Cooper M, Corton JC, Fostel J, Hester S, Lee JS, Liu F, Liu J, Qian HR, **Quackenbush J**, Pettit S, Thompson KL. Sources of variation in baseline gene expression levels from toxicogenomics study control animals across multiple laboratories. BMC genomics. 2008;9:285. Epub 2008/06/14. doi: 10.1186/1471-2164-9-285. PubMed PMID: 18549499; PubMed Central PMCID: PMC2453529.
141. Cheng AS, Culhane AC, Chan MW, Venkataramu CR, Ehrich M, Nasir A, Rodriguez BA, Liu J, Yan PS, **Quackenbush J**, Nephew KP, Yeatman TJ, Huang TH. Epithelial progeny of estrogen-exposed breast progenitor cells display a cancer-like methylome. Cancer research. 2008;68(6):1786-96. Epub 2008/03/15. doi: 10.1158/0008-5472.CAN-07-5547. PubMed PMID: 18339859; PubMed Central PMCID: PMC4172329.
142. Chittenden TW, Howe EA, Culhane AC, Sultana R, Taylor JM, Holmes C, **Quackenbush J**. Functional classification analysis of somatically mutated genes in human breast and colorectal cancers. Genomics. 2008;91(6):508-11. Epub 2008/04/25. doi: 10.1016/j.ygeno.2008.03.002. PubMed PMID: 18434084; PubMed Central PMCID: PMC2492759.
143. Deutsch EW, Ball CA, Berman JJ, Bova GS, Brazma A, Bumgarner RE, Campbell D, Causton HC, Christiansen JH, Daian F, Dauga D, Davidson DR, Gimenez G, Goo YA, Grimmond S, Henrich T, Herrmann BG, Johnson MH, Korb M, Mills JC, Oudes AJ, Parkinson HE, Pascal LE, Pollet N, **Quackenbush J**, Ramialison M, Ringwald M, Salgado D, Sansone SA, Sherlock G, Stoeckert CJ, Jr., Swedlow J, Taylor RC, Walashek L, Warford A, Wilkinson DG, Zhou Y, Zon LI, Liu AY, True LD. Minimum information specification for in situ hybridization and immunohistochemistry experiments (MISFISHIE). Nature biotechnology. 2008;26(3):305-12. Epub 2008/03/11. doi: 10.1038/nbt1391. PubMed PMID: 18327244; PubMed Central PMCID: PMC4367930.
144. Djebbari A, **Quackenbush J**. Seeded Bayesian Networks: constructing genetic networks from microarray data. BMC systems biology. 2008;2:57. Epub 2008/07/08. doi: 10.1186/1752-0509-2-57. PubMed PMID: 18601736; PubMed Central PMCID: PMC2474592.
145. Kahvejian A, **Quackenbush J**, Thompson JF. What would you do if you could sequence everything? Nature biotechnology. 2008;26(10):1125-33. Epub

2008/10/11. doi: 10.1038/nbt1494. PubMed PMID: 18846086; PubMed Central PMCID: PMC4153598.

146. Kitami T, Rubio R, O'Brien W, **Quackenbush J**, Nadeau JH. Gene-environment interactions reveal a homeostatic role for cholesterol metabolism during dietary folate perturbation in mice. Physiological genomics. 2008;35(2):182-90. Epub 2008/08/14. doi: 10.1152/physiolgenomics.00294.2007. PubMed PMID: 18697859; PubMed Central PMCID: PMC2575958.
147. Nosho K, Irahara N, Shima K, Kure S, Kirkner GJ, Schernhammer ES, Hazra A, Hunter DJ, **Quackenbush J**, Spiegelman D, Giovannucci EL, Fuchs CS, Ogino S. Comprehensive biostatistical analysis of CpG island methylator phenotype in colorectal cancer using a large population-based sample. PloS one. 2008;3(11):e3698. Epub 2008/11/13. doi: 10.1371/journal.pone.0003698. PubMed PMID: 19002263; PubMed Central PMCID: PMC2579485.
148. Taoufik E, Petit E, Divoux D, Tseveleki V, Mengozzi M, Roberts ML, Valable S, Ghezzi P, **Quackenbush J**, Brines M, Cerami A, Probert L. TNF receptor I sensitizes neurons to erythropoietin- and VEGF-mediated neuroprotection after ischemic and excitotoxic injury. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(16):6185-90. Epub 2008/04/17. doi: 10.1073/pnas.0801447105. PubMed PMID: 18413601; PubMed Central PMCID: PMC2299225.
149. Taylor CF, Field D, Sansone SA, Aerts J, Apweiler R, Ashburner M, Ball CA, Binz PA, Bogue M, Booth T, Brazma A, Brinkman RR, Michael Clark A, Deutsch EW, Fiehn O, Fostel J, Ghazal P, Gibson F, Gray T, Grimes G, Hancock JM, Hardy NW, Hermjakob H, Julian RK, Jr., Kane M, Kettner C, Kinsinger C, Kolker E, Kuiper M, Le Novere N, Leebens-Mack J, Lewis SE, Lord P, Mallon AM, Marthandan N, Masuya H, McNally R, Mehrle A, Morrison N, Orchard S, **Quackenbush J**, Reecy JM, Robertson DG, Rocca-Serra P, Rodriguez H, Rosenfelder H, Santoyo-Lopez J, Scheuermann RH, Schober D, Smith B, Snape J, Stoeckert CJ, Jr., Tipton K, Sterk P, Untergasser A, Vandesompele J, Wiemann S. Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. Nature biotechnology. 2008;26(8):889-96. Epub 2008/08/09. doi: 10.1038/nbt.1411. PubMed PMID: 18688244; PubMed Central PMCID: PMC2771753.
150. April C, Klotzle B, Royce T, Wickham-Garcia E, Boyaniwsky T, Izzo J, Cox D, Jones W, Rubio R, Holton K, Matulonis U, **Quackenbush J**, Fan JB. Whole-genome gene expression profiling of formalin-fixed, paraffin-embedded tissue samples. PloS one. 2009;4(12):e8162. Epub 2009/12/10. doi: 10.1371/journal.pone.0008162. PubMed PMID: 19997620; PubMed Central PMCID: PMC2780295.
151. Aryee MJ, Gutierrez-Pabello JA, Kramnik I, Maiti T, **Quackenbush J**. An improved empirical bayes approach to estimating differential gene expression in microarray time-course data: BETR (Bayesian Estimation of Temporal Regulation). BMC bioinformatics. 2009;10:409. Epub 2009/12/17. doi: 10.1186/1471-2105-10-409. PubMed PMID: 20003283; PubMed Central PMCID: PMC2801687.
152. Bateman A, **Quackenbush J**. Bioinformatics for next generation sequencing. Bioinformatics. 2009;25(4):429. Epub 2009/02/10. doi: 10.1093/bioinformatics/btp037. PubMed PMID: 19202193.

153. Bodedingheimer MJ, Chou JW, Corton JC, Fostel J, O'Lone R, Pine PS, **Quackenbush J**, Thompson KL, Wolfinger RD. Variance due to Smooth Bias in Rat Liver and Kidney Baseline Gene Expression in a Large Multi-laboratory Data Set. In: Scherer A, editor. Batch Effects and Noise in Microarray Experiments: Sources and Solutions John Wiley and Sons, Ltd; 2009. p. 87-99.
154. Colak D, Kaya N, Al-Zahrani J, Al Bakheet A, Muiya P, Andres E, **Quackenbush J**, Dzimiri N. Left ventricular global transcriptional profiling in human end-stage dilated cardiomyopathy. Genomics. 2009;94(1):20-31. Epub 2009/04/01. doi: 10.1016/j.ygeno.2009.03.003. PubMed PMID: 19332114; PubMed Central PMCID: PMC4152850.
155. Consortium F, Suzuki H, Forrest AR, van Nimwegen E, Daub CO, Balwierz PJ, Irvine KM, Lassmann T, Ravasi T, Hasegawa Y, de Hoon MJ, Katayama S, Schroder K, Carninci P, Tomaru Y, Kanamori-Katayama M, Kubosaki A, Akalin A, Ando Y, Arner E, Asada M, Asahara H, Bailey T, Bajic VB, Bauer D, Beckhouse AG, Bertin N, Bjorkegren J, Brombacher F, Bulger E, Chalk AM, Chiba J, Cloonan N, Dawe A, Dostie J, Engstrom PG, Essack M, Faulkner GJ, Fink JL, Fredman D, Fujimori K, Furuno M, Gojobori T, Gough J, Grimmond SM, Gustafsson M, Hashimoto M, Hashimoto T, Hatakeyama M, Heinzel S, Hide W, Hofmann O, Hornquist M, Huminiecki L, Ikeo K, Imamoto N, Inoue S, Inoue Y, Ishihara R, Iwayanagi T, Jacobsen A, Kaur M, Kawaji H, Kerr MC, Kimura R, Kimura S, Kimura Y, Kitano H, Koga H, Kojima T, Kondo S, Konno T, Krogh A, Kruger A, Kumar A, Lenhard B, Lennartsson A, Lindow M, Lizio M, Macpherson C, Maeda N, Maher CA, Maqungo M, Mar J, Matigian NA, Matsuda H, Mattick JS, Meier S, Miyamoto S, Miyamoto-Sato E, Nakabayashi K, Nakachi Y, Nakano M, Nygaard S, Okayama T, Okazaki Y, Okuda-Yabukami H, Orlando V, Otomo J, Pachkov M, Petrovsky N, Plessy C, **Quackenbush J**, Radovanovic A, Rehli M, Saito R, Sandelin A, Schmeier S, Schonbach C, Schwartz AS, Semple CA, Sera M, Severin J, Shirahige K, Simons C, St Laurent G, Suzuki M, Suzuki T, Sweet MJ, Taft RJ, Takeda S, Takenaka Y, Tan K, Taylor MS, Teasdale RD, Tegner J, Teichmann S, Valen E, Wahlestedt C, Waki K, Waterhouse A, Wells CA, Winther O, Wu L, Yamaguchi K, Yanagawa H, Yasuda J, Zavolan M, Hume DA, Riken Omics Science C, Arakawa T, Fukuda S, Imamura K, Kai C, Kaiho A, Kawashima T, Kawazu C, Kitazume Y, Kojima M, Miura H, Murakami K, Murata M, Ninomiya N, Nishiyori H, Noma S, Ogawa C, Sano T, Simon C, Tagami M, Takahashi Y, Kawai J, Hayashizaki Y. The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. Nature genetics. 2009;41(5):553-62. Epub 2009/04/21. doi: 10.1038/ng.375. PubMed PMID: 19377474.
156. Culhane AC, **Quackenbush J**. Confounding effects in "A six-gene signature predicting breast cancer lung metastasis". Cancer research. 2009;69(18):7480-5. Epub 2009/09/03. doi: 10.1158/0008-5472.CAN-08-3350. PubMed PMID: 19723662; PubMed Central PMCID: PMC3128918.
157. Dutta B, Kanani H, **Quackenbush J**, Klapa MI. Time-series integrated "omic" analyses to elucidate short-term stress-induced responses in plant liquid cultures. Biotechnology and bioengineering. 2009;102(1):264-79. Epub 2008/10/30. doi: 10.1002/bit.22036. PubMed PMID: 18958862; PubMed Central PMCID: PMC4134133.

158. Fedorov VB, Goropashnaya AV, Toien O, Stewart NC, Gracey AY, Chang C, Qin S, Pertea G, **Quackenbush J**, Showe LC, Showe MK, Boyer BB, Barnes BM. Elevated expression of protein biosynthesis genes in liver and muscle of hibernating black bears (Ursus americanus). Physiological genomics. 2009;37(2):108-18. Epub 2009/02/26. doi: 10.1152/physiolgenomics.90398.2008. PubMed PMID: 19240299; PubMed Central PMCID: PMC3774579.
159. Gorgun G, Ramsay AG, Holderried TA, Zahrieh D, Le Dieu R, Liu F, **Quackenbush J**, Croce CM, Gribben JG. E(mu)-TCL1 mice represent a model for immunotherapeutic reversal of chronic lymphocytic leukemia-induced T-cell dysfunction. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(15):6250-5. Epub 2009/04/01. doi: 10.1073/pnas.0901166106. PubMed PMID: 19332800; PubMed Central PMCID: PMC2669383.
160. Hellner K, Mar J, Fang F, **Quackenbush J**, Munger K. HPV16 E7 oncogene expression in normal human epithelial cells causes molecular changes indicative of an epithelial to mesenchymal transition. Virology. 2009;391(1):57-63. Epub 2009/06/26. doi: 10.1016/j.virol.2009.05.036. PubMed PMID: 19552933; PubMed Central PMCID: PMC2728214.
161. Mar JC, Kimura Y, Schroder K, Irvine KM, Hayashizaki Y, Suzuki H, Hume D, **Quackenbush J**. Data-driven normalization strategies for high-throughput quantitative RT-PCR. BMC bioinformatics. 2009;10:110. Epub 2009/04/21. doi: 10.1186/1471-2105-10-110. PubMed PMID: 19374774; PubMed Central PMCID: PMC2680405.
162. Mar JC, **Quackenbush J**. Decomposition of gene expression state space trajectories. PLoS computational biology. 2009;5(12):e1000626. Epub 2009/12/31. doi: 10.1371/journal.pcbi.1000626. PubMed PMID: 20041215; PubMed Central PMCID: PMC2791157.
163. **Quackenbush J**. Data reporting standards: making the things we use better. Genome medicine. 2009;1(11):111. Epub 2009/11/27. doi: 10.1186/gm111. PubMed PMID: 19939294; PubMed Central PMCID: PMC2808746.
164. Santagata S, Maire CL, Idbaih A, Geffers L, Correll M, Holton K, **Quackenbush J**, Ligon KL. CRX is a diagnostic marker of retinal and pineal lineage tumors. PloS one. 2009;4(11):e7932. Epub 2009/11/26. doi: 10.1371/journal.pone.0007932. PubMed PMID: 19936203; PubMed Central PMCID: PMC2775954.
165. Speers C, Tsimelzon A, Sexton K, Herrick AM, Gutierrez C, Culhane A, **Quackenbush J**, Hilsenbeck S, Chang J, Brown P. Identification of novel kinase targets for the treatment of estrogen receptor-negative breast cancer. Clinical cancer research : an official journal of the American Association for Cancer Research. 2009;15(20):6327-40. Epub 2009/10/08. doi: 10.1158/1078-0432.CCR-09-1107. PubMed PMID: 19808870; PubMed Central PMCID: PMC2763053.
166. Antonescu C, Antonescu V, Sultana R, **Quackenbush J**. Using the DFCI gene index databases for biological discovery. Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al]. 2010;Chapter 1:Unit1 6 1-36. Epub 2010/03/06. doi: 10.1002/0471250953.bi0106s29. PubMed PMID: 20205187; PubMed Central PMCID: PMC4205239.
167. Chen DT, Nasir A, Culhane A, Venkataramu C, Fulp W, Rubio R, Wang T, Agrawal D, McCarthy SM, Gruidl M, Bloom G, Anderson T, White J, **Quackenbush J**, Yeatman T. Proliferative genes dominate malignancy-risk gene

signature in histologically-normal breast tissue. Breast cancer research and treatment. 2010;119(2):335-46. Epub 2009/03/07. doi: 10.1007/s10549-009-0344-y. PubMed PMID: 19266279; PubMed Central PMCID: PMC2796276.

168. Chittenden TW, Pak J, Rubio R, Cheng H, Holton K, Prendergast N, Glinskii V, Cai Y, Culhane A, Bentink S, Schwede M, Mar JC, Howe EA, Aryee M, Sultana R, Lanahan AA, Taylor JM, Holmes C, Hahn WC, Zhao JJ, Iglehart JD, **Quackenbush J**. Therapeutic implications of GIPC1 silencing in cancer. PloS one. 2010;5(12):e15581. Epub 2011/01/07. doi: 10.1371/journal.pone.0015581. PubMed PMID: 21209904; PubMed Central PMCID: PMC3012716.
169. Colak D, Chishti MA, Al-Bakheet AB, Al-Qahtani A, Shoukri MM, Goyns MH, Ozand PT, **Quackenbush J**, Park BH, Kaya N. Integrative and comparative genomics analysis of early hepatocellular carcinoma differentiated from liver regeneration in young and old. Molecular cancer. 2010;9:146. Epub 2010/06/15. doi: 10.1186/1476-4598-9-146. PubMed PMID: 20540791; PubMed Central PMCID: PMC2898705.
170. Culhane AC, Schwarzl T, Sultana R, Picard KC, Picard SC, Lu TH, Franklin KR, French SJ, Papenhausen G, Correll M, **Quackenbush J**. GeneSigDB--a curated database of gene expression signatures. Nucleic acids research. 2010;38(Database issue):D716-25. Epub 2009/11/26. doi: 10.1093/nar/gkp1015. PubMed PMID: 19934259; PubMed Central PMCID: PMC2808880.
171. Tanaka N, Huttenhower C, Nosho K, Baba Y, Shima K, **Quackenbush J**, Haigis KM, Giovannucci E, Fuchs CS, Ogino S. Novel application of structural equation modeling to correlation structure analysis of CpG island methylation in colorectal cancer. The American journal of pathology. 2010;177(6):2731-40. Epub 2010/11/03. doi: 10.2353/ajpath.2010.100361. PubMed PMID: 21037082; PubMed Central PMCID: PMC2993288.
172. Tseveleki V, Rubio R, Vamvakas SS, White J, Taoufik E, Petit E, **Quackenbush J**, Probert L. Comparative gene expression analysis in mouse models for multiple sclerosis, Alzheimer's disease and stroke for identifying commonly regulated and disease-specific gene changes. Genomics. 2010;96(2):82-91. Epub 2010/05/04. doi: 10.1016/j.ygeno.2010.04.004. PubMed PMID: 20435134; PubMed Central PMCID: PMC4205236.
173. Bontempi G, Haibe-Kains B, Desmedt C, Sotiriou C, **Quackenbush J**. Multiple-input multiple-output causal strategies for gene selection. BMC bioinformatics. 2011;12:458. Epub 2011/11/29. doi: 10.1186/1471-2105-12-458. PubMed PMID: 22118187; PubMed Central PMCID: PMC3323860.
174. Cannon EK, Birkett SM, Braun BL, Kodavali S, Jennewein DM, Yilmaz A, Antonescu V, Antonescu C, Harper LC, Gardiner JM, Schaeffer ML, Campbell DA, Andorf CM, Andorf D, Lisch D, Koch KE, McCarty DR, **Quackenbush J**, Grotewold E, Lushbough CM, Sen TZ, Lawrence CJ. POPcorn: An Online Resource Providing Access to Distributed and Diverse Maize Project Data. International journal of plant genomics. 2011;2011:923035. Epub 2012/01/19. doi: 10.1155/2011/923035. PubMed PMID: 22253616; PubMed Central PMCID: PMC3255282.
175. Chervitz SA, Deutsch EW, Field D, Parkinson H, **Quackenbush J**, Rocca-Serra P, Sansone SA, Stoeckert CJ, Jr., Taylor CF, Taylor R, Ball CA. Data standards for Omics data: the basis of data sharing and reuse. Methods Mol Biol. 2011;719:31-

69. Epub 2011/03/04. doi: 10.1007/978-1-61779-027-0_2. PubMed PMID: 21370078; PubMed Central PMCID: PMC4152841.
176. Gewurz BE, Mar JC, Padi M, Zhao B, Shinners NP, Takasaki K, Bedoya E, Zou JY, Cahir-McFarland E, **Quackenbush J**, Kieff E. Canonical NF-kappaB activation is essential for Epstein-Barr virus latent membrane protein 1 TES2/CTAR2 gene regulation. Journal of virology. 2011;85(13):6764-73. Epub 2011/05/06. doi: 10.1128/JVI.00422-11. PubMed PMID: 21543491; PubMed Central PMCID: PMC3126495.
177. Howe EA, Sinha R, Schlauch D, **Quackenbush J**. RNA-Seq analysis in MeV. Bioinformatics. 2011;27(22):3209-10. Epub 2011/10/07. doi: 10.1093/bioinformatics/btr490. PubMed PMID: 21976420; PubMed Central PMCID: PMC3208390.
178. Jirawatnotai S, Hu Y, Michowski W, Elias JE, Becks L, Bienvenu F, Zagozdzon A, Goswami T, Wang YE, Clark AB, Kunkel TA, van Harn T, Xia B, Correll M, **Quackenbush J**, Livingston DM, Gygi SP, Sicinski P. A function for cyclin D1 in DNA repair uncovered by protein interactome analyses in human cancers. Nature. 2011;474(7350):230-4. Epub 2011/06/10. doi: 10.1038/nature10155. PubMed PMID: 21654808; PubMed Central PMCID: PMC3134411.
179. Konstantinopoulos PA, Cannistra SA, Fountzilas H, Culhane A, Pillay K, Rueda B, Cramer D, Seiden M, Birrer M, Coukos G, Zhang L, **Quackenbush J**, Spentzos D. Integrated analysis of multiple microarray datasets identifies a reproducible survival predictor in ovarian cancer. PloS one. 2011;6(3):e18202. Epub 2011/04/12. doi: 10.1371/journal.pone.0018202. PubMed PMID: 21479231; PubMed Central PMCID: PMC3066217.
180. Liu F, White JA, Antonescu C, Gusenleitner D, **Quackenbush J**. GCOD - GeneChip Oncology Database. BMC bioinformatics. 2011;12:46. Epub 2011/02/05. doi: 10.1186/1471-2105-12-46. PubMed PMID: 21291543; PubMed Central PMCID: PMC3045303.
181. Mar JC, Matigian NA, Mackay-Sim A, Mellick GD, Sue CM, Silburn PA, McGrath JJ, **Quackenbush J**, Wells CA. Variance of gene expression identifies altered network constraints in neurological disease. PLoS genetics. 2011;7(8):e1002207. Epub 2011/08/20. doi: 10.1371/journal.pgen.1002207. PubMed PMID: 21852951; PubMed Central PMCID: PMC3154954.
182. Mar JC, Matigian NA, **Quackenbush J**, Wells CA. attract: A method for identifying core pathways that define cellular phenotypes. PloS one. 2011;6(10):e25445. Epub 2011/10/25. doi: 10.1371/journal.pone.0025445. PubMed PMID: 22022396; PubMed Central PMCID: PMC3194807.
183. Mar JC, Wells CA, **Quackenbush J**. Defining an informativeness metric for clustering gene expression data. Bioinformatics. 2011;27(8):1094-100. Epub 2011/02/19. doi: 10.1093/bioinformatics/btr074. PubMed PMID: 21330289; PubMed Central PMCID: PMC3072547.
184. Marko NF, **Quackenbush J**, Weil RJ. Why is there a lack of consensus on molecular subgroups of glioblastoma? Understanding the nature of biological and statistical variability in glioblastoma expression data. PloS one. 2011;6(7):e20826. Epub 2011/08/11. doi: 10.1371/journal.pone.0020826. PubMed PMID: 21829433; PubMed Central PMCID: PMC3145641.
185. **Quackenbush J**. The Human Genome: Book of Essential Knowledge (Curiosity Guides): Imagine! Publishing, An Imprint of Charlesbridge; 2011. 176 p.

186. Sathirapongsasuti JF, Lee H, Horst BA, Brunner G, Cochran AJ, Binder S, **Quackenbush J**, Nelson SF. Exome sequencing-based copy-number variation and loss of heterozygosity detection: ExomeCNV. Bioinformatics. 2011;27(19):2648-54. Epub 2011/08/11. doi: 10.1093/bioinformatics/btr462. PubMed PMID: 21828086; PubMed Central PMCID: PMC3179661.
187. Schroder MS, Culhane AC, **Quackenbush J**, Haibe-Kains B. survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. Bioinformatics. 2011;27(22):3206-8. Epub 2011/09/10. doi: 10.1093/bioinformatics/btr511. PubMed PMID: 21903630; PubMed Central PMCID: PMC3208391.
188. Zhao B, Mar JC, Maruo S, Lee S, Gewurz BE, Johannsen E, Holton K, Rubio R, Takada K, **Quackenbush J**, Kieff E. Epstein-Barr virus nuclear antigen 3C regulated genes in lymphoblastoid cell lines. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(1):337-42. Epub 2010/12/22. doi: 10.1073/pnas.1017419108. PubMed PMID: 21173222; PubMed Central PMCID: PMC3017193.
189. Zhao B, Zou J, Wang H, Johannsen E, Peng CW, **Quackenbush J**, Mar JC, Morton CC, Freedman ML, Blacklow SC, Aster JC, Bernstein BE, Kieff E. Epstein-Barr virus exploits intrinsic B-lymphocyte transcription programs to achieve immortal cell growth. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(36):14902-7. Epub 2011/07/13. doi: 10.1073/pnas.1108892108. PubMed PMID: 21746931; PubMed Central PMCID: PMC3169132.
190. Bentink S, Haibe-Kains B, Risch T, Fan JB, Hirsch MS, Holton K, Rubio R, April C, Chen J, Wickham-Garcia E, Liu J, Culhane A, Drapkin R, **Quackenbush J**, Matulonis UA. Angiogenic mRNA and microRNA gene expression signature predicts a novel subtype of serous ovarian cancer. PloS one. 2012;7(2):e30269. Epub 2012/02/22. doi: 10.1371/journal.pone.0030269. PubMed PMID: 22348002; PubMed Central PMCID: PMC3278409.
191. Chittenden TW, Howe EA, Taylor JM, Mar JC, Aryee MJ, Gomez H, Sultana R, Braisted J, Nair SJ, **Quackenbush J**, Holmes C. nEASE: a method for gene ontology subclassification of high-throughput gene expression data. Bioinformatics. 2012;28(5):726-8. Epub 2012/01/17. doi: 10.1093/bioinformatics/bts011. PubMed PMID: 22247278.
192. Culhane AC, Schroder MS, Sultana R, Picard SC, Martinelli EN, Kelly C, Haibe-Kains B, Kapushesky M, St Pierre AA, Flahive W, Picard KC, Gusenleitner D, Papenhausen G, O'Connor N, Correll M, **Quackenbush J**. GeneSigDB: a manually curated database and resource for analysis of gene expression signatures. Nucleic acids research. 2012;40(Database issue):D1060-6. Epub 2011/11/24. doi: 10.1093/nar/gkr901. PubMed PMID: 22110038; PubMed Central PMCID: PMC3245038.
193. Desmedt C, Majjaj S, Kheddoumi N, Singhal SK, Haibe-Kains B, El Ouriaghli F, Chaboteaux C, Michiels S, Lallemand F, Journe F, Duvillier H, Loi S, **Quackenbush J**, Dekoninck S, Blanpain C, Lagneaux L, Houhou N, Delorenzi M, Larsimont D, Piccart M, Sotiriou C. Characterization and clinical evaluation of CD10+ stroma cells in the breast cancer microenvironment. Clinical cancer research : an official journal of the American Association for Cancer Research.

2012;18(4):1004-14. Epub 2012/01/12. doi: 10.1158/1078-0432.CCR-11-0383. PubMed PMID: 22235100.

194. Fountzilas E, Kelly AD, Perez-Atayde AR, Goldsmith J, Konstantinopoulos PA, Francoeur N, Correll M, Rubio R, Hu L, Gebhardt MC, **Quackenbush J**, Spentzos D. A microRNA activity map of human mesenchymal tumors: connections to oncogenic pathways; an integrative transcriptomic study. BMC genomics. 2012;13:332. Epub 2012/07/25. doi: 10.1186/1471-2164-13-332. PubMed PMID: 22823907; PubMed Central PMCID: PMC3443663.
195. Gewurz BE, Towfic F, Mar JC, Shinners NP, Takasaki K, Zhao B, Cahir-McFarland ED, **Quackenbush J**, Xavier RJ, Kieff E. Genome-wide siRNA screen for mediators of NF-kappaB activation. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(7):2467-72. Epub 2012/02/07. doi: 10.1073/pnas.1120542109. PubMed PMID: 22308454; PubMed Central PMCID: PMC3289371.
196. Gulbahce N, Yan H, Dricot A, Padi M, Byrdsong D, Franchi R, Lee DS, Rozenblatt-Rosen O, Mar JC, Calderwood MA, Baldwin A, Zhao B, Santhanam B, Braun P, Simonis N, Huh KW, Hellner K, Grace M, Chen A, Rubio R, Marto JA, Christakis NA, Kieff E, Roth FP, Roecklein-Canfield J, Decaprio JA, Cusick ME, **Quackenbush J**, Hill DE, Munger K, Vidal M, Barabasi AL. Viral perturbations of host networks reflect disease etiology. PLoS computational biology. 2012;8(6):e1002531. Epub 2012/07/05. doi: 10.1371/journal.pcbi.1002531. PubMed PMID: 22761553; PubMed Central PMCID: PMC3386155.
197. Gusenleitner D, Howe EA, Bentink S, **Quackenbush J**, Culhane AC. iBBiG: iterative binary bi-clustering of gene sets. Bioinformatics. 2012;28(19):2484-92. Epub 2012/07/14. doi: 10.1093/bioinformatics/bts438. PubMed PMID: 22789589; PubMed Central PMCID: PMC3463116.
198. Haibe-Kains B, Desmedt C, Loi S, Culhane AC, Bontempi G, **Quackenbush J**, Sotiriou C. A three-gene model to robustly identify breast cancer molecular subtypes. Journal of the National Cancer Institute. 2012;104(4):311-25. Epub 2012/01/21. doi: 10.1093/jnci/djr545. PubMed PMID: 22262870; PubMed Central PMCID: PMC3283537.
199. Haibe-Kains B, Olsen C, Djebbari A, Bontempi G, Correll M, Bouton C, **Quackenbush J**. Predictive networks: a flexible, open source, web application for integration and analysis of human gene networks. Nucleic acids research. 2012;40(Database issue):D866-75. Epub 2011/11/19. doi: 10.1093/nar/gkr1050. PubMed PMID: 22096235; PubMed Central PMCID: PMC3245161.
200. Milbury CA, Correll M, **Quackenbush J**, Rubio R, Makrigiorgos GM. COLD-PCR enrichment of rare cancer mutations prior to targeted amplicon resequencing. Clinical chemistry. 2012;58(3):580-9. Epub 2011/12/24. doi: 10.1373/clinchem.2011.176198. PubMed PMID: 22194627; PubMed Central PMCID: PMC3418918.
201. Rozenblatt-Rosen O, Deo RC, Padi M, Adelmant G, Calderwood MA, Rolland T, Grace M, Dricot A, Askenazi M, Tavares M, Pevzner SJ, Abderazzaq F, Byrdsong D, Carvunis AR, Chen AA, Cheng J, Correll M, Duarte M, Fan C, Feltkamp MC, Ficarro SB, Franchi R, Garg BK, Gulbahce N, Hao T, Holthaus AM, James R, Korkhin A, Litovchick L, Mar JC, Pak TR, Rabello S, Rubio R, Shen Y, Singh S, Spangle JM, Tasan M, Wanamaker S, Webber JT, Roecklein-Canfield J, Johannsen E, Barabasi AL, Beroukhim R, Kieff E, Cusick ME, Hill DE, Munger K, Marto JA,

Quackenbush J, Roth FP, DeCaprio JA, Vidal M. Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. Nature. 2012;487(7408):491-5. Epub 2012/07/20. doi: 10.1038/nature11288. PubMed PMID: 22810586; PubMed Central PMCID: PMC3408847.

202. Skiadas CC, Duan S, Correll M, Rubio R, Karaca N, Ginsburg ES, **Quackenbush J**, Racowsky C. Ovarian reserve status in young women is associated with altered gene expression in membrana granulosa cells. Molecular human reproduction. 2012;18(7):362-71. Epub 2012/02/23. doi: 10.1093/molehr/gas008. PubMed PMID: 22355044; PubMed Central PMCID: PMC3378309.
203. Vidal M, Chan DW, Gerstein M, Mann M, Omenn GS, Tagle D, Sechi S, Apweiler R, Bader J, Barnes S, Bettauer R, Chan MM, Coon J, Dolinski K, Dowell BL, Edmonds CG, Gaudet S, Gerszten R, Kagan J, Mazurchuk R, McShane L, Moritz RL, Palmer-Toy D, Pandey A, Polanski M, **Quackenbush J**, Sheeley D, Skates S, Srivastava S, Wang W, Weng Z, White F. The human proteome - a scientific opportunity for transforming diagnostics, therapeutics, and healthcare. Clinical proteomics. 2012;9(1):6. Epub 2012/05/16. doi: 10.1186/1559-0275-9-6. PubMed PMID: 22583803; PubMed Central PMCID: PMC3388576.
204. Waldron L, Ogino S, Hoshida Y, Shima K, McCart Reed AE, Simpson PT, Baba Y, Nosho K, Segata N, Vargas AC, Cummings MC, Lakhani SR, Kirkner GJ, Giovannucci E, **Quackenbush J**, Golub TR, Fuchs CS, Parmigiani G, Huttenhower C. Expression profiling of archival tumors for long-term health studies. Clinical cancer research : an official journal of the American Association for Cancer Research. 2012;18(22):6136-46. Epub 2012/11/09. doi: 10.1158/1078-0432.CCR-12-1915. PubMed PMID: 23136189; PubMed Central PMCID: PMC3500412.
205. Wang ZC, Birkbak NJ, Culhane AC, Drapkin R, Fatima A, Tian R, Schwede M, Alsop K, Daniels KE, Piao H, Liu J, Etemadmoghadam D, Miron A, Salvesen HB, Mitchell G, DeFazio A, **Quackenbush J**, Berkowitz RS, Iglehart JD, Bowtell DD, Australian Ovarian Cancer Study G, Matulonis UA. Profiles of genomic instability in high-grade serous ovarian cancer predict treatment outcome. Clinical cancer research : an official journal of the American Association for Cancer Research. 2012;18(20):5806-15. Epub 2012/08/23. doi: 10.1158/1078-0432.CCR-12-0857. PubMed PMID: 22912389; PubMed Central PMCID: PMC4205235.
206. Beck AH, Knoblauch NW, Hefti MM, Kaplan J, Schnitt SJ, Culhane AC, Schroeder MS, Risch T, **Quackenbush J**, Haibe-Kains B. Significance analysis of prognostic signatures. PLoS computational biology. 2013;9(1):e1002875. Epub 2013/02/01. doi: 10.1371/journal.pcbi.1002875. PubMed PMID: 23365551; PubMed Central PMCID: PMC3554539.
207. Correll M, Johnson CK, Ferrari G, Brizzio M, Mak AW, **Quackenbush J**, Shaw RE, Zapolanski A, Grau JB. Mutational analysis clopidogrel resistance and platelet function in patients scheduled for coronary artery bypass grafting. Genomics. 2013;101(6):313-7. Epub 2013/03/07. doi: 10.1016/j.ygeno.2013.01.005. PubMed PMID: 23462555; PubMed Central PMCID: PMC4149181.
208. De Rienzo A, Richards WG, Yeap BY, Coleman MH, Sugarbaker PE, Chirieac LR, Wang YE, **Quackenbush J**, Jensen RV, Bueno R. Sequential binary gene ratio tests define a novel molecular diagnostic strategy for malignant pleural mesothelioma. Clinical cancer research : an official journal of the American Association for Cancer Research. 2013;19(9):2493-502. Epub 2013/03/16. doi: 10.1158/1078-

0432.CCR-12-2117. PubMed PMID: 23493352; PubMed Central PMCID: PMC3644001.

209. Glass K, Huttenhower C, **Quackenbush J**, Yuan GC. Passing messages between biological networks to refine predicted interactions. PloS one. 2013;8(5):e64832. Epub 2013/06/07. doi: 10.1371/journal.pone.0064832. PubMed PMID: 23741402; PubMed Central PMCID: PMC3669401.
210. Haibe-Kains B, El-Hachem N, Birkbak NJ, Jin AC, Beck AH, Aerts HJ, **Quackenbush J**. Inconsistency in large pharmacogenomic studies. Nature. 2013;504(7480):389-93. Epub 2013/11/29. doi: 10.1038/nature12831. PubMed PMID: 24284626; PubMed Central PMCID: PMC4237165.
211. Kelly AD, Haibe-Kains B, Janeway KA, Hill KE, Howe E, Goldsmith J, Kurek K, Perez-Atayde AR, Francoeur N, Fan JB, April C, Schneider H, Gebhardt MC, Culhane A, **Quackenbush J**, Spentzos D. MicroRNA paraffin-based studies in osteosarcoma reveal reproducible independent prognostic profiles at 14q32. Genome medicine. 2013;5(1):2. Epub 2013/01/24. doi: 10.1186/gm406. PubMed PMID: 23339462; PubMed Central PMCID: PMC3706900.
212. Kelly AD, Hill KE, Correll M, Hu L, Wang YE, Rubio R, Duan S, **Quackenbush J**, Spentzos D. Next-generation sequencing and microarray-based interrogation of microRNAs from formalin-fixed, paraffin-embedded tissue: preliminary assessment of cross-platform concordance. Genomics. 2013;102(1):8-14. Epub 2013/04/09. doi: 10.1016/j.ygeno.2013.03.008. PubMed PMID: 23562991; PubMed Central PMCID: PMC4116671.
213. Lam HC, Cloonan SM, Bhashyam AR, Haspel JA, Singh A, Sathirapongsasuti JF, Cervo M, Yao H, Chung AL, Mizumura K, An CH, Shan B, Franks JM, Haley KJ, Owen CA, Tesfaigzi Y, Washko GR, **Quackenbush J**, Silverman EK, Rahman I, Kim HP, Mahmood A, Biswal SS, Ryter SW, Choi AM. Histone deacetylase 6-mediated selective autophagy regulates COPD-associated cilia dysfunction. The Journal of clinical investigation. 2013;123(12):5212-30. Epub 2013/11/10. doi: 10.1172/JCI69636. PubMed PMID: 24200693; PubMed Central PMCID: PMC3859407.
214. Larson JL, Huttenhower C, **Quackenbush J**, Yuan GC. A tiered hidden Markov model characterizes multi-scale chromatin states. Genomics. 2013;102(1):1-7. Epub 2013/04/11. doi: 10.1016/j.ygeno.2013.03.009. PubMed PMID: 23570996; PubMed Central PMCID: PMC3676702.
215. Merritt MA, Bentink S, Schwede M, Iwanicki MP, **Quackenbush J**, Woo T, Agoston ES, Reinhardt F, Crum CP, Berkowitz RS, Mok SC, Witt AE, Jones MA, Wang B, Ince TA. Gene expression signature of normal cell-of-origin predicts ovarian tumor outcomes. PloS one. 2013;8(11):e80314. Epub 2013/12/05. doi: 10.1371/journal.pone.0080314. PubMed PMID: 24303006; PubMed Central PMCID: PMC3841174.
216. Papillon-Cavanagh S, De Jay N, Hachem N, Olsen C, Bontempi G, Aerts HJ, **Quackenbush J**, Haibe-Kains B. Comparison and validation of genomic predictors for anticancer drug sensitivity. Journal of the American Medical Informatics Association : JAMIA. 2013;20(4):597-602. Epub 2013/01/29. doi: 10.1136/amiajnl-2012-001442. PubMed PMID: 23355484; PubMed Central PMCID: PMC3721163.
217. Portal D, Zhou H, Zhao B, Kharchenko PV, Lowry E, Wong L, **Quackenbush J**, Holloway D, Jiang S, Lu Y, Kieff E. Epstein-Barr virus nuclear antigen leader protein localizes to promoters and enhancers with cell transcription factors and

EBNA2. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(46):18537-42. Epub 2013/10/30. doi: 10.1073/pnas.1317608110. PubMed PMID: 24167291; PubMed Central PMCID: PMC3832032.

218. Schroder MS, Gusenleitner D, **Quackenbush J**, Culhane AC, Haibe-Kains B. RamiGO: an R/Bioconductor package providing an AmiGO visualize interface. Bioinformatics. 2013;29(5):666-8. Epub 2013/01/09. doi: 10.1093/bioinformatics/bts708. PubMed PMID: 23297033; PubMed Central PMCID: PMC3582261.
219. Schwede M, Spentzos D, Bentink S, Hofmann O, Haibe-Kains B, Harrington D, **Quackenbush J**, Culhane AC. Stem cell-like gene expression in ovarian cancer predicts type II subtype and prognosis. PloS one. 2013;8(3):e57799. Epub 2013/03/29. doi: 10.1371/journal.pone.0057799. PubMed PMID: 23536770; PubMed Central PMCID: PMC3594231.
220. Zhou X, Qiu W, Sathirapongsasuti JF, Cho MH, Mancini JD, Lao T, Thibault DM, Litonjua AA, Bakke PS, Gulsvik A, Lomas DA, Beaty TH, Hersh CP, Anderson C, Geigenmuller U, Raby BA, Rennard SI, Perrella MA, Choi AM, **Quackenbush J**, Silverman EK. Gene expression analysis uncovers novel hedgehog interacting protein (HHIP) effects in human bronchial epithelial cells. Genomics. 2013;101(5):263-72. Epub 2013/03/06. doi: 10.1016/j.ygeno.2013.02.010. PubMed PMID: 23459001; PubMed Central PMCID: PMC3659826.
221. Adamia S, Bar-Natan M, Haibe-Kains B, Pilarski PM, Bach C, Pevzner S, Calimeri T, Avet-Loiseau H, Lode L, Verselis S, Fox EA, Galinsky I, Mathews S, Dagogo-Jack I, Wadleigh M, Steensma DP, Motyckova G, Deangelo DJ, **Quackenbush J**, Tenen DG, Stone RM, Griffin JD. NOTCH2 and FLT3 gene mis-splicings are common events in patients with acute myeloid leukemia (AML): new potential targets in AML. Blood. 2014;123(18):2816-25. Epub 2014/02/28. doi: 10.1182/blood-2013-02-481507. PubMed PMID: 24574459; PubMed Central PMCID: PMC4007608.
222. Adamia S, Haibe-Kains B, Pilarski PM, Bar-Natan M, Pevzner S, Avet-Loiseau H, Lode L, Verselis S, Fox EA, Burke J, Galinsky I, Dagogo-Jack I, Wadleigh M, Steensma DP, Motyckova G, Deangelo DJ, **Quackenbush J**, Stone R, Griffin JD. A genome-wide aberrant RNA splicing in patients with acute myeloid leukemia identifies novel potential disease markers and therapeutic targets. Clinical cancer research : an official journal of the American Association for Cancer Research. 2014;20(5):1135-45. Epub 2013/11/29. doi: 10.1158/1078-0432.CCR-13-0956. PubMed PMID: 24284058.
223. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Cavalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D, Hoebers F, Rietbergen MM, Leemans CR, Dekker A, **Quackenbush J**, Gillies RJ, Lambin P. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nature communications. 2014;5:4006. Epub 2014/06/04. doi: 10.1038/ncomms5006. PubMed PMID: 24892406; PubMed Central PMCID: PMC4059926.
224. Chu JH, Hersh CP, Castaldi PJ, Cho MH, Raby BA, Laird N, Bowler R, Rennard S, Loscalzo J, **Quackenbush J**, Silverman EK. Analyzing networks of phenotypes in complex diseases: methodology and applications in COPD. BMC systems biology. 2014;8:78. Epub 2014/06/27. doi: 10.1186/1752-0509-8-78. PubMed PMID: 24964944; PubMed Central PMCID: PMC4105829.

225. Ferrari G, **Quackenbush J**, Strobeck J, Hu L, Johnson CK, Mak A, Shaw RE, Sayles K, Brizzio ME, Zapolanski A, Grau JB. Comparative genome-wide transcriptional analysis of human left and right internal mammary arteries. Genomics. 2014;104(1):36-44. Epub 2014/05/27. doi: 10.1016/j.ygeno.2014.04.005. PubMed PMID: 24858532; PubMed Central PMCID: PMC4152843.
226. Glass K, **Quackenbush J**, Silverman EK, Celli B, Rennard SI, Yuan GC, DeMeo DL. Sexually-dimorphic targeting of functionally-related genes in COPD. BMC systems biology. 2014;8:118. Epub 2014/11/29. doi: 10.1186/s12918-014-0118-y. PubMed PMID: 25431000; PubMed Central PMCID: PMC4269917.
227. Hatzis C, Bedard PL, Birkbak NJ, Beck AH, Aerts HJ, Stem DF, Shi L, Clarke R, **Quackenbush J**, Haibe-Kains B. Enhancing reproducibility in cancer drug screening: how do we move forward? Cancer research. 2014;74(15):4016-23. Epub 2014/07/13. doi: 10.1158/0008-5472.CAN-14-0725. PubMed PMID: 25015668; PubMed Central PMCID: PMC4119520.
228. Jirawatnotai S, Sharma S, Michowski W, Suktitipat B, Geng Y, **Quackenbush J**, Elias JE, Gygi SP, Wang YE, Sicinski P. The cyclin D1-CDK4 oncogenic interactome enables identification of potential novel oncogenes and clinical prognosis. Cell Cycle. 2014;13(18):2889-900. Epub 2014/12/09. doi: 10.4161/15384101.2014.946850. PubMed PMID: 25486477.
229. Lee JW, Norden AD, Ligon KL, Golby AJ, Beroukhim R, **Quackenbush J**, Wells W, Oelschlager K, Maetzold D, Wen PY. Tumor associated seizures in glioblastomas are influenced by survival gene expression in a region-specific manner: a gene expression imaging study. Epilepsy research. 2014;108(5):843-52. Epub 2014/04/03. doi: 10.1016/j.eplepsyres.2014.02.021. PubMed PMID: 24690158; PubMed Central PMCID: PMC4076930.
230. Li WQ, Han J, Widlund HR, Correll M, Wang YE, **Quackenbush J**, Mihm MC, Canales AL, Wu S, Golub T, Hoshida Y, Hunter DJ, Murphy G, Kupper TS, Qureshi AA. CXCR4 pathway associated with family history of melanoma. Cancer causes & control : CCC. 2014;25(1):125-32. Epub 2013/10/26. doi: 10.1007/s10552-013-0315-9. PubMed PMID: 24158781; PubMed Central PMCID: PMC4100594.
231. Mason EA, Mar JC, Laslett AL, Pera MF, **Quackenbush J**, Wolvetang E, Wells CA. Gene expression variability as a unifying element of the pluripotency network. Stem cell reports. 2014;3(2):365-77. Epub 2014/09/26. doi: 10.1016/j.stemcr.2014.06.008. PubMed PMID: 25254348; PubMed Central PMCID: PMC4175554.
232. Olsen C, Bontempi G, Emmert-Streib F, **Quackenbush J**, Haibe-Kains B. Relevance of different prior knowledge sources for inferring gene interaction networks. Frontiers in genetics. 2014;5:177. Epub 2014/07/11. doi: 10.3389/fgene.2014.00177. PubMed PMID: 25009552; PubMed Central PMCID: PMC4067568.
233. Olsen C, Fleming K, Prendergast N, Rubio R, Emmert-Streib F, Bontempi G, Haibe-Kains B, **Quackenbush J**. Inference and validation of predictive gene networks from biomedical literature and gene expression data. Genomics. 2014;103(5-6):329-36. Epub 2014/04/03. doi: 10.1016/j.ygeno.2014.03.004. PubMed PMID: 24691108; PubMed Central PMCID: PMC4119824.

234. **Quackenbush J**. Learning to share. Scientific American. 2014;311(1):S22. Epub 2014/07/01. PubMed PMID: 24974706.
235. **Quackenbush J**. Perspective: Learning to share. Nature. 2014;509(7502):S68. Epub 2014/05/30. doi: 10.1038/509S68a. PubMed PMID: 24870827; PubMed Central PMCID: PMC4205237.
236. Sharron Lin X, Hu L, Sandy K, Correll M, **Quackenbush J**, Wu CL, Scott McDougal W. Differentiating progressive from nonprogressive T1 bladder cancer by gene expression profiling: applying RNA-sequencing analysis on archived specimens. Urologic oncology. 2014;32(3):327-36. Epub 2013/09/24. doi: 10.1016/j.urolonc.2013.06.014. PubMed PMID: 24055427.
237. Yang IV, Pedersen BS, Rabinovich E, Hennessy CE, Davidson EJ, Murphy E, Guardela BJ, Tedrow JR, Zhang Y, Singh MK, Correll M, Schwarz MI, Geraci M, Sciurba FC, **Quackenbush J**, Spira A, Kaminski N, Schwartz DA. Relationship of DNA methylation and gene expression in idiopathic pulmonary fibrosis. American journal of respiratory and critical care medicine. 2014;190(11):1263-72. Epub 2014/10/22. doi: 10.1164/rccm.201408-1452OC. PubMed PMID: 25333685; PubMed Central PMCID: PMC4315819.
238. Campbell JD, Liu G, Luo L, Xiao J, Gerrein J, Juan-Guardela B, Tedrow J, Alekseyev YO, Yang IV, Correll M, Geraci M, **Quackenbush J**, Sciurba F, Schwartz DA, Kaminski N, Johnson WE, Monti S, Spira A, Beane J, Lenburg ME. Assessment of microRNA differential expression and detection in multiplexed small RNA sequencing data. RNA. 2015;21(2):164-71. Epub 2014/12/19. doi: 10.1261/rna.046060.114. PubMed PMID: 25519487; PubMed Central PMCID: PMC4338344.
239. Castaldi PJ, Cho MH, Zhou X, Qiu W, McGeachie M, Celli B, Bakke P, Gulsvik A, Lomas DA, Crapo JD, Beaty TH, Rennard S, Harshfield B, Lange C, Singh D, Tal-Singer R, Riley JH, **Quackenbush J**, Raby BA, Carey VJ, Silverman EK, Hersh CP. Genetic control of gene expression at novel and established chronic obstructive pulmonary disease loci. Human molecular genetics. 2015;24(4):1200-10. Epub 2014/10/16. doi: 10.1093/hmg/ddu525. PubMed PMID: 25315895.
240. Glass K, **Quackenbush J**, Spentzos D, Haibe-Kains B, Yuan GC. A network model for angiogenesis in ovarian cancer. BMC bioinformatics. 2015;16:115. Epub 2015/04/19. doi: 10.1186/s12859-015-0551-y. PubMed PMID: 25888305; PubMed Central PMCID: PMC4408593.
241. Lao T, Glass K, Qiu W, Polverino F, Gupta K, Morrow J, Mancini JD, Vuong L, Perrella MA, Hersh CP, Owen CA, **Quackenbush J**, Yuan GC, Silverman EK, Zhou X. Haploinsufficiency of Hedgehog interacting protein causes increased emphysema induced by cigarette smoke through network rewiring. Genome medicine. 2015;7(1):12. Epub 2015/03/13. doi: 10.1186/s13073-015-0137-3. PubMed PMID: 25763110; PubMed Central PMCID: PMC4355149.
242. Olsen C, Fleming K, Prendergast N, Rubio R, Emmert-Streib F, Bontempi G, **Quackenbush J**, Haibe-Kains B. Using shRNA experiments to validate gene regulatory networks. Genomics data. 2015;4:123-6. doi: 10.1016/j.gdata.2015.03.011. PubMed PMID: 26484195; PubMed Central PMCID: PMC4535466.
243. Padi M, **Quackenbush J**. Integrating transcriptional and protein interaction networks to prioritize condition-specific master regulators. BMC systems biology.

2015;9:80. doi: 10.1186/s12918-015-0228-1. PubMed PMID: 26576632; PubMed Central PMCID: PMC4650867.

244. Berrios C, Padi M, Keibler MA, Park DE, Molla V, Cheng J, Lee SM, Stephanopoulos G, **Quackenbush J**, DeCaprio JA. Merkel Cell Polyomavirus Small T Antigen Promotes Pro-Glycolytic Metabolic Perturbations Required for Transformation. PLoS Pathog. 2016;12(11):e1006020. doi: 10.1371/journal.ppat.1006020. PubMed PMID: 27880818; PubMed Central PMCID: PMCPMC5120958.
245. Cloonan SM, Glass K, Laucho-Contreras ME, Bhashyam AR, Cervo M, Pabon MA, Konrad C, Polverino F, Siempos, II, Perez E, Mizumura K, Ghosh MC, Parameswaran H, Williams NC, Rooney KT, Chen ZH, Goldklang MP, Yuan GC, Moore SC, Demeo DL, Rouault TA, D'Armiento JM, Schon EA, Manfredi G, **Quackenbush J**, Mahmood A, Silverman EK, Owen CA, Choi AM. Mitochondrial iron chelation ameliorates cigarette smoke-induced bronchitis and emphysema in mice. Nature medicine. 2016;22(2):163-74. doi: 10.1038/nm.4021. PubMed PMID: 26752519; PubMed Central PMCID: PMCPMC4742374.
246. De Rienzo A, Archer MA, Yeap BY, Dao N, Sciaranghella D, Sideris AC, Zheng Y, Holman AG, Wang YE, Dal Cin PS, Fletcher JA, Rubio R, Croft L, **Quackenbush J**, Sugarbaker PE, Munir KJ, Battilana JR, Gustafson CE, Chirieac LR, Ching SM, Wong J, Tay LC, Rudd S, Hercus R, Sugarbaker DJ, Richards WG, Bueno R. Gender-Specific Molecular and Clinical Features Underlie Malignant Pleural Mesothelioma. Cancer research. 2016;76(2):319-28. doi: 10.1158/0008-5472.CAN-15-0751. PubMed PMID: 26554828; PubMed Central PMCID: PMCPMC4715909.
247. Delaney SK, Hultner ML, Jacob HJ, Ledbetter DH, McCarthy JJ, Ball M, Beckman KB, Belmont JW, Bloss CS, Christman MF, Cosgrove A, Damiani SA, Danis T, Delledonne M, Dougherty MJ, Dudley JT, Faucett WA, Friedman JR, Haase DH, Hays TS, Heilsberg S, Huber J, Kaminsky L, Ledbetter N, Lee WH, Levin E, Libiger O, Linderman M, Love RL, Magnus DC, Martland A, McClure SL, Megill SE, Messier H, Nussbaum RL, Palaniappan L, Patay BA, Popovich BW, **Quackenbush J**, Savant MJ, Su MM, Terry SF, Tucker S, Wong WT, Green RC. Toward clinical genomics in everyday medicine: perspectives and recommendations. Expert review of molecular diagnostics. 2016;16(5):521-32. doi: 10.1586/14737159.2016.1146593. PubMed PMID: 26810587; PubMed Central PMCID: PMCPMC4841021.
248. Elias KM, Emori MM, Westerling T, Long H, Budina-Kolomets A, Li F, MacDuffie E, Davis MR, Holman A, Lawney B, Freedman ML, **Quackenbush J**, Brown M, Drapkin R. Epigenetic remodeling regulates transcriptional changes between ovarian cancer and benign precursors. JCI Insight. 2016;1(13). doi: 10.1172/jci.insight.87988. PubMed PMID: 27617304; PubMed Central PMCID: PMCPMC5017158.
249. Kusko RL, Brothers JF, 2nd, Tedrow J, Pandit K, Huleihel L, Perdomo C, Liu G, Juan-Guardela B, Kass D, Zhang S, Lenburg M, Martinez F, **Quackenbush J**, Sciurba F, Limper A, Geraci M, Yang I, Schwartz DA, Beane J, Spira A, Kaminski N. Integrated Genomics Reveals Convergent Transcriptomic Networks Underlying Chronic Obstructive Pulmonary Disease and Idiopathic Pulmonary Fibrosis. American journal of respiratory and critical care medicine. 2016;194(8):948-60.

doi: 10.1164/rccm.201510-2026OC. PubMed PMID: 27104832; PubMed Central PMCID: PMCPMC5067817.

250. Malleshaiah M, Padi M, Rue P, **Quackenbush J**, Martinez-Arias A, Gunawardena J. Nac1 Coordinates a Sub-network of Pluripotency Factors to Regulate Embryonic Stem Cell Differentiation. Cell reports. 2016;14(5):1181-94. doi: 10.1016/j.celrep.2015.12.101. PubMed PMID: 26832399; PubMed Central PMCID: PMCPMC4749452.
251. Manimaran S, Selby HM, Okrah K, Ruberman C, Leek JT, **Quackenbush J**, Haibe-Kains B, Bravo HC, Johnson WE. BatchQC: interactive software for evaluating sample and batch effects in genomic data. Bioinformatics. 2016;32(24):3836-8. doi: 10.1093/bioinformatics/btw538. PubMed PMID: 27540268; PubMed Central PMCID: PMCPMC5167063.
252. Morrow JD, Cho MH, Hersh CP, Pinto-Plata V, Celli B, Marchetti N, Criner G, Bueno R, Washko G, Glass K, Choi AM, **Quackenbush J**, Silverman EK, DeMeo DL. DNA methylation profiling in human lung tissue identifies genes associated with COPD. Epigenetics. 2016:1-10. doi: 10.1080/15592294.2016.1226451. PubMed PMID: 27564456; PubMed Central PMCID: PMCPMC5094634.
253. Platig J, Castaldi PJ, DeMeo D, **Quackenbush J**. Bipartite Community Structure of eQTLs. PLoS computational biology. 2016;12(9):e1005033. doi: 10.1371/journal.pcbi.1005033. PubMed PMID: 27618581; PubMed Central PMCID: PMCPMC5019382.
254. Safikhani Z, El-Hachem N, Quevedo R, Smirnov P, Goldenberg A, Juul Birkbak N, Mason C, Hatzis C, Shi L, Aerts HJ, **Quackenbush J**, Haibe-Kains B. Assessment of pharmacogenomic agreement. F1000Res. 2016;5:825. doi: 10.12688/f1000research.8705.1. PubMed PMID: 27408686; PubMed Central PMCID: PMCPMC4926729.
255. Safikhani Z, El-Hachem N, Smirnov P, Freeman M, Goldenberg A, Birkbak NJ, Beck AH, Aerts HJ, **Quackenbush J**, Haibe-Kains B. Safikhani et al. reply. Nature. 2016;540(7631):E2-E4. doi: 10.1038/nature19839. PubMed PMID: 27905430.
256. Safikhani Z, El-Hachem N, Smirnov P, Freeman M, Goldenberg A, Birkbak NJ, Beck AH, Aerts HJ, **Quackenbush J**, Haibe-Kains B. Safikhani et al. reply. Nature. 2016;540(7631):E11-E2. doi: 10.1038/nature20581. PubMed PMID: 27905423.
257. Safikhani Z, El-Hachem N, Smirnov P, Freeman M, Goldenberg A, Birkbak NJ, Beck AH, Aerts HJ, **Quackenbush J**, Haibe-Kains B. Safikhani et al. reply. Nature. 2016;540(7631):E6-E8. doi: 10.1038/nature20172. PubMed PMID: 27905416.
258. van IDG, Glass K, **Quackenbush J**, Kuijjer ML. PyPanda: a Python package for gene regulatory network reconstruction. Bioinformatics. 2016;32(21):3363-5. doi: 10.1093/bioinformatics/btw422. PubMed PMID: 27402905; PubMed Central PMCID: PMCPMC5079480.
259. Vargas AJ, **Quackenbush J**, Glass K. Diet-induced weight loss leads to a switch in gene regulatory network control in the rectal mucosa. Genomics. 2016;108(3-4):126-33. doi: 10.1016/j.ygeno.2016.08.001. PubMed PMID: 27524493; PubMed Central PMCID: PMCPMC5121035.
260. Wu W, Parmar C, Grossmann P, **Quackenbush J**, Lambin P, Bussink J, Mak R, Aerts HJ. Exploratory Study to Identify Radiomics Classifiers for Lung Cancer Histology. Frontiers in oncology. 2016;6:71. doi: 10.3389/fonc.2016.00071. PubMed PMID: 27064691; PubMed Central PMCID: PMCPMC4811956.

261. Campbell PT, Rebbeck TR, Nishihara R, Beck AH, Begg CB, Bogdanov AA, Cao Y, Coleman HG, Freeman GJ, Heng YJ, Huttenhower C, Irizarry RA, Kip NS, Michor F, Nevo D, Peters U, Phipps AI, Poole EM, Qian ZR, **Quackenbush J**, Robins H, Rogan PK, Slattery ML, Smith-Warner SA, Song M, VanderWeele TJ, Xia D, Zabor EC, Zhang X, Wang M, Ogino S. Proceedings of the third international molecular pathological epidemiology (MPE) meeting. Cancer causes & control : CCC. 2017;28(2):167-76. doi: 10.1007/s10552-016-0845-z. PubMed PMID: 28097472; PubMed Central PMCID: PMC5303153; https://www.ncbi.nlm.nih.gov/pubmed/28097472.
262. Domenyuk V, Zhong Z, Stark A, Xiao N, O'Neill HA, Wei X, Wang J, Tinder TT, Tonapi S, Duncan J, Hornung T, Hunter A, Miglarese MR, Schorr J, Halbert DD, **Quackenbush J**, Poste G, Berry DA, Mayer G, Famulok M, Spetzler D. Plasma Exosome Profiling of Cancer Patients by a Next Generation Systems Biology Approach. Sci Rep. 2017;7:42741. doi: 10.1038/srep42741. PubMed PMID: 28218293; PubMed Central PMCID: PMC5316983; https://www.ncbi.nlm.nih.gov/pubmed/28218293.
263. Fagny M, Paulson JN, Kuijjer ML, Sonawane AR, Chen CY, Lopes-Ramos CM, Glass K, **Quackenbush J**, Platig J. Exploring regulation in tissues with eQTL networks. Proceedings of the National Academy of Sciences of the United States of America. 2017;114(37):E7841-E50. doi: 10.1073/pnas.1707375114. PubMed PMID: 28851834; PubMed Central PMCID: PMC5604022; https://www.ncbi.nlm.nih.gov/pubmed/28851834.
264. Hill KE, Kelly AD, Kuijjer ML, Barry W, Rattani A, Garbutt CC, Kissick H, Janeway K, Perez-Atayde A, Goldsmith J, Gebhardt MC, Arredouani MS, Cote G, Hornicek F, Choy E, Duan Z, **Quackenbush J**, Haibe-Kains B, Spentzos D. An imprinted non-coding genomic cluster at 14q32 defines clinically relevant molecular subtypes in osteosarcoma across multiple independent datasets. J Hematol Oncol. 2017;10(1):107. doi: 10.1186/s13045-017-0465-4. PubMed PMID: 28506242; PubMed Central PMCID: PMC5433149; https://www.ncbi.nlm.nih.gov/pubmed/28506242.
265. Kibbe W, Klemm J, **Quackenbush J**. Cancer Informatics: New Tools for a Data-Driven Age in Cancer Research. Cancer research. 2017;77(21):e1-e2. doi: 10.1158/0008-5472.CAN-17-2212. PubMed PMID: 29092926; https://www.ncbi.nlm.nih.gov/pubmed/29092926.
266. Lopes-Ramos CM, Paulson JN, Chen CY, Kuijjer ML, Fagny M, Platig J, Sonawane AR, DeMeo DL, **Quackenbush J**, Glass K. Regulatory network changes between cell lines and their tissues of origin. BMC genomics. 2017;18(1):723. doi: 10.1186/s12864-017-4111-x. PubMed PMID: 28899340; PubMed Central PMCID: PMC5596945; https://www.ncbi.nlm.nih.gov/pubmed/28899340.
267. Morrow JD, Zhou X, Lao T, Jiang Z, DeMeo DL, Cho MH, Qiu W, Cloonan S, Pinto-Plata V, Celli B, Marchetti N, Criner GJ, Bueno R, Washko GR, Glass K, **Quackenbush J**, Choi AM, Silverman EK, Hersh CP. Functional interactors of three genome-wide association study genes are differentially expressed in severe chronic obstructive pulmonary disease lung tissue. Sci Rep. 2017;7:44232. doi: 10.1038/srep44232. PubMed PMID: 28287180; PubMed Central PMCID: PMC5347019; https://www.ncbi.nlm.nih.gov/pubmed/28287180.
268. Nishihara R, Glass K, Mima K, Hamada T, Nowak JA, Qian ZR, Kraft P, Giovannucci EL, Fuchs CS, Chan AT, **Quackenbush J**, Ogino S, Onnela JP.

Biomarker correlation network in colorectal carcinoma by tumor anatomic location. BMC bioinformatics. 2017;18(1):304. doi: 10.1186/s12859-017-1718-5. PubMed PMID: 28623901; PubMed Central PMCID: PMC5474023; https://www.ncbi.nlm.nih.gov/pubmed/28623901.

269. Paulson JN, Chen CY, Lopes-Ramos CM, Kuijjer ML, Platig J, Sonawane AR, Fagny M, Glass K, **Quackenbush J**. Tissue-aware RNA-Seq processing and normalization for heterogeneous and sparse data. BMC bioinformatics. 2017;18(1):437. doi: 10.1186/s12859-017-1847-x. PubMed PMID: 28974199; PubMed Central PMCID: PMC5627434; https://www.ncbi.nlm.nih.gov/pubmed/28974199.
270. Qiu W, Guo F, Glass K, Yuan GC, **Quackenbush J**, Zhou X, Tantisira KG. Differential connectivity of gene regulatory networks distinguishes corticosteroid response in asthma. J Allergy Clin Immunol. 2017. doi: 10.1016/j.jaci.2017.05.052. PubMed PMID: 28736268; https://www.ncbi.nlm.nih.gov/pubmed/28736268.
271. Quiroz-Zarate A, Harshfield BJ, Hu R, Knoblauch N, Beck AH, Hankinson SE, Carey V, Tamimi RM, Hunter DJ, **Quackenbush J**, Hazra A. Expression Quantitative Trait loci (QTL) in tumor adjacent normal breast tissue and breast tumor tissue. PloS one. 2017;12(2):e0170181. doi: 10.1371/journal.pone.0170181. PubMed PMID: 28152060; PubMed Central PMCID: PMC5289428; https://www.ncbi.nlm.nih.gov/pubmed/28152060.
272. Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, Ye Z, Makrigiorgos M, Fennessy F, Mak RH, Gillies R, **Quackenbush J**, Aerts H. Somatic Mutations Drive Distinct Imaging Phenotypes in Lung Cancer. Cancer research. 2017;77(14):3922-30. doi: 10.1158/0008-5472.CAN-17-0122. PubMed PMID: 28566328; PubMed Central PMCID: PMC5528160; https://www.ncbi.nlm.nih.gov/pubmed/28566328.
273. Schlauch D, Paulson JN, Young A, Glass K, **Quackenbush J**. Estimating gene regulatory networks with pandaR. Bioinformatics. 2017;33(14):2232-4. doi: 10.1093/bioinformatics/btx139. PubMed PMID: 28334344; https://www.ncbi.nlm.nih.gov/pubmed/28334344.
274. Sonawane AR, Platig J, Fagny M, Chen CY, Paulson JN, Lopes-Ramos CM, DeMeo DL, **Quackenbush J**, Glass K, Kuijjer ML. Understanding Tissue-Specific Gene Regulation. Cell reports. 2017;21(4):1077-88. doi: 10.1016/j.celrep.2017.10.001. PubMed PMID: 29069589; https://www.ncbi.nlm.nih.gov/pubmed/29069589.
275. Wang YE, Kutnetsov L, Partensky A, Farid J, **Quackenbush J**. WebMeV: A Cloud Platform for Analyzing and Visualizing Cancer Genomic Data. Cancer research. 2017;77(21):e11-e4. doi: 10.1158/0008-5472.CAN-17-0802. PubMed PMID: 29092929; https://www.ncbi.nlm.nih.gov/pubmed/29092929.
276. Yuan GC, Cai L, Elowitz M, Enver T, Fan G, Guo G, Irizarry R, Kharchenko P, Kim J, Orkin S, **Quackenbush J**, Saadatpour A, Schroeder T, Shivdasani R, Tirosh I. Challenges and emerging directions in single-cell analysis. Genome biology. 2017;18(1):84. doi: 10.1186/s13059-017-1218-y. PubMed PMID: 28482897; PubMed Central PMCID: PMC5421338; https://www.ncbi.nlm.nih.gov/pubmed/28482897.
277. Barry JD, Fagny M, Paulson JN, Aerts H, Platig J, **Quackenbush J**. Histopathological Image QTL Discovery of Immune Infiltration Variants. iScience.

2018;5:80-9. doi: 10.1016/j.isci.2018.07.001. PubMed PMID: 30240647; PubMed Central PMCID: PMC6123851; https://www.ncbi.nlm.nih.gov/pubmed/30240647.

278. Carette X, Platig J, Young DC, Helmel M, Young AT, Wang Z, Potluri LP, Moody CS, Zeng J, Prisic S, Paulson JN, Muntel J, Madduri AVR, Velarde J, Mayfield JA, Locher C, Wang T, **Quackenbush J**, Rhee KY, Moody DB, Steen H, Husson RN. Multisystem Analysis of Mycobacterium tuberculosis Reveals Kinase-Dependent Remodeling of the Pathogen-Environment Interface. MBio. 2018;9(2). doi: 10.1128/mBio.02333-17. PubMed PMID: 29511081; PubMed Central PMCID: PMC5845002; https://www.ncbi.nlm.nih.gov/pubmed/29511081.
279. Domenyuk V, Gatalica Z, Santhanam R, Wei X, Stark A, Kennedy P, Toussaint B, Levenberg S, Wang J, Xiao N, Greil R, Rinnerthaler G, Gampenrieder SP, Heimberger AB, Berry DA, Barker A, **Quackenbush J**, Marshall JL, Poste G, Vacirca JL, Vidal GA, Schwartzberg LS, Halbert DD, Voss A, Magee D, Miglarese MR, Famulok M, Mayer G, Spetzler D. Poly-ligand profiling differentiates trastuzumab-treated breast cancer patients according to their outcomes. Nature communications. 2018;9(1):1219. doi: 10.1038/s41467-018-03631-z. PubMed PMID: 29572535; PubMed Central PMCID: PMC5865185; https://www.ncbi.nlm.nih.gov/pubmed/29572535.
280. Driesbaugh KH, Branchetti E, Grau JB, Keeney SJ, Glass K, Oyama MA, Rioux N, Ayoub S, Sacks MS, **Quackenbush J**, Levy RJ, Ferrari G. Serotonin receptor 2B signaling with interstitial cell activation and leaflet remodeling in degenerative mitral regurgitation. J Mol Cell Cardiol. 2018;115:94-103. doi: 10.1016/j.yjmcc.2017.12.014. PubMed PMID: 29291394; PubMed Central PMCID: PMC5856457; https://www.ncbi.nlm.nih.gov/pubmed/29291394.
281. Hicks SC, Okrah K, Paulson JN, **Quackenbush J**, Irizarry RA, Bravo HC. Smooth quantile normalization. Biostatistics. 2018;19(2):185-98. doi: 10.1093/biostatistics/kxx028. PubMed PMID: 29036413; PubMed Central PMCID: PMC5862355; https://www.ncbi.nlm.nih.gov/pubmed/29036413.
282. Hosny A, Parmar C, **Quackenbush J**, Schwartz LH, Aerts H. Artificial intelligence in radiology. Nat Rev Cancer. 2018;18(8):500-10. doi: 10.1038/s41568-018-0016-5. PubMed PMID: 29777175; PubMed Central PMCID: PMC6268174; https://www.ncbi.nlm.nih.gov/pubmed/29777175.
283. Kuijjer ML, Paulson JN, Salzman P, Ding W, **Quackenbush J**. Cancer subtype identification using somatic mutation data. Br J Cancer. 2018;118(11):1492-501. doi: 10.1038/s41416-018-0109-7. PubMed PMID: 29765148; PubMed Central PMCID: PMC5988673; https://www.ncbi.nlm.nih.gov/pubmed/29765148.
284. Lopes-Ramos CM, Kuijjer ML, Ogino S, Fuchs CS, DeMeo DL, Glass K, **Quackenbush J**. Gene Regulatory Network Analysis Identifies Sex-Linked Differences in Colon Cancer Drug Metabolism. Cancer research. 2018;78(19):5538-47. doi: 10.1158/0008-5472.CAN-18-0454. PubMed PMID: 30275053; PubMed Central PMCID: PMC6169995; https://www.ncbi.nlm.nih.gov/pubmed/30275053.
285. Morrow JD, Cho MH, Platig J, Zhou X, DeMeo DL, Qiu W, Celli B, Marchetti N, Criner GJ, Bueno R, Washko GR, Glass K, **Quackenbush J**, Silverman EK, Hersh

CP. Ensemble genomic analysis in human lung tissue identifies novel genes for chronic obstructive pulmonary disease. Hum Genomics. 2018;12(1):1. doi: 10.1186/s40246-018-0132-z. PubMed PMID: 29335020; PubMed Central PMCID: PMC5769240; https://www.ncbi.nlm.nih.gov/pubmed/29335020.

286. Morrow JD, Glass K, Cho MH, Hersh CP, Pinto-Plata V, Celli B, Marchetti N, Criner G, Bueno R, Washko G, Choi AMK, **Quackenbush J**, Silverman EK, DeMeo DL. Human Lung DNA Methylation Quantitative Trait Loci Colocalize with Chronic Obstructive Pulmonary Disease Genome-Wide Association Loci. American journal of respiratory and critical care medicine. 2018;197(10):1275-84. doi: 10.1164/rccm.201707-1434OC. PubMed PMID: 29313708; PubMed Central PMCID: PMC5955059; https://www.ncbi.nlm.nih.gov/pubmed/29313708.
287. Padi M, **Quackenbush J**. Detecting phenotype-driven transitions in regulatory network structure. NPJ Syst Biol Appl. 2018;4:16. doi: 10.1038/s41540-018-0052-5. PubMed PMID: 29707235; PubMed Central PMCID: PMC5908977; https://www.ncbi.nlm.nih.gov/pubmed/29707235.
288. Parmar C, Barry JD, Hosny A, **Quackenbush J**, Aerts H. Data Analysis Strategies in Medical Imaging. Clinical cancer research: an official journal of the American Association for Cancer Research. 2018;24(15):3492-9. doi: 10.1158/1078-0432.CCR-18-0385. PubMed PMID: 29581134; PubMed Central PMCID: PMC6082690; https://www.ncbi.nlm.nih.gov/pubmed/29581134.
289. Qiu W, Guo F, Glass K, Yuan GC, **Quackenbush J**, Zhou X, Tantisira KG. Differential connectivity of gene regulatory networks distinguishes corticosteroid response in asthma. J Allergy Clin Immunol. 2018;141(4):1250-8. doi: 10.1016/j.jaci.2017.05.052. PubMed PMID: 28736268; PubMed Central PMCID: PMC5775928; https://www.ncbi.nlm.nih.gov/pubmed/28736268.
290. St Hilaire MA, Kristal BS, Rahman SA, Sullivan JP, **Quackenbush J**, Duffy JF, Barger LK, Gooley JJ, Czeisler CA, Lockley SW. Using a Single Daytime Performance Test to Identify Most Individuals at High-Risk for Performance Impairment during Extended Wake. Sci Rep. 2019;9(1):16681. Epub 2019/11/15. doi: 10.1038/s41598-019-52930-y. PubMed PMID: 31723161; PubMed Central PMCID: PMC6853981.
291. Schwede M, Waldron L, Mok SC, Wei W, Basunia A, Merritt MA, Mitsiades CS, Parmigiani G, Harrington DP, **Quackenbush J**, Birrer MJ, Culhane AC. The impact of stroma admixture on molecular subtypes and prognostic gene signatures in serous ovarian cancer. Cancer Epidemiol Biomarkers Prev. 2019. Epub 2019/12/25. doi: 10.1158/1055-9965.EPI-18-1359. PubMed PMID: 31871106.
292. Lozano Calderon SA, Garbutt C, Kim J, Lietz CE, Chen YL, Bernstein K, Chebib I, Nielsen GP, Deshpande V, Rubio R, Wang YE, **Quackenbush J**, Delaney T, Raskin K, Schwab J, Cote G, Spentzos D. Clinical and Molecular Analysis of Pathologic Fracture-associated Osteosarcoma: MicroRNA profile Is Different and Correlates with Prognosis. Clin Orthop Relat Res. 2019;477(9):2114-26. Epub 2019/08/08. doi: 10.1097/CORR.0000000000000867. PubMed PMID: 31389890; PubMed Central PMCID: PMC7000091.

293. Lopes-Ramos CM, Kuijjer ML, Ogino S, Fuchs CS, DeMeo DL, Glass K, **Quackenbush J**. Correction: Gene Regulatory Network Analysis Identifies Sex-Linked Differences in Colon Cancer Drug Metabolism. Cancer Res. 2019;79(8):2084. Epub 2019/04/17. doi: 10.1158/0008-5472.CAN-19-0678. PubMed PMID: 30987980.
294. Kuijjer ML, Tung MG, Yuan G, **Quackenbush J**, Glass K. Estimating Sample-Specific Regulatory Networks. iScience. 2019;14:226-40. Epub 2019/04/15. doi: 10.1016/j.isci.2019.03.021. PubMed PMID: 30981959; PubMed Central PMCID: PMC6463816.
295. Kuijjer ML, Hsieh PH, **Quackenbush J**, Glass K. lionessR: single sample network inference in R. BMC Cancer. 2019;19(1):1003. Epub 2019/10/28. doi: 10.1186/s12885-019-6235-7. PubMed PMID: 31653243; PubMed Central PMCID: PMC6815019.
296. Hersh CP, Adcock IM, Celedon JC, Cho MH, Christiani DC, Himes BE, Kaminski N, Mathias RA, Meyers DA, **Quackenbush J**, Redline S, Steiling KA, Tabor HK, Tobin MD, Wurfel MM, Yang IV, Koppelman GH. High-Throughput Sequencing in Respiratory, Critical Care, and Sleep Medicine Research. An Official American Thoracic Society Workshop Report. Ann Am Thorac Soc. 2019;16(1):1-16. Epub 2018/12/29. doi: 10.1513/AnnalsATS.201810-716WS. PubMed PMID: 30592451; PubMed Central PMCID: PMC6812157.
297. Gaynor SM, Sun R, Lin X, **Quackenbush J**. Identification of differentially expressed gene sets using the Generalized Berk-Jones statistic. Bioinformatics. 2019;35(22):4568-76. Epub 2019/05/08. doi: 10.1093/bioinformatics/btz277. PubMed PMID: 31062858; PubMed Central PMCID: PMC6853697.
298. Campbell PT, Ambrosone CB, Nishihara R, Aerts H, Bondy M, Chatterjee N, Garcia-Closas M, Giannakis M, Golden JA, Heng YJ, Kip NS, Koshiol J, Liu XS, Lopes-Ramos CM, Mucci LA, Nowak JA, Phipps AI, **Quackenbush J**, Schoen RE, Sholl LM, Tamimi RM, Wang M, Weijenberg MP, Wu CJ, Wu K, Yao S, Yu KH, Zhang X, Rebbeck TR, Ogino S. Proceedings of the fourth international molecular pathological epidemiology (MPE) meeting. Cancer Causes Control. 2019;30(8):799-811. Epub 2019/05/10. doi: 10.1007/s10552-019-01177-z. PubMed PMID: 31069578; PubMed Central PMCID: PMC6614001.
299. Altenbuchinger M, Weihs A, **Quackenbush J**, Grabe HJ, Zacharias HU. Gaussian and Mixed Graphical Models as (multi-)omics data analysis tools. Biochim Biophys Acta Gene Regul Mech. 2019:194418. Epub 2019/10/23. doi: 10.1016/j.bbagrm.2019.194418. PubMed PMID: 31639475.
300. Silverman EK, Schmidt H, Anastasiadou E, Altucci L, Angelini M, Badimon L, Balligand JL, Benincasa G, Capasso G, Conte F, Di Costanzo A, Farina L, Fiscon G, Gatto L, Gentili M, Loscalzo J, Marchese C, Napoli C, Paci P, Petti M, **Quackenbush J**, Tieri P, Viggiano D, Vilahur G, Glass K, Baumbach J. Molecular networks in Network Medicine: Development and applications. Wiley Interdiscip Rev Syst Biol Med. 2020:e1489. Epub 2020/04/21. doi: 10.1002/wsbm.1489. PubMed PMID: 32307915.

301. Schwede M, Waldron L, Mok SC, Wei W, Basunia A, Merritt MA, Mitsiades CS, Parmigiani G, Harrington DP, **Quackenbush J**, Birrer MJ, Culhane AC. The Impact of Stroma Admixture on Molecular Subtypes and Prognostic Gene Signatures in Serous Ovarian Cancer. Cancer Epidemiol Biomarkers Prev. 2020;29(2):509-19. Epub 2019/12/25. doi: 10.1158/1055-9965.EPI-18-1359. PubMed PMID: 31871106.
302. Lietz CE, Garbutt C, Barry WT, Deshpande V, Chen YL, Lozano-Calderon SA, Wang Y, Lawney B, Ebb D, Cote GM, Duan Z, Hornicek FJ, Choy E, Petur Nielsen G, Haibe-Kains B, **Quackenbush J**, Spentzos D. MicroRNA-mRNA networks define translatable molecular outcome phenotypes in osteosarcoma. Sci Rep. 2020;10(1):4409. Epub 2020/03/12. doi: 10.1038/s41598-020-61236-3. PubMed PMID: 32157112.
303. Fagny M, Platig J, Kuijjer ML, Lin X, **Quackenbush J**. Nongenic cancer-risk SNPs affect oncogenes, tumour-suppressor genes, and immune function. Br J Cancer. 2020;122(4):569-77. Epub 2019/12/07. doi: 10.1038/s41416-019-0614-3. PubMed PMID: 31806877; PubMed Central PMCID: PMC7028992.
304. Altenbuchinger M, Weihs A, **Quackenbush J**, Grabe HJ, Zacharias HU. Gaussian and Mixed Graphical Models as (multi-)omics data analysis tools. Biochim Biophys Acta Gene Regul Mech. 2020;1863(6):194418. Epub 2019/10/23. doi: 10.1016/j.bbagrm.2019.194418. PubMed PMID: 31639475; PubMed Central PMCID: PMC7166149.

Preprints:

1. Glass K, **Quackenbush J**, Kepner J. High Performance Computing of Gene Regulatory Networks using a Message-Passing Model. 2015.
2. Kuijjer M, Tung M, Yuan G, **Quackenbush J**, Glass K. Estimating sample-specific regulatory networks. eprint arXiv:150506440. 2015. doi: citeulike-article-id:13659997. https://arxiv.org/abs/1505.06440.
3. Safikhani Z, Freeman M, Smirnov P, El-Hachem N, She A, Quevedo R, Goldenberg A, Juul Birkbak N, Hatzis C, Shi L, Beck AH, Aerts HJ, **Quackenbush J**, Haibe-Kains B. Revisiting inconsistency in large pharmacogenomic studies. bioRxiv. 2015. doi: 10.1101/026153. http://biorxiv.org/content/early/2015/09/11/026153.
4. Chen C-Y, Lopes-Ramos CM, Kuijjer ML, Paulson JN, Sonawane AR, Fagny M, Platig J, Glass K, **Quackenbush J**, DeMeo DL. Sexual dimorphism in gene expression and regulatory networks across human tissues. bioRxiv. 2016. doi: 10.1101/082289. http://biorxiv.org/content/early/2016/10/20/082289.
5. Fagny M, Paulson JN, Kuijjer ML, Sonawane AR, Chen C-Y, Lopes-Ramos CM, Glass K, **Quackenbush J**, Platig J. A network-based approach to eQTL interpretation and SNP functional characterization. bioRxiv. 2016. doi: 10.1101/086587. http://biorxiv.org/content/early/2016/11/09/086587.

6. Hicks SC, Okrah K, Paulson JN, **Quackenbush J**, Irizarry RA, Corrada Bravo H. Smooth Quantile Normalization. bioRxiv. 2016. doi: 10.1101/085175. http://biorxiv.org/content/early/2016/11/03/085175.
7. Lopes-Ramos CM, Paulson JN, Chen C-Y, Kuijjer ML, Fagny M, Platig J, Sonawane AR, DeMeo DL, **Quackenbush J**, Glass K. Transcriptional landscape of cell lines and their tissues of origin. bioRxiv. 2016. doi: 10.1101/082065. http://biorxiv.org/content/early/2016/10/20/082065.
8. Safikhani Z, El-Hachem N, Quevedo R, Smirnov P, Goldenberg A, Juul Birkbak N, Mason CE, Hatzis C, Shi L, Aerts HJ, **Quackenbush J**, Haibe-Kains B. Assessment of pharmacogenomic agreement. bioRxiv. 2016. doi: 10.1101/048470. http://biorxiv.org/content/early/2016/04/13/048470.
9. Schlauch D, Glass K, Hersh CP, Silverman EK, **Quackenbush J**. Estimating Drivers of Cell State Transitions Using Gene Regulatory Network Models. bioRxiv. 2016. doi: 10.1101/089003. http://biorxiv.org/content/early/2016/11/21/089003.
10. Barry JD, Fagny M, Paulson JN, Aerts H, Platig J, **Quackenbush J**. Histopathological image QTL discovery of thyroid autoimmune disease variants. bioRxiv. 2017. doi: 10.1101/126730. https://www.biorxiv.org/content/biorxiv/early/2017/08/08/126730.full.pdf.
11. Kuijjer ML, Paulson JN, Salzman P, Ding W, **Quackenbush J**. Cancer subtype identification using somatic mutation data. bioRxiv. 2017. doi: 10.1101/228031. https://www.biorxiv.org/content/biorxiv/early/2017/12/03/228031.full.pdf.
12. Padi M, **Quackenbush J**. Phenotype-Driven Transitions In Regulatory Network Structure. bioRxiv. 2017. doi: 10.1101/142281. http://www.biorxiv.org/content/early/2017/05/25/142281.
13. Qiang J, Ding W, **Quackenbush J**, Chen P. Network-based Distance Metric with Application to Discover Disease Subtypes in Cancer. 2017.
14. Sonawane AR, Platig J, Fagny M, Chen C-Y, Paulson JN, Lopes-Ramos CM, DeMeo DL, **Quackenbush J**, Glass K, Kuijjer ML. Understanding Tissue-specific Gene Regulation. bioRxiv. 2017. doi: 10.1101/110601. http://biorxiv.org/content/early/2017/02/21/110601.
15. Wang YE, Kuznetsov L, Partensky A, Farid J, **Quackenbush J**. WebMeV: A Cloud Platform for Analyzing and Visualizing Cancer Genomic Data. bioRxiv. 2017. doi: 10.1101/147884.
16. Yuan G-C, Cai L, Elowitz M, Enver T, Fan G, Guo G, Irizarry R, Kharchenko P, Kim J, Orkin S, **Quackenbush J**, Saadatpour A, Schroeder T, Shivdasani R, Tirosh I. Challenges And Emerging Directions In Single-Cell Analysis. bioRxiv. 2017. doi: 10.1101/127761. https://www.biorxiv.org/content/biorxiv/early/2017/04/15/127761.full.pdf.
17. Fagny M, Platig J, Kuijjer ML, Lin X, **Quackenbush J**. Nongenic cancer-risk SNPs affect oncogenes, tumor suppressor genes, and immune function. bioRxiv. 2018.
18. Lopes-Ramos CM, Kuijjer ML, Ogino S, Fuchs CS, DeMeo DL, Glass K, **Quackenbush J**. Gene regulatory network analysis identifies sex-linked differences in colon cancer drug metabolism processes. bioRxiv. 2018. doi: 10.1101/277186. https://www.biorxiv.org/content/biorxiv/early/2018/03/07/277186.full.pdf.

19. Schwede M, Waldron L, Mok SC, Wei W, Basunia A, Merritt MA, Parmigiani G, Harrington D, **Quackenbush J**, Birrer MJ, Culhane A. The impact of stroma on the discovery of molecular subtypes and prognostic gene signatures in serous ovarian cancer. bioRxiv. 2018.
20. Gaynor S, Fagny M, Lin X, Platig J, **Quackenbush J**. Connectivity of variants in eQTL networks dictates reproducibility and functionality. bioRxiv. 2019.

Reviews:

1. **Quackenbush J**. Viva la revolution! A report from the FANTOM meeting. Nature Genet. 2000;26(3):255-6.
2. **Quackenbush J**. Computational analysis of microarray data. Nat Rev Genet. 2001;2(6):418-27.
3. **Quackenbush J**. Microarray data normalization and transformation. Nat Genet. 2002;32 Suppl:496-501
4. Dudoit S, Gentleman RC, **Quackenbush J**. Open source software for the analysis of microarray data. Biotechniques. 2003;Suppl:45-51.
5. **Quackenbush J**. Extracting meaning from functional genomics experiments. Toxicol Appl Pharmacol. 2005;207(2 Suppl):195-9.
6. **Quackenbush J**. Microarray analysis and tumor classification. N Engl J Med. 2006;354(23):2463-72.
7. **Quackenbush J**. From 'omes to biology. Anim Genet. 2006;37 Suppl 1:48-56.
8. **Quackenbush J**. Computational approaches to analysis of DNA microarray data. Methods Inf Med. 2006;45 Suppl 1:91-103.
9. **Quackenbush J**. Extracting biology from high-dimensional biological data. J Exp Biol. 2007;210(Pt 9):1507-17.
10. Bateman, A., and **Quackenbush, J. (2009).** Bioinformatics for next generation sequencing. Bioinformatics 25, 429.

Editorials:

1. **Quackenbush J**. Data standards for 'omic' science. Nat Biotechnol. 2004;22(5):613-4.
2. **Quackenbush J**, Salzberg SL. It is time to end the patenting of software. Bioinformatics. 2006;22(12):1416-7.
3. **Quackenbush J**. Weighing our measures of gene expression. Mol Syst Biol. 2006;2:63.
4. **Quackenbush J.** Standardizing the standards. Mol Syst Biol. 2006;2:2006 0010.
5. **Quackenbush J**, Irizarry RA. Response to Shields: 'MIAME, we have a problem'. Trends Genet. 2006;22(9):471-2.
6. **Quackenbush J.** Learning to share. Scientific American. 2014;311(1):S22. Epub 2014/07/01. PubMed PMID: 24974706.
7. **Quackenbush J.** Perspective: Learning to share. Nature. 2014;509(7502):S68. Epub 2014/05/30. doi: 10.1038/509S68a. PubMed PMID: 24870827; PubMed Central PMCID: PMC4205237.

Books:

1. Coulston H. C., **Quackenbush J.,** Brazma A., **Microarray Gene Expression Data Analysis: A Beginner's Guide**, Blackwell Publishing, 2003.
2. Kelley, J. M. and **Quackenbush J.,** "Sequencing PCR Products," in **PCR Methods Manual,** Second edition, M. Innis, D. Gelfand, and J. Sninsky, eds., Academic Press, 1999.
3. **Quackenbush J.,** Adams M. D., Fraser C. M., and Venter J. C., "Genome Projects:From Microbes to Man, From Sequence to Function," in **Genomics: Commercial Opportunities from a Scientific Revolution,** G. K. Dixon, L. G. Copping, and D. Livingstone, eds. BIOS Scientific Publishers, 1998.
4. **Quackenbush, J.** Using DNA Microarrays to Assay Gene Expression. *In:* A. Baxevanis and B. Ouellette (eds.), Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, third edition: John Wiley and Sons, 2004.
5. **Quackenbush, J.** The Human Genome: Book of Essential Knowledge. Imagine! Publishing, 2011.

Educational Material: I have taught numerous courses and workshops and make the materials freely available. I developed and co-taught a HarvardX course on reproducible research (https://www.edx.org/course/principles-statistical-computational-harvardx-ph527x).

Thesis: Gauge Field Theory in Two Space-Time Dimensions: Anomalies and Applications to String Models.

Nonprint Materials, i.e., film strips, films, videotapes and computer-based materials relevant to appointee's academic field. *Include description of item, who commissioned, purpose, users, penetration in field.*

None

Patents

None

Abstracts (*only those containing data not yet published in complete form.)*

None

EXHIBIT 2

EXHIBIT 2 - MATERIALS CONSIDERED BY JOHN QUACKENBUSH, PH.D.

PATENTS	BEG PAGE NO.	END PAGE NO.
U.S. Patent No. 8,703,652	B0001	B0024
U.S. Patent No. 9,845,497	B0025	B0048
U.S. Patent No. 10,329,607	B0049	B0073
PATENT APPLICATION PUBLICATIONS		
U.S. Patent Application Publication No. 2005/0042654 A1 (February 24, 2005)	B0191	B0212
U.S. Patent Application Publication No. 2007/0207466 A1 (September 6, 2007)	B0401	B0427
U.S. Patent Application Publication No. 2009/0087847 A1 (April 2, 2009)	B1220	B1290
International Publication No. WO 2005/023091 A2 (March 17, 2005)	B0995	B0995
International Publication No. WO 2008/118988 A1 (October 2, 2008)	B1050	B1067
International Publication No. WO 2009/013492 A1 (January 29, 2009)	B1105	B1118
PRODUCED DOCUMENTS		
CAREDXVI00001865-CAREDXVI00001868 - Declaration by John Beausang under 37 C.F.R. § 1.132 that appears in the prosecution history file of the U.S. Patent No. 8,703,652	B0260	B0263
ARTICLES, BOOKS, MANUSCRIPTS, PRESS RELEASES, AND PUBLICATIONS		
T. Nordstrom, et al., *Direct Analysis of Single-Nucleotide Polymorphism on Double-Stranded DNA by Pyrosequencing*, Biotechnol. Appl. Biochem., 31:107–112 (2000)	B0074	B0079
Speicher, Antnarakis, and Motulsky, *Vogel and Motlsky's Human Genetics, Problems and Approaches*, 4th Ed. (June 28, 2009)	B0481	B0580
Y. Lo, et al., *Presence of Donor-Specific DNA In Plasma of Kidney and Liver-Transplant Recipients*, The Lancet, Vol. 351, pp. 1329–1330 (May 2, 1998)	B0080	B0081
Helicos Biosciences Corp.'s Form 10-K (December 31, 2007)	B0581	B0606
The International HapMap Project, Nature, Vol. 426, pp. 789–796 (December 2003)	B0086	B0093
E. Lander, et al., *Initial Sequencing and Analysis of the Human Genome, International Human Genome Sequencing Consortium*, Nature, Vol. 409, pp. 860–921 (February, 15, 2001)	B0094	B0155
Frequently Asked Questions and Answers about the Human Genome Project and its Impact on the Field of Genomics	B0156	B0163
A Global Reference for Human Genetic Variation, The 1000 Genomes Project Consortium, published in final edited form as: *Nature*, 526 (7571) pp. 68–74 (October 1, 2015)	B0164	B0183
K.J. Livak, *Towards Fully Automated Genome-Wide Polymorphism Screening*, Nature Genetics, Vol. 9, pp. 341–342 (April 1995)	B0184	B0185
S. Germer, et al., *High-Throughput SNP Allele-Frequency Determination in Pooled DNA Samples by Kinetic PCR*, Genome Research, 10:258–266 (2000)	B1291	B1300
A. Lichtenstein, et al., *Circulating Nucleic Acids and Apotosis*, Annals New York Academy of Sciences, pp. 239–249 (2001)	B0213	B0223
RainDance Technologies Introduces the RDT 1000, RainDance Technologies (November 12, 2008)	B0224	B0225
K. Voelkerding, et al., *Next-Generation Sequencing: From Basic Research to Diagnostics*, Clinical Chemistry 55:4, pp. 641–658 (2009)	B0226	B0243
M. Ogren, *Whole-Genome SNP Genotyping*, The Scientist, (June 1, 2003)	B0244	B0250
A. Kahvejian, et al., *What Would You Do if You Could Sequence Everything?*, Nature Biotechnology, Vol. 26, No. 10, 1125–1133 (October 2008)	B0251	B0259
C. Hutchison III, *DNA Sequencing: Bench to Bedside and Beyond*, Nucleic Acids Research, Vol. 35, No. 18, pp. 6227–6237 (2007)	B0264	B0274
E. Green, *Strategies for the Systematic Sequencing of Complex Genomes*, Nature Reviews, Vol. 2, pp. 573–583 (August 2001)	B0275	B0285
P. Nyrén, et al., *Solid Phase DNA Minisequencing by an Enzymatic Luminometric Inorganic Pyrophosphate Detection Assay*, Analytical Biochemistry, 208 pp. 171–175 (1993)	B0286	B0290
M. Ronaghi, *Pyrosequencing Sheds Light on DNA Sequencing*, Genome Research, 11:3–11 (2001)	B0291	B0300
S. Marsh, *Pyrosequencing Protocols*, Methods in Molecular Biology No. 373 (2007)	B0694	B0713
J. Heather, *The Sequence of Sequencers: The History of Sequencing DNA*, Genomics, Vol. 107, pp. 1–8 (2016)	B0301	B0308
A. Coombs, *The Sequencing Shakeup*, Nature Biotechnology, Vol. 26, No. 10, pp. 1109–1112 (October 2008)	B0309	B0312
A. Ahmadian, et al., *Single-Nucleotide Polymorphism Analysis by Pyrosequencing*, Analytical Biochemistry 280, pp. 103–110 (2000)	B0313	B0320
History of Sequencing by Synthesis, Illumina (2020)	B0321	B0323

EXHIBIT 2 - MATERIALS CONSIDERED BY JOHN QUACKENBUSH, PH.D.

ARTICLES, BOOKS, MANUSCRIPTS, PRESS RELEASES, AND PUBLICATIONS		
Illumina, Inc.'s Form 10-K for the fiscal year ended (December 28, 2008)	B0905	B945
O. Morozova and M. Marra, *Applications of Next-Generation Sequencing Technologies in Functional Genomics*, Genomics, 92:255–264 (2008)	B0324	B0333
S. Quake, et al., *Digital PCR Provides Sensitive and Absolute Calibration for High Throughput Sequencing*, BMC Genomics, 10:116 (March 19, 2009)	B0334	B0345
Y. Shen, et al., *A SNP Discovery Method to Assess Variant Allele Probability from Next-Generation Resequencing Data*, Genome Research, 20:273–280 (2010)	B0346	B0354
S, Quake, T. Harris, et al., *Single-Molecule DNA Sequencing of a Viral Genome*, Science, Vol. 320 (April 4, 2008)	B0355	B0359
J. Sampson and H. Zhao, *Identifying Individuals in a Complex Mixture of DNA with Unknown Ancestry*, Statistical Applications in Genetics and Molecular Biology, Vol. 8(1), pp. 1–27 (2009)	B0360	B0388
K. Conger, *To Better Detect Heart Transplant Rejections, Scientist Test for Traces of Donor's Genome*, Stanford Medicine (March 28, 2011)	B0389	B0391
K. Conger, *New Stanford Blood Test Identifies Heart-Transplant Rejection Earlier than Biopsy can*, Stanford Engineering (June 18, 2014)	B0392	B0394
C. Ding, et al., *MS Analysis of Single-Nucleotide Differences in Circulating Nucleic Acids: Application to Noninvasive Prenatal Diagnosis*, PNAS Vol. 101, No. 29, pp. 10762–10767 (July 10, 2004)	B0395	B0400
R. Dhallan, et al., *A Non-Invasive Test for Prenatal Diagnosis Based on Fetal DNA Present in Maternal Blood: A Preliminary Study*, The Lancet, 369: 474–481 (February 10, 2007)	B0428	B0435
J. Beck, et al., *Profile of the Circulating DNA in Apparently Healthy Individuals*, Clinical Chemistry 55:4, pp. 730–738 (2009)	B0436	B0444
J. Beck, et al., *Next Generation Sequencing of Serum Circulating Nucleic Acids from Patients with Invasive Ductal Breast Cancer Reveals Differences to Healthy and Nonmalignant Controls*, Molecular Cancer Research, 8:335–342 (March 9, 2010)	B0445	B0453
P. Gordon, et al., *Disease-Specific Motifs can be Identified in Circulating Nucleic Acids from Live Elk and Cattle Infected with Transmissible Spongiform Encephalopathies*, Nucleic Acids Research, Vol. 37, No. 2, pp. 550–556 (2009)	B0454	B0460
K. Holt and Y. Teo, *Detecting SNPs and Estimating Allele Frequencies in Clonal Bacterial Populations by Sequencing Pooled DNA*, Bioinformatics, Vol. 25 No. 16, pp. 2074–2075 (June 3, 2009)	B0461	B0462
Genome Analyzer Sy stem, Illumina (2007)	B0463	B0466
Multiplexed Sequencing with the Illumina Genome Analyzer System, Illumina (2008)	B0467	B0470
R. Li, et al., *SNP Detection for Massively Parallel Whole-Genome Resequencing*, Genome Research, 19:1124–1132 (2009)	B0471	B0480
R. Thomas, *Sensitive Mutation Detection in Heterogeneous Cancer Specimens by Massively Parallel Picoliter Reactor Sequencing*, Nature Medicine, Vol. 12, No. 7, pp. 852–855 (July 2006)	B1301	B1304.1
M. Margulies, *Genome Sequencing in Microfabricated High-Density Picolitre Reactors*, Nature, (2005)	B1305	B1309