Illumina Inc.

v.

Molecular Loop Biosciences Inc.

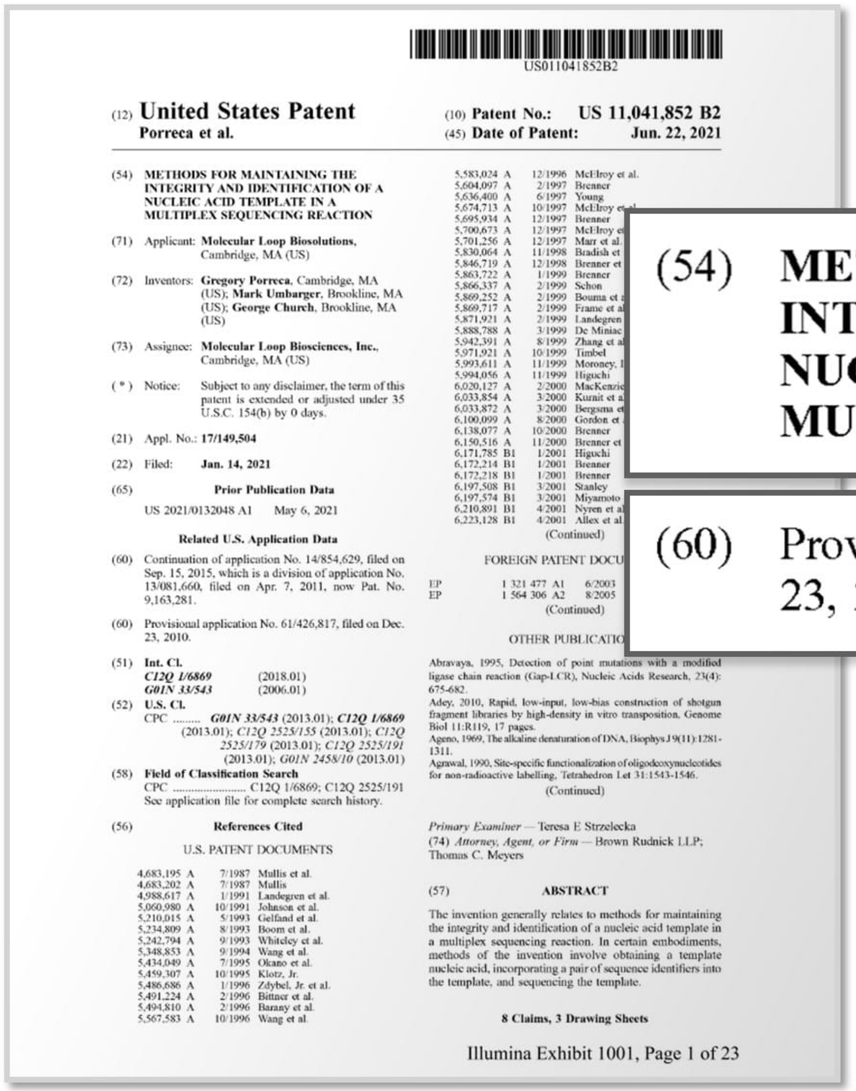
IPR2024-00964 (US Patent 11,041,852)
IPR2024-00965 (US Patent 11,768,200)

Petitioner Illumina's Demonstratives

September 23, 2025

illumina

'852 and '200 Patents



US011041852B2

(12) **United States Patent**
Porreca et al.

(10) **Patent No.: US 11,041,852 B2**
(45) **Date of Patent: Jun. 22, 2021**

(54) **METHODS FOR MAINTAINING THE INTEGRITY AND IDENTIFICATION OF A NUCLEIC ACID TEMPLATE IN A MULTIPLEX SEQUENCING REACTION**

(71) Applicant: **Molecular Loop Biosolutions**, Cambridge, MA (US)

(72) Inventors: **Gregory Porreca**, Cambridge, MA (US); **Mark Umbarger**, Brookline, MA (US); **George Church**, Brookline, MA (US)

(73) Assignee: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **17/149,504**

(22) Filed: **Jan. 14, 2021**

(65) **Prior Publication Data**

US 2021/0132048 A1 May 6, 2021

Related U.S. Application Data

(60) Continuation of application No. 14/854,629, filed on Sep. 15, 2015, which is a division of application No. 13/081,660, filed on Apr. 7, 2011, now Pat. No. 9,163,281.

(60) Provisional application No. 61/426,817, filed on Dec. 23, 2010.

(51) **Int. Cl.**
C12Q 1/6869 (2018.01)
G01N 33/543 (2006.01)

(52) **U.S. Cl.**
CPC ***G01N 33/543*** (2013.01); ***C12Q 1/6869*** (2013.01); *C12Q 2525/155* (2013.01); *C12Q 2525/179* (2013.01); *C12Q 2525/191* (2013.01); *G01N 2458/10* (2013.01)

(58) **Field of Classification Search**
CPC C12Q 1/6869; C12Q 2525/191
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,683,195 A 7/1987 Mullis et al.
4,683,202 A 7/1987 Mullis
4,988,617 A 1/1991 Landegren et al.
5,060,980 A 10/1991 Johnston et al.
5,210,015 A 5/1993 Gelfand et al.
5,234,809 A 8/1993 Boom et al.
5,242,794 A 9/1993 Whiteley et al.
5,348,853 A 9/1994 Wang et al.
5,434,049 A 7/1995 Okano et al.
5,459,307 A 10/1995 Klotz, Jr.
5,486,686 A 1/1996 Zdybel, Jr. et al.
5,491,224 A 2/1996 Bittner et al.
5,494,810 A 2/1996 Barany et al.
5,567,583 A 10/1996 Wang et al.
5,583,024 A 12/1996 McElroy et al.
5,604,097 A 2/1997 Brenner
5,636,400 A 6/1997 Young
5,674,713 A 10/1997 McElroy et al.
5,695,934 A 12/1997 Brenner
5,700,673 A 12/1997 McElroy et
5,701,256 A 12/1997 Marr et al.
5,830,064 A 11/1998 Bradish et
5,846,719 A 12/1998 Brenner et
5,863,722 A 1/1999 Brenner
5,866,337 A 2/1999 Schon
5,869,252 A 2/1999 Bouma et
5,869,717 A 2/1999 Frame et al.
5,871,921 A 2/1999 Landegren
5,888,788 A 3/1999 De Miniac
5,942,391 A 8/1999 Zhang et al.
5,971,921 A 10/1999 Timbel
5,993,611 A 11/1999 Moroney, I
5,994,056 A 11/1999 Higuchi
6,020,127 A 2/2000 MacKenzie
6,033,854 A 3/2000 Kurnit et a
6,033,872 A 3/2000 Bergsma et
6,100,099 A 8/2000 Gordon et
6,138,077 A 10/2000 Brenner
6,150,516 A 11/2000 Brenner et
6,171,785 B1 1/2001 Higuchi
6,172,214 B1 1/2001 Brenner
6,172,218 B1 1/2001 Brenner
6,197,508 B1 3/2001 Stanley
6,197,574 B1 3/2001 Miyamoto
6,210,891 B1 4/2001 Nyren et al.
6,223,128 B1 4/2001 Allex et al.
(Continued)

FOREIGN PATENT DOCU

EP 1 321 477 A1 6/2003
EP 1 564 306 A2 8/2005
(Continued)

OTHER PUBLICATIO

Abravaya, 1995, Detection of point mutations with a modified ligase chain reaction (Gap-LCR), Nucleic Acids Research, 23(4): 675-682.
Adey, 2010, Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition, Genome Biol 11:R119, 17 pages.
Ageno, 1969, The alkaline denaturation of DNA, Biophys J 9(11):1281-1311.
Agrawal, 1990, Site-specific functionalization of oligodeoxynucleotides for non-radioactive labelling, Tetrahedron Let 31:1543-1546.
(Continued)

Primary Examiner — Teresa E Strzelecka
(74) *Attorney, Agent, or Firm* — Brown Rudnick LLP; Thomas C. Meyers

(57) **ABSTRACT**

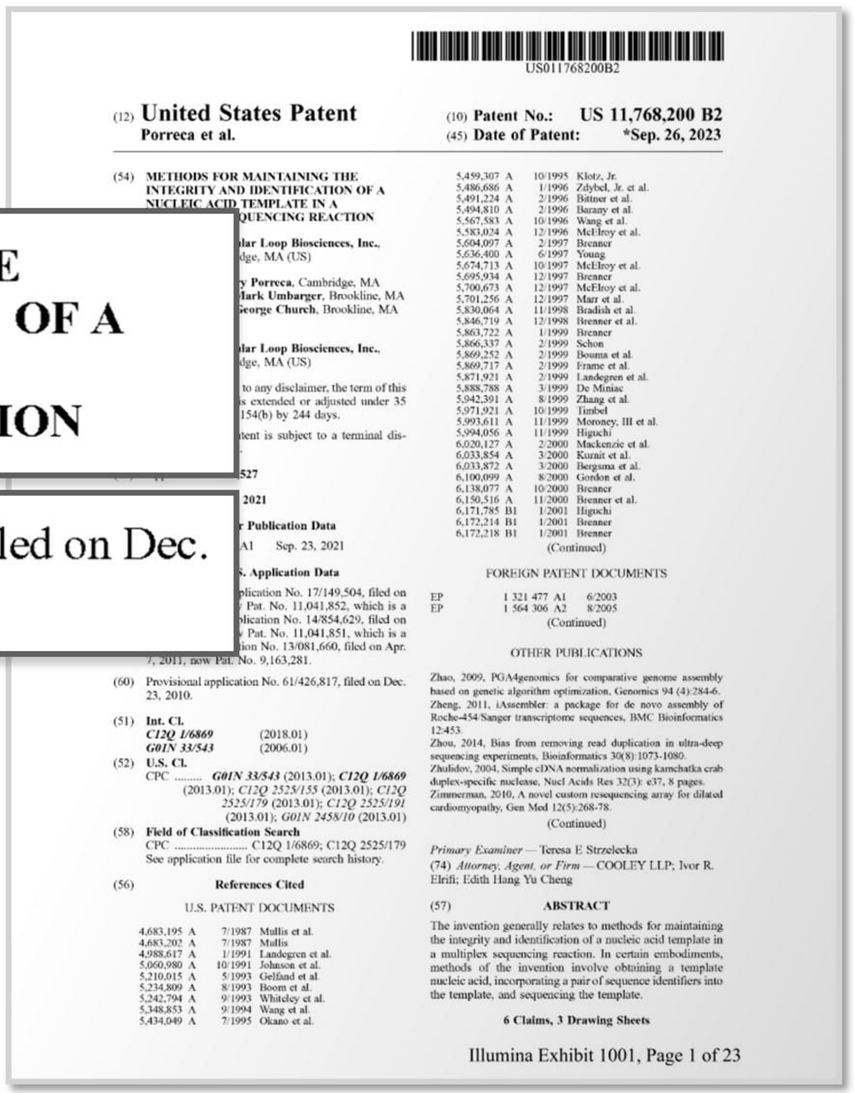
The invention generally relates to methods for maintaining the integrity and identification of a nucleic acid template in a multiplex sequencing reaction. In certain embodiments, methods of the invention involve obtaining a template nucleic acid, incorporating a pair of sequence identifiers into the template, and sequencing the template.

8 Claims, 3 Drawing Sheets

Illumina Exhibit 1001, Page 1 of 23

(54) **METHODS FOR MAINTAINING THE INTEGRITY AND IDENTIFICATION OF A NUCLEIC ACID TEMPLATE IN A MULTIPLEX SEQUENCING REACTION**

(60) Provisional application No. 61/426,817, filed on Dec. 23, 2010.



US011768200B2

(12) **United States Patent**
Porreca et al.

(10) **Patent No.: US 11,768,200 B2**
(45) **Date of Patent: *Sep. 26, 2023**

(54) **METHODS FOR MAINTAINING THE INTEGRITY AND IDENTIFICATION OF A NUCLEIC ACID TEMPLATE IN A QUENCING REACTION**

lar Loop Biosciences, Inc., dge, MA (US)

y Porreca, Cambridge, MA
lark Umbarger, Brookline, MA
eorge Church, Brookline, MA

lar Loop Biosciences, Inc., dge, MA (US)

to any disclaimer, the term of this s extended or adjusted under 35 154(b) by 244 days.

ent is subject to a terminal dis-

527

2021

r Publication Data

A1 Sep. 23, 2021

. Application Data

plication No. 17/149,504, filed on Pat. No. 11,041,852, which is a blication No. 14/854,629, filed on Pat. No. 11,041,851, which is a ion No. 13/081,660, filed on Apr. 7, 2011, now Pat. No. 9,163,281.

(60) Provisional application No. 61/426,817, filed on Dec. 23, 2010.

(51) **Int. Cl.**
C12Q 1/6869 (2018.01)
G01N 33/543 (2006.01)

(52) **U.S. Cl.**
CPC ***G01N 33/543*** (2013.01); ***C12Q 1/6869*** (2013.01); *C12Q 2525/155* (2013.01); *C12Q 2525/179* (2013.01); *C12Q 2525/191* (2013.01); *G01N 2458/10* (2013.01)

(58) **Field of Classification Search**
CPC C12Q 1/6869; C12Q 2525/179
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,683,195 A 7/1987 Mullis et al.
4,683,202 A 7/1987 Mullis
4,988,617 A 1/1991 Landegren et al.
5,060,980 A 10/1991 Johnston et al.
5,210,015 A 5/1993 Gelfand et al.
5,234,809 A 8/1993 Boom et al.
5,242,794 A 9/1993 Whiteley et al.
5,348,853 A 9/1994 Wang et al.
5,434,049 A 7/1995 Okano et al.
5,459,307 A 10/1995 Klotz, Jr.
5,486,686 A 1/1996 Zdybel, Jr. et al.
5,491,224 A 2/1996 Bittner et al.
5,494,810 A 2/1996 Barany et al.
5,567,583 A 10/1996 Wang et al.
5,583,024 A 12/1996 McElroy et al.
5,604,097 A 2/1997 Brenner
5,636,400 A 6/1997 Young
5,674,713 A 10/1997 McElroy et al.
5,695,934 A 12/1997 Brenner
5,700,673 A 12/1997 McElroy et al.
5,701,256 A 12/1997 Marr et al.
5,830,064 A 11/1998 Bradish et al.
5,846,719 A 12/1998 Brenner et al.
5,863,722 A 1/1999 Brenner
5,866,337 A 2/1999 Schon
5,869,252 A 2/1999 Bouma et al.
5,869,717 A 2/1999 Frame et al.
5,871,921 A 2/1999 Landegren et al.
5,888,788 A 3/1999 De Miniac
5,942,391 A 8/1999 Zhang et al.
5,971,921 A 10/1999 Timbel
5,993,611 A 11/1999 Moroney, III et al.
5,994,056 A 11/1999 Higuchi
6,020,127 A 2/2000 Mackenzie et al.
6,033,854 A 3/2000 Kurnit et al.
6,033,872 A 3/2000 Bergsma et al.
6,100,099 A 8/2000 Gordon et al.
6,138,077 A 10/2000 Brenner
6,150,516 A 11/2000 Brenner et al.
6,171,785 B1 1/2001 Higuchi
6,172,214 B1 1/2001 Brenner
6,172,218 B1 1/2001 Brenner
(Continued)

FOREIGN PATENT DOCUMENTS

EP 1 321 477 A1 6/2003
EP 1 564 306 A2 8/2005
(Continued)

OTHER PUBLICATIONS

Zhao, 2009, PGA4genomics for comparative genome assembly based on genetic algorithm optimization, Genomics 94 (4):284-6.
Zheng, 2011, iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequences, BMC Bioinformatics 12:453.
Zhou, 2014, Bias from removing read duplication in ultra-deep sequencing experiments, Bioinformatics 30(8):1073-1080.
Zhulidov, 2004, Simple cDNA normalization using kamchatka crab duplex-specific nuclease, Nucl Acids Res 32(3): e37, 8 pages.
Zimmerman, 2010, A novel custom resequencing array for dilated cardiomyopathy, Gen Med 12(5):268-78.
(Continued)

Primary Examiner — Teresa E Strzelecka
(74) *Attorney, Agent, or Firm* — COOLEY LLP; Ivor R. Elrifi; Edith Hang Yu Cheng

(57) **ABSTRACT**

The invention generally relates to methods for maintaining the integrity and identification of a nucleic acid template in a multiplex sequencing reaction. In certain embodiments, methods of the invention involve obtaining a template nucleic acid, incorporating a pair of sequence identifiers into the template, and sequencing the template.

6 Claims, 3 Drawing Sheets

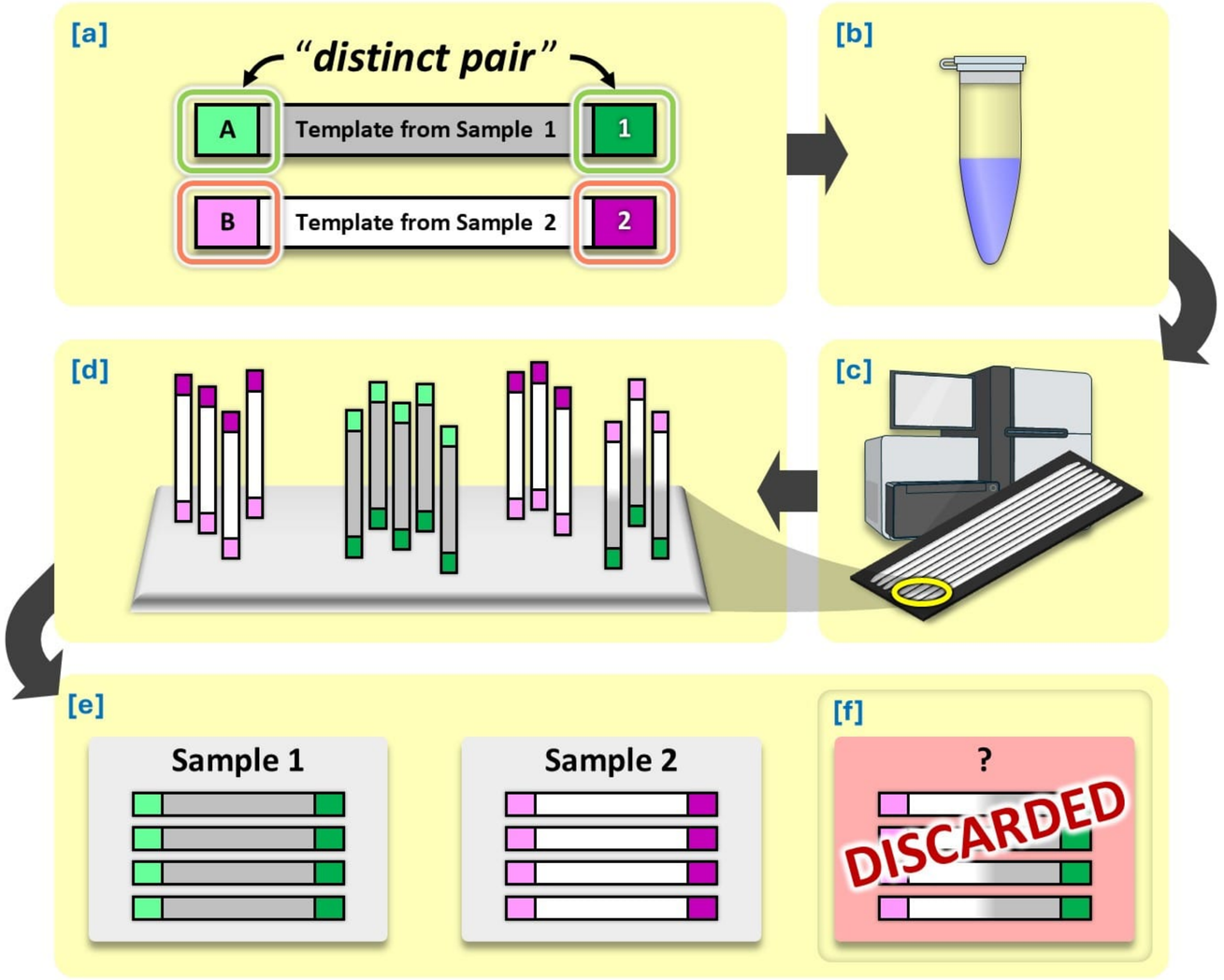
Illumina Exhibit 1001, Page 1 of 23

illumina

'852 Patent (IPR2024-00964)

Challenged Claims – '852 Patent

1 [pre]	A method for reducing template cross-over error introduced during sequencing workflow, the method comprising the steps of
[a]	**incorporating** at least two members of a plurality of identifier sequences into template nucleic acids obtained from at least two different samples, wherein said two members constitute **a distinct pair** associated with said template;
[b]	**combining** said nucleic acid templates into a single sample;
[c]	**amplifying** said nucleic acid templates on **a surface of a flow cell thereby to form clusters,**
[d]	**at least one of which comprises a chimeric sequence** comprising a combination of identifier sequences that are different than any of said distinct pairs;
[e]	**sequencing** amplicons obtained in said amplifying step; and
[f]	**discarding** sequence reads obtained in said sequencing step that contain said **combination of identifier sequences that are different than any of said distinct pairs**, thereby to reduce cross-over error.



illumina

Prosecution History of the '852 Patent

47 [pre]	A method for reducing template cross-over error introduced during sequencing workflow, the method comprising the steps of
[a]	incorporating at least two members of a plurality of identifier sequences into template nucleic acids obtained from at least two different samples, wherein said two members constitute a distinct pair associated with said template;
[b]	combining said nucleic acid templates into a single sample;
[c]	amplifying said nucleic acid templates on a surface of a flow cell thereby to form clusters,
[d]	at least one of which comprises a chimeric sequence comprising a combination of identifier sequences that are different than any of said distinct pairs;
[e]	sequencing amplicons obtained in said amplifying step; and
[f]	discarding sequence reads obtained in said sequencing step that contain said combination of identifier sequences that are different than any of said distinct pairs, thereby to reduce cross-over error.

Office Action on Feb. 26, 2021 (EX1007, 192-197)

8. Claims 47-54 are rejected under pre-AIA 35 U.S.C. 102(b) as being anticipated by Parameswaran et al. (Nucl. Acids Res., vol. 35, e130, pp. 1-9, 2007; cited in the IDS).

Regarding claim 47, Parameswaran et al. teach a method for reducing template cross-over error introduced during sequencing workflow, the method comprising the steps of

Response on April 13, 2021 (EX1007, 301-306)

inherent in pyrosequencing. *See*, Parameswaran at page 4. However, Parameswaran does not mention forming clusters in a flow cell. That is not at all surprising given that Parameswaran is all about pyrosequencing. In pyrosequencing, one prepares a series of bead-bound templates and ultimately monitors the liberation of diphosphates as and indication of base incorporation during sequencing.

Notice of Allowance on Jul. 27, 2021 (EX1007, 312-318)

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

NOTICE OF ALLOWANCE AND FEE(S) DUE

Grounds for IPR2024-00964 ('852 Patent)

Parameswaran (EX1004)

Published online 11 October 2007 *Nucleic Acids Research, 2007, Vol. 35, No. 19 e130*
doi:10.1093/nar/gkm760

A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing

Poornima Parameswaran[1], Roxana Jalili[3], Li Tao[4], Shadi Shokralla[3], Baback Gharizadeh[3], Mostafa Ronaghi[3] and Andrew Z. Fire[2,*]

[1]Department of Microbiology and Immunology, [2]Department of Pathology and Department of Genetics, Stanford University School of Medicine, [3]Stanford Genome Technology Center and [4]Department of Biological Sciences, Stanford University, Stanford, CA-94305, USA

Received July 30, 2007; Revised September 10, 2007; Accepted September 11, 2007

ABSTRACT

Multiplexed currently limi sequenced in sequences c segregation number of ch obscures ava these limitati ing approach DNA from in subsequent Forty-eight described: ea unique with improved rea nations of fo used to sequ libraries for each set of '*n*' forward and '*n*' reverse barcodes, for each defined set of cloning-linkers. In two pilot series of barcoded sequencing using the GS20 Sequencer (454/Roche), we found that over 99.8% of obtained sequences could be assigned to 25 independent, uniquely barcoded libraries based on the presence of either a perfect forward or a perfect reverse barcode. The false-discovery rate, as measured by the percentage of sequences with unexpected perfect pairings of unmatched forward and reverse barcodes, was estimated to be <0.005%.

quencers because their sequencing technology is based on the detection of pyrophosphates released during DNA synthesis (3). A few sequencing runs using 454/Roche's pyrosequencing platform can generate sufficient coverage for assembling entire microbial genomes (4,5), for the discovery, identification and quantitation of small RNAs (6,7), and for the detection of rare variations in cancers (8), among many other applications. For analysis of multiple libraries, the currently available 454/Roche pyrosequencers can accommodate a maximum of only 16 independent samples, which have to be physically separated using manifolds on the sequencing medium. These separation manifolds occlude wells on the

*To whom correspondence should be addressed. Tel: +1 650 723 2885; Fax: +1 650 724 9070; Email: afire@stanford.edu

© 2007 The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Detected and discarded **chimeric sequences** formed during multiplex sequencing using **unique dual-indexing**

+

Gloor (EX1005)

arXiv:1007.5075v1 [q-bio.GN] 28 Jul 2010

Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products

Abstract

Introduction

Ground 1

Dual-indexed, multiplex sequencing on **Illumina sequencer**

+

Bentley (EX1006)

nature

ARTICLES

Accurate whole hu sequencing using r chemistry

Ground 2

Genomic sequencing using **Illumina sequencer**

illumina

Illumina Sequencing Was One of Three Major NGS Platforms

Parameswaran (EX1004, 1)

Published online 11 October 2007 Nucleic Acids Research, 2007, Vol. 35, No. 19 e130

A pyr
desig
samp

Poornima
Baback

[1]Departmer
University S
Stanford U

Received July

ABSTRACT

Multiplexe
currently li
sequenced
sequences
segregatio
number of
obscures
these limit
ing approa
DNA from
subsequen
Forty-eight
described:
unique wit
improved
nations of forward and reverse barcodes may be used to sequence from as many as n^2 independent libraries for each set of 'n' forward and 'n' reverse barcodes, for each defined set of cloning-linkers. In two pilot series of barcoded sequencing using the GS20 Sequencer (454/Roche), we found that over 99.8% of obtained sequences could be assigned to 25 independent, uniquely barcoded libraries based on the presence of either a perfect forward or a perfect reverse barcode. The false-discovery rate, as measured by the percentage of sequences with unexpected perfect pairings of unmatched forward and reverse barcodes, was estimated to be <0.005%.

barcodes and sequences of interest.

The 454/Roche Genome Sequencers are called pyrosequencers because their sequencing technology is based on the detection of pyrophosphates released during DNA synthesis (3). A few sequencing runs using 454/Roche's pyrosequencing platform can generate sufficient coverage for assembling entire microbial genomes (4,5), for the discovery, identification and quantitation of small RNAs (6,7), and for the detection of rare variations in cancers (8), among many other applications. For analysis of multiple libraries, the currently available 454/Roche pyrosequencers can accommodate a maximum of only 16 independent samples, which have to be physically separated using manifolds on the sequencing medium. These separation manifolds occlude wells on the

*To whom correspondence should be addressed. Tel: +1 650 723 2885; Fax: +1 650 724 9070; Email: afire@stanford.edu

© 2007 The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

With the advent of high-throughput sequencing, characterization of the nucleic acid world is proceeding at an accelerated pace. Three major high-throughput sequencing platforms are in use today: the Genome Sequencers from (1) Roche/454 Life Sciences [GS-20 or GS-FLX; (1)], the 1G Analyzer from (2) Illumina/Solexa (2) and (3) the SOLiD System from Applied Biosystems (http://solid.applied biosystems.com). Comparison across the three platforms

samples. The Illumina sequencing method generated ~ 12 million useable reads at a cost of ~ 0.03 cents per read, an approximate order of magnitude cheaper than the per-read cost of pyrosequencing. The cheaper per-read costs allows economical experiments on large numbers of samples at very large sequencing depths. Since Illumina sequencing is now capable of ~ 100 nt long reads from

Gloor (EX1005, 2)

:1007.5075v1 [q-bio.GN] 28 Jul 2010

Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products

Gregory B. Gloor [1,*], Ruben Hummelen [2,3], Jean M. Macklaim [1,2], Russell J. Dickson [1] Andrew D. Fernandes [1,4], Roderick MacPhee [2,5], Gregor Reid [1,2,5,6]
1 Department of Biochemistry, University of Western Ontario, London, ON, Canada
2 Canadian Research & Development Centre for Probiotics, Lawson Health Research Institute, London, Canada
3 Department of Public Health, Erasmus MC, University Medical Centre Rotterdam, The Netherlands
4 Department Applied Mathematics
5 Department of Microbiology & Immunology, University of Western Ontario, London, ON, Canada
6 Department of Surgery, University of Western Ontario, London, ON, Canada
* E-mail: ggloor@uwo.ca

Abstract

We developed a low-cost, high-throughput microbiome profiling method that uses combinatorial sequence tags attached to PCR primers that amplify the rRNA V6 region. Amplified PCR products are sequenced using an Illumina paired-end protocol to generate millions of overlapping reads. Combinatorial sequence tagging can be used to examine hundreds of samples with far fewer primers than is required when sequence tags are incorporated at only a single end. The number of reads generated permitted saturating or near-saturating analysis of samples of the vaginal microbiome. The large number of reads allowed an in-depth analysis of errors, and we found that PCR-induced errors composed the vast majority of non-organism derived species variants, an observation that has significant implications for sequence clustering of similar high-throughput data. We show that the short reads are sufficient to assign

illumina

NGS Used PCR at Multiple Sequencing Stages

Meyer (EX1015, 8-9)

Indexing PCR and Pooling

* * *

22. Prepare a PCR master mix for the required number of reactions. Dispense the master mix into a 96-well reaction plate, and then add template DNA and a different indexing primer to each well using a multichannel pipette.

* * *

23. Mix and perform cycling using the following temperature profile:

* * *

26. Determine the DNA concentration, and pool the indexed libraries in equimolar ratios.

* * *

Target Capture and/or Sequencing on the Illumina Platform

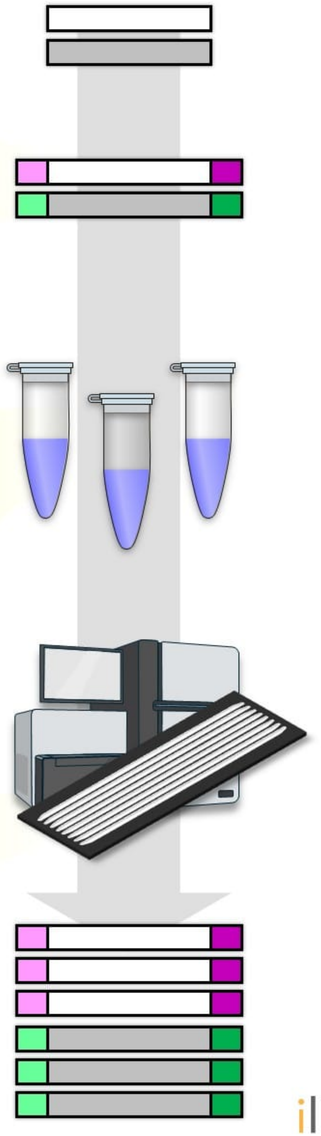
27. For target capture on microarrays, follow, for example, the exact procedure given in the protocol of Hodges et al. (2009) with the following modifications:

    i. Use a different set of blocking oligos (BO1-BO6).

    ii. Use primers IS5 and IS6 at an annealing temperature of 60°C for amplifying the library pool after capture.

28. For sequencing and data analysis, use the recipes, kits, and analysis tools for multiplex sequencing provided by Illumina.

Illumina Exhibit 1015, Page 1 of 25



illumina

NGS Was an "Ideal Breeding Ground" for Chimera Formation

Lahr **(EX1017, 1)**	"One worrisome aspect of PCR-based studies is the phenomenon of PCR-mediated recombination, or chimera formation."
Meyerhans (EX1018, 1)	"[W]hen target sequences, such as multigene families, alleles or RNA viruses are amplified, the PCR products will be heterogeneous. In such instances recombinant or 'shuffled' molecules, i.e. molecules composed of parts of two different sequences artificially combined during PCR may be produced."
Paabo **(EX1019, 1)**	"[T]he prematurely terminated product then jumps to another template and polymerization continues, creating an *in vitro* recombination product."
Odelberg **(EX1020, 1)**	"[R]ecombinants can be generated during a single round of primer extension in the absence of subsequent heat denaturation."

Dai **(EX1033, 8-9)**



PROCEEDINGS Open Access

NGSQC: cross-platform quality analysis pipeline for deep sequencing data

Illumina Exhibit 1033, Page 1 of 9

alternative splicing. A key issue in paired-end/mate pair data analysis is the false chimeric sequences, whose two ends are not from the original continuous sequence fragments during library preparation. NGSQC can gen-

In many ways, metagenetic ultrasequencing experiments are the ideal 'breeding ground' for recombinant DNA molecules. They are based on the amplification of homologous regions from a large number of potentially highly related organisms (von Wintzingerode *et al.* 1997; Qiu *et al.* 2001). The actual species richness and

Creer **(EX1034, 10)**



MOLECULAR ECOLOGY

Ultrasequencing of the meiofaunal biosphere: practice, pitfalls and promises

Illumina Exhibit 1034, Page 1 of 17

Multiplex Sequencing Was Particularly Vulnerable to Chimera Formation

Meyer (EX1015, 9)



Protocol

Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing

Matthias Meyer[1] and Martin Kircher

Illumina Exhibit 1015, Page 1 of 25

For sequencing and data analysis, use the recipes, kits, and analysis tools for multiplex sequencing provided by Illumina.

* * *

recommended to carry the blank library all the way through target capture and/or sequencing. To avoid cross-contamination of samples through jumping PCR (Meyerhans et al. 1990), pools of indexed libraries should be amplified with a minimum number of PCR cycles or sequenced without amplification if possible.

Gloor (EX1005, 1, 8)



Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products

Illumina Exhibit 1005, Page 1 of 28

Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products

* * *

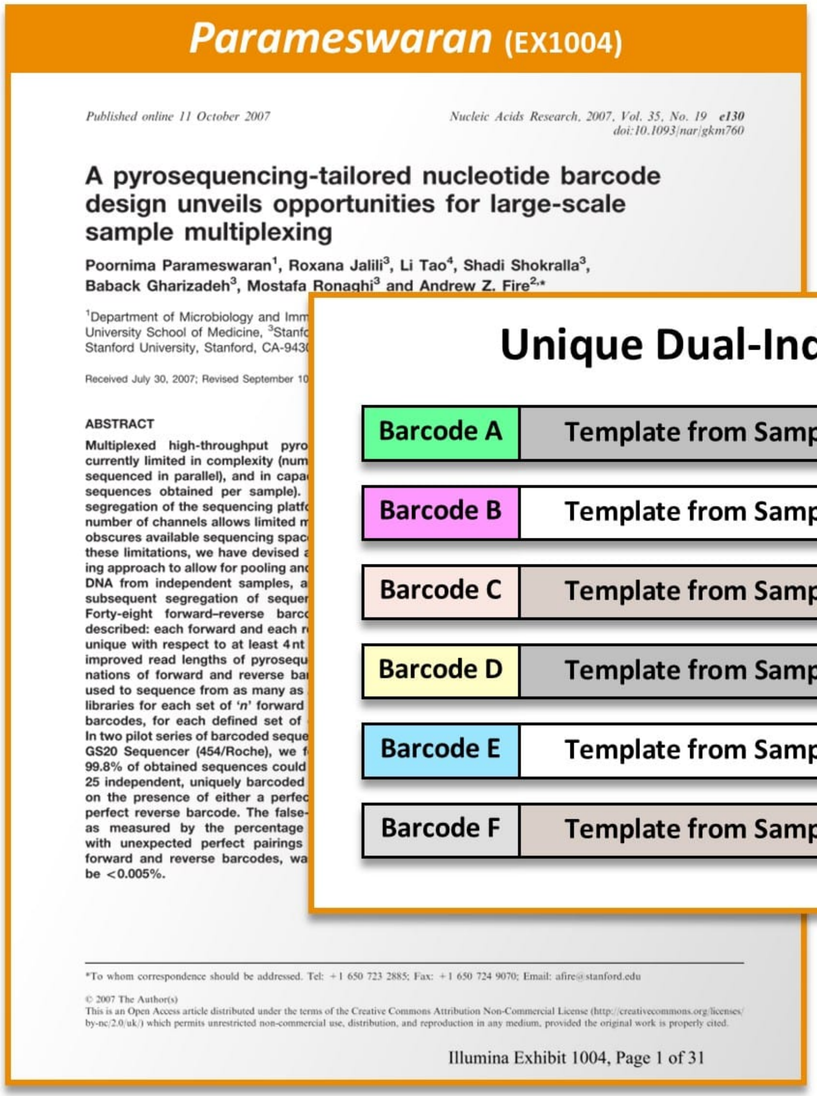
The occurrence of chimeric sequences was examined using UCHIME, a part of the UCLUST package (REF, http://www.drive5.com/uclust/). Chimeric sequences can arise during PCR [21] or because of overlapping spots on the solid support when imaged during DNA sequencing [6]. The default settings of

Lahr (EX1017, 1)	"Chimeras are formed when incompletely extended DNA fragments anneal to closely related sequences generating recombinants between starting templates."
Görzer (EX1060, 1)	"When PCR amplification is carried out using a mixture of two or more non-identical but similar genotype target sequences as templates, this can lead potentially to the formation of artificial chimeras between the amplification products."
Smyth (EX1063, 2)	"Unfortunately, highly similar sequences are the most likely to form chimeras during PCR, meaning the most commonly formed chimeras are the most difficult to computationally detect."

illumina

Dual-Indexing Was Used for Multiplex Sequencing

Parameswaran (EX1004)



Published online 11 October 2007

Nucleic Acids Research, 2007, Vol. 35, No. 19 e130
doi:10.1093/nar/gkm760

A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing

Poornima Parameswaran[1], Roxana Jalili[3], Li Tao[4], Shadi Shokralla[3], Baback Gharizadeh[3], Mostafa Ronaghi[3] and Andrew Z. Fire[2,*]

[1]Department of Microbiology and Imm
University School of Medicine, [3]Stanfo
Stanford University, Stanford, CA-943

Received July 30, 2007; Revised September 10

ABSTRACT

Multiplexed high-throughput pyro
currently limited in complexity (num
sequenced in parallel), and in capa
sequences obtained per sample).
segregation of the sequencing platfo
number of channels allows limited m
obscures available sequencing spac
these limitations, we have devised a
ing approach to allow for pooling an
DNA from independent samples, a
subsequent segregation of sequen
Forty-eight forward–reverse barc
described: each forward and each r
unique with respect to at least 4 nt
improved read lengths of pyrosequ
nations of forward and reverse ba
used to sequence from as many as
libraries for each set of '*n*' forward
barcodes, for each defined set of
In two pilot series of barcoded seque
GS20 Sequencer (454/Roche), we f
99.8% of obtained sequences could
25 independent, uniquely barcoded
on the presence of either a perfec
perfect reverse barcode. The false-
as measured by the percentage
with unexpected perfect pairings
forward and reverse barcodes, wa
be <0.005%.

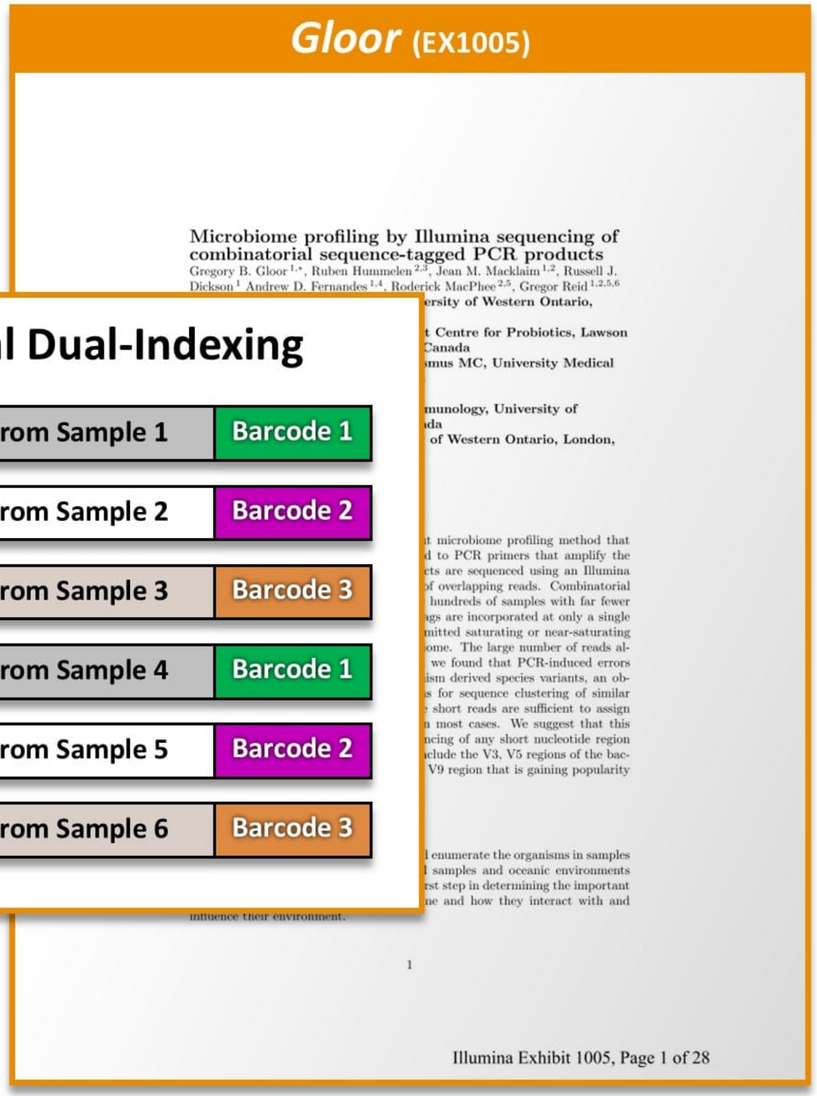
*To whom correspondence should be addressed. Tel: +1 650 723 2885; Fax: +1 650 724 9070; Email: afire@stanford.edu

© 2007 The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Unique Dual-Indexing

Barcode A	Template from Sample 1	Barcode 1
Barcode B	Template from Sample 2	Barcode 2
Barcode C	Template from Sample 3	Barcode 3
Barcode D	Template from Sample 4	Barcode 4
Barcode E	Template from Sample 5	Barcode 5
Barcode F	Template from Sample 6	Barcode 6

Gloor (EX1005)



Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products

Gregory B. Gloor[1,*], Ruben Hummelen[2,3], Jean M. Macklaim[1,2], Russell J. Dickson[1], Andrew D. Fernandes[1,4], Roderick MacPhee[2,5], Gregor Reid[1,2,5,6]

ersity of Western Ontario,

t Centre for Probiotics, Lawson
Canada
mus MC, University Medical

munology, University of
da
of Western Ontario, London,

t microbiome profiling method that
d to PCR primers that amplify the
cts are sequenced using an Illumina
of overlapping reads. Combinatorial
hundreds of samples with far fewer
gs are incorporated at only a single
mitted saturating or near-saturating
ome. The large number of reads al-
we found that PCR-induced errors
ism derived species variants, an ob-
s for sequence clustering of similar
short reads are sufficient to assign
most cases. We suggest that this
ncing of any short nucleotide region
clude the V3, V5 regions of the bac-
V9 region that is gaining popularity

l enumerate the organisms in samples
l samples and oceanic environments
rst step in determining the important
ne and how they interact with and
influence their environment.

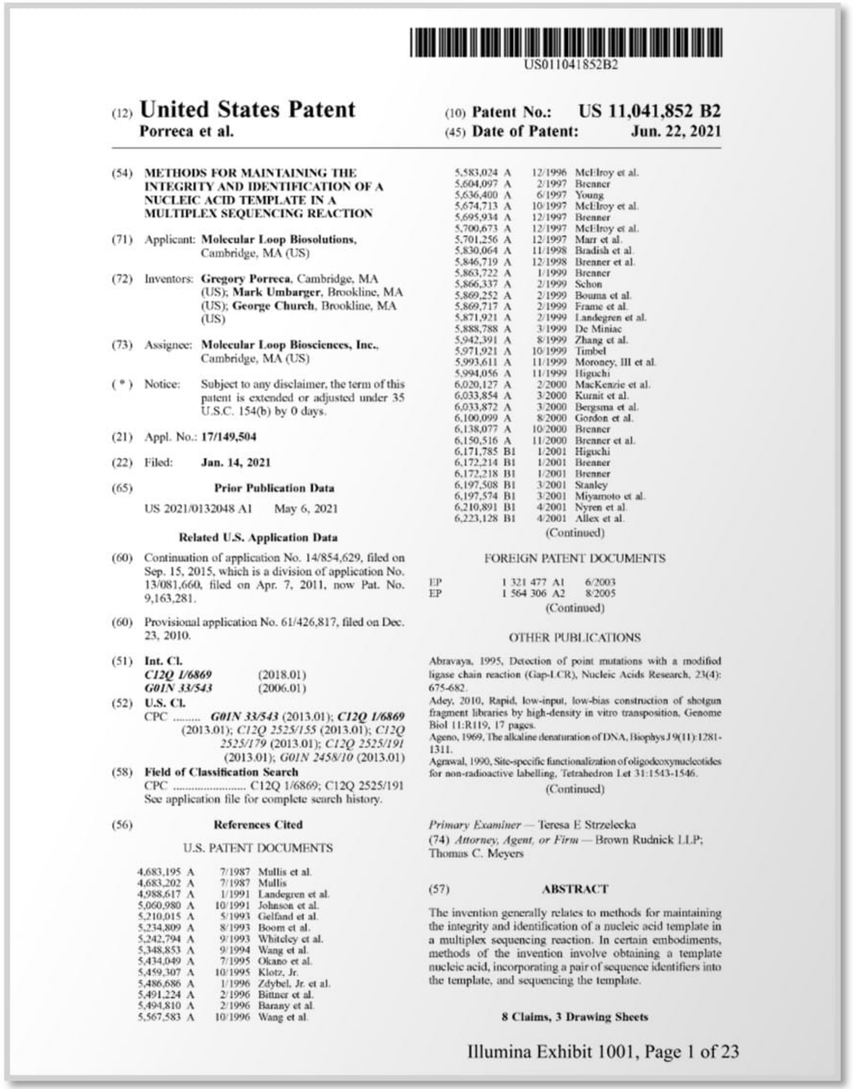
1

Combinatorial Dual-Indexing

Barcode A	Template from Sample 1	Barcode 1
Barcode A	Template from Sample 2	Barcode 2
Barcode A	Template from Sample 3	Barcode 3
Barcode B	Template from Sample 4	Barcode 1
Barcode B	Template from Sample 5	Barcode 2
Barcode B	Template from Sample 6	Barcode 3

illumina®

Ground 1: *Parameswaran* and *Gloor*



US011041852B2

(12) **United States Patent**
Porreca et al.

(10) **Patent No.: US 11,041,852 B2**
(45) **Date of Patent: Jun. 22, 2021**

(54) **METHODS FOR MAINTAINING THE INTEGRITY AND IDENTIFICATION OF A NUCLEIC ACID TEMPLATE IN A MULTIPLEX SEQUENCING REACTION**

(71) Applicant: **Molecular Loop Biosolutions**, Cambridge, MA (US)

(72) Inventors: **Gregory Porreca**, Cambridge, MA (US); **Mark Umbarger**, Brookline, MA (US); **George Church**, Brookline, MA (US)

(73) Assignee: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **17/149,504**

(22) Filed: **Jan. 14, 2021**

(65) **Prior Publication Data**

US 2021/0132048 A1 May 6, 2021

Related U.S. Application Data

(60) Continuation of application No. 14/854,629, filed on Sep. 15, 2015, which is a division of application No. 13/081,660, filed on Apr. 7, 2011, now Pat. No. 9,163,281.

(60) Provisional application No. 61/426,817, filed on Dec. 23, 2010.

(51) **Int. Cl.**
C12Q 1/6869 (2018.01)
G01N 33/543 (2006.01)

(52) **U.S. Cl.**
CPC ***G01N 33/543*** (2013.01); ***C12Q 1/6869*** (2013.01); *C12Q 2525/155* (2013.01); *C12Q 2525/179* (2013.01); *C12Q 2525/191* (2013.01); *G01N 2458/10* (2013.01)

(58) **Field of Classification Search**
CPC C12Q 1/6869; C12Q 2525/191
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,683,195 A 7/1987 Mullis et al.
4,683,202 A 7/1987 Mullis
4,988,617 A 1/1991 Landegren et al.
5,060,980 A 10/1991 Johnson et al.
5,210,015 A 5/1993 Gelfand et al.
5,234,809 A 8/1993 Boom et al.
5,242,794 A 9/1993 Whiteley et al.
5,348,853 A 9/1994 Wang et al.
5,434,049 A 7/1995 Okano et al.
5,459,307 A 10/1995 Klotz, Jr.
5,486,686 A 1/1996 Zdybel, Jr. et al.
5,491,224 A 2/1996 Bittner et al.
5,494,810 A 2/1996 Barany et al.
5,567,583 A 10/1996 Wang et al.
5,583,024 A 12/1996 McElroy et al.
5,604,097 A 2/1997 Brenner
5,636,400 A 6/1997 Young
5,674,713 A 10/1997 McElroy et al.
5,695,934 A 12/1997 Brenner
5,700,673 A 12/1997 McElroy et al.
5,701,256 A 12/1997 Marr et al.
5,830,064 A 11/1998 Bradish et al.
5,846,719 A 12/1998 Brenner et al.
5,863,722 A 1/1999 Brenner
5,866,337 A 2/1999 Schon
5,869,252 A 2/1999 Bouma et al.
5,869,717 A 2/1999 Frame et al.
5,871,921 A 2/1999 Landegren et al.
5,888,788 A 3/1999 De Mintac
5,942,391 A 8/1999 Zhang et al.
5,971,921 A 10/1999 Timbel
5,993,611 A 11/1999 Moroney, III et al.
5,994,056 A 11/1999 Higuchi
6,020,127 A 2/2000 MacKenzie et al.
6,033,854 A 3/2000 Kurait et al.
6,033,872 A 3/2000 Bergsma et al.
6,100,099 A 8/2000 Gordon et al.
6,138,077 A 10/2000 Brenner
6,150,516 A 11/2000 Brenner et al.
6,171,785 B1 1/2001 Higuchi
6,172,214 B1 1/2001 Brenner
6,172,218 B1 1/2001 Brenner
6,197,508 B1 3/2001 Stanley
6,197,574 B1 3/2001 Miyamoto et al.
6,210,891 B1 4/2001 Nyren et al.
6,223,128 B1 4/2001 Allex et al.
(Continued)

FOREIGN PATENT DOCUMENTS

EP 1 321 477 A1 6/2003
EP 1 564 306 A2 8/2005
(Continued)

OTHER PUBLICATIONS

Abravaya, 1995, Detection of point mutations with a modified ligase chain reaction (Gap-LCR), Nucleic Acids Research, 23(4): 675-682.
Adey, 2010, Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition, Genome Biol 11:R119, 17 pages.
Ageno, 1969, The alkaline denaturation of DNA, Biophys J 9(11):1281-1311.
Agrawal, 1990, Site-specific functionalization of oligodeoxynucleotides for non-radioactive labelling, Tetrahedron Let 31:1543-1546.
(Continued)

Primary Examiner — Teresa E Strzelecka
(74) *Attorney, Agent, or Firm* — Brown Rudnick LLP; Thomas C. Meyers

(57) **ABSTRACT**

The invention generally relates to methods for maintaining the integrity and identification of a nucleic acid template in a multiplex sequencing reaction. In certain embodiments, methods of the invention involve obtaining a template nucleic acid, incorporating a pair of sequence identifiers into the template, and sequencing the template.

8 Claims, 3 Drawing Sheets

Illumina Exhibit 1001, Page 1 of 23

Ground	References	Basis	Claims
1	*Parameswaran* (EX1004) and *Gloor* (EX1005)	§103	1-8
2	*Parameswaran* (EX1004) and *Bentley* (EX1006)	§103	1-8

illumina

Claim-Construction: "Cross-Over Error"

1 [pre] A method for reducing template cross-over error introduced during sequencing workflow, the method comprising the steps of

[a] incorporating at least two members of a plurality of identifier sequences into template nucleic acids obtained from at least two different samples, wherein said two members constitute a distinct pair associated with said template;

[b] combining said nucleic acid templates into a single sample;

[c] amplifying said nucleic acid templates on a surface of a flow cell thereby to form clusters,

[d] at least one of which comprises a chimeric sequence comprising a combination of identifier sequences that are different than any of said distinct pairs;

[e] sequencing amplicons obtained in said amplifying step; and

[f] **discarding** sequence reads obtained in said sequencing step that contain said **combination of identifier sequences that are different than any of said distinct pairs**, thereby to reduce cross-over error.



Petitioner's Expert, Dr. Spellman (EX1003, ¶¶122-123)

"Cross-over error" means: "**sequences of chimeric molecules made from template nucleic acids derived from different samples.**"

"The '852 patent uses the term 'crossover' to describe the formation of chimeric molecules that contain a combination of two template nucleic acids derived from different samples. *See* EX1001 ('852 patent), 5:41-6:6 (stating that 'a crossover has occurred' when a 'chimeric molecule formed' between template nucleic acids from patient A and patient B due to 'improper molecular overlap')."

illumina

EX1001 ('852 Pat.), cl. 1; EX1003, ¶¶122-123; Pet., 22-23.

Patent Owner's Narrow Construction Contradicts the Claim Language

1 [pre]	A method for reducing template cross-over error introduced during sequencing workflow, the method comprising the steps of
[a]	**incorporating** at least two members of a plurality of identifier sequences into template nucleic acids obtained from at least two different samples, wherein said two members constitute **a distinct pair** associated with said template;
[b]	**combining** said nucleic acid templates into a single sample;
[c]	**amplifying** said nucleic acid templates on **a surface of a flow cell thereby to form clusters,**
[d]	**at least one of which comprises a chimeric sequence** comprising a combination of identifier sequences that are different than any of said distinct pairs;
[e]	**sequencing** amplicons obtained in said amplifying step; and
[f]	**discarding** sequence reads obtained in said sequencing step that contain said **combination of identifier sequences that are different than any of said distinct pairs,** thereby to reduce cross-over error.

Patent Owner Response (POR, 13)

"The Board should construe these terms to mean '**error from** sequences of chimeric molecules **created during cluster amplification and** made from template nucleic acids derived from different samples.'"

Does Not Require Chimeric Sequences to Be "Created" During Cluster Amplification (Reply, 3-4)

Sisvel Int'l S.A. v. Sierra Wireless, Inc.,
81 F.4th 1231, 1236-37 (Fed. Cir. 2023)

"[T]he intrinsic evidence provides no persuasive basis to limit the claims," especially when "the claim language itself is not so limited."

illumina

Patent Owner's Narrow Construction Contradicts the Intrinsic Evidence

1 [pre]	A method for reducing template cross-over error introduced during sequencing workflow, the method comprising the steps of
[a]	**incorporating** at least two members of a plurality of identifier sequences into template nucleic acids obtained from at least two different samples, wherein said two members constitute **a distinct pair** associated with said template;
[b]	**combining** said nucleic acid templates into a single sample;
[c]	**amplifying** said nucleic acid templates on **a surface of a flow cell thereby to form clusters,**
[d]	**at least one of which comprises a chimeric sequence** comprising a combination of identifier sequences that are different than any of said distinct pairs;
[e]	**sequencing** amplicons obtained in said amplifying step; and
[f]	**discarding** sequence reads obtained in said sequencing step that contain said **combination of identifier sequences that are different than any of said distinct pairs,** thereby to reduce cross-over error.

Hill-Rom Servs., Inc. v. Stryker Corp.,
755 F.3d 1367, 1371 (Fed. Cir. 2014)

"[W]e do not read limitations from the embodiments in the specification into the claims."

'852 Patent **(EX1001, 6:25-29)**

primers at discrete locations on a flow cell. Without being limited by any particular theory or mechanism of action, the data in FIG. **3** suggest that the nucleic acid from sample A and the nucleic acid from sample B landed within a cluster radius on the solid support. When the amplification process

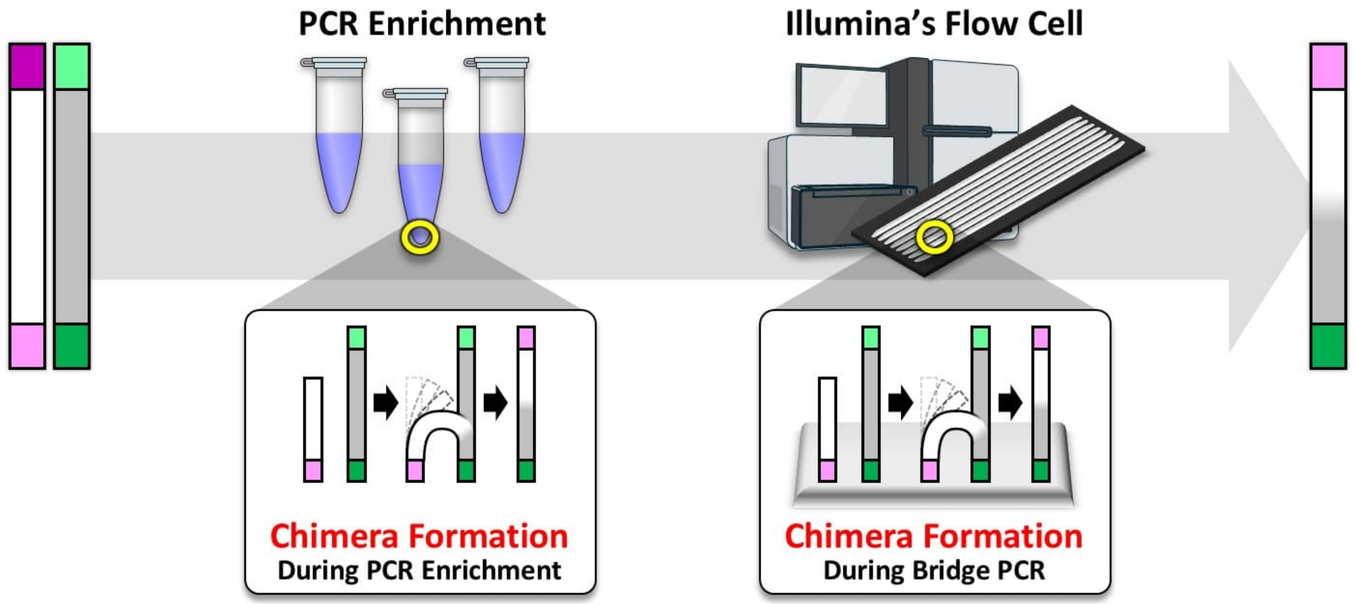
'852 Patent Prosecution **(EX1007, 305)**

inherent in pyrosequencing. *See*, Parameswaran at page 4. However, Parameswaran does not mention forming clusters in a flow cell. That is not at all surprising given that Parameswaran is all about pyrosequencing. In pyrosequencing, one prepares a series of bead-bound templates and ultimately monitors the liberation of diphosphates as and indication of base incorporation during sequencing.

illumina

Patent Owner's Narrow Construction Contradicts the Claimed Method

1 [pre]	A method for reducing template cross-over error introduced during sequencing workflow, the method comprising the steps of
[a]	**incorporating** at least two members of a plurality of identifier sequences into template nucleic acids obtained from at least two different samples, wherein said two members constitute **a distinct pair** associated with said template;
[b]	**combining** said nucleic acid templates into a single sample;
[c]	**amplifying** said nucleic acid templates on **a surface of a flow cell thereby to form clusters,**
[d]	**at least one of which comprises a chimeric sequence** comprising a combination of identifier sequences that are different than any of said distinct pairs;
[e]	**sequencing** amplicons obtained in said amplifying step; and
[f]	**discarding** sequence reads obtained in said sequencing step that contain said **combination of identifier sequences that are different than any of said distinct pairs,** thereby to reduce cross-over error.





Petitioner's Expert, Dr. Spellman (EX1054, ¶24)

"[T]he '852 patent ***does not teach*** how to selectively reduce cross-over errors 'created during cluster amplification' instead of cross-over errors created at other points during the sequencing workflow[.]"

illumina

Parameswaran and *Gloor* Disclose Each Element of Claim 1

1 [pre] A method for reducing template cross-over error introduced during sequencing workflow, the method comprising the steps of

[a] **incorporating** at least two members of a plurality of identifier sequences into template nucleic acids obtained from at least two different samples, wherein said two members constitute **a distinct pair** associated with said template;

[b] **combining** said nucleic acid templates into a single sample;

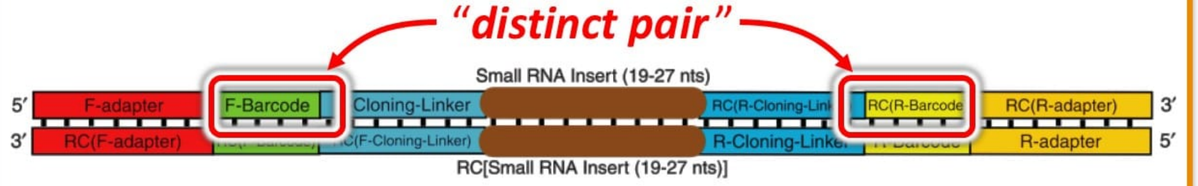
[c] amplifying said nucleic acid templates on a surface of a flow cell thereby to form clusters,

[d] at least one of which comprises a chimeric sequence comprising a combination of identifier sequences that are different than any of said distinct pairs;

[e] sequencing amplicons obtained in said amplifying step; and

[f] discarding sequence reads obtained in said sequencing step that contain said combination of identifier sequences that are different than any of said distinct pairs, thereby to reduce cross-over error.

Parameswaran **(EX1004, 3, 7)**



dent datasets. These barcode + linker sequences are referred to herein as 'search motifs' or 'barcode motifs'.

* * *

DNA from two independent sets of 25 barcoded libraries (prepared with 19–27 nts RNAs from different *in vitro* and *in vivo* virus-infected systems) were pooled for sequencing

illumina

Parameswaran and *Gloor* Disclose Each Element of Claim 1

1 [pre] A method for reducing template cross-over error introduced during sequencing workflow, the method comprising the steps of

[a] **incorporating** at least two members of a plurality of identifier sequences into template nucleic acids obtained from at least two different samples, wherein said two members constitute a distinct pair associated with said template;

[b] **combining** said nucleic acid templates into a single sample;

[c] **amplifying** said nucleic acid templates on a surface of a flow cell thereby to form clusters,

[d] at least one of which comprises a chimeric sequence comprising a combination of identifier sequences that are different than any of said distinct pairs;

[e] **sequencing** amplicons obtained in said amplifying step; and

[f] **discarding** sequence reads obtained in said sequencing step that contain said combination of identifier sequences that are different than any of said distinct pairs, thereby to reduce cross-over error.

Parameswaran **(EX1004, 4)**

Processing of pooled sample for high-throughput sequencing

Amplicon libraries prepared as described above may be entered directly into the 454/Roche sequencing pipeline, with the only major variable being the amount of DNA from the pooled sample used to seed the emulsion PCR reaction. Approximately 50 ng of the purified pool was set aside for each pyrosequencing run. For run I, DNA

illumina

EX1001 ('852 Pat.), cl. 1; EX1004, 1-2, 4, 8, 10-11; EX1005, 5-7, 22-24; EX1003, ¶¶143-146, 150-151; Pet., 31-34.

Parameswaran and *Gloor* Disclose Each Element of Claim 1

1 [pre] A method for reducing template cross-over error introduced during sequencing workflow, the method comprising the steps of

[a] incorporating at least two members of a plurality of identifier sequences into template nucleic acids obtained from at least two different samples, wherein said two members constitute a distinct pair associated with said template;

[b] combining said nucleic acid templates into a single sample;

[c] amplifying said nucleic acid templates on a surface of a flow cell thereby to form clusters,

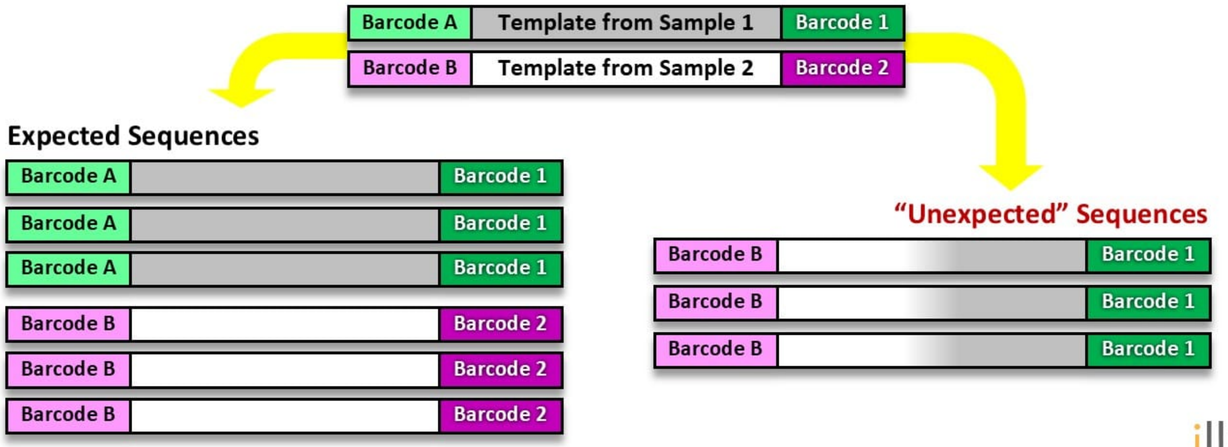
[d] at least one of which comprises a chimeric sequence comprising a combination of identifier sequences that are different than any of said distinct pairs;

[e] sequencing amplicons obtained in said amplifying step; and

[f] discarding sequence reads obtained in said sequencing step that contain said combination of identifier sequences that are different than any of said distinct pairs, thereby to reduce cross-over error.

Parameswaran (EX1004, 7)

Sequences that have a perfect match to only one of the expected motifs have been further subdivided into three classes based on whether the motif at the other end has: (i) perfect identity to an unexpected motif, resulting in an unassigned 5′–3′ motif combination (perfect, unexpected occurrence), (ii) a high-scoring partial match to the expected motif for that library (imperfect, expected occurrence) and (iii) no strong match to any barcode motif (imperfect, ambiguous occurrence). Perfect, 'unexpected'



illumina

Parameswaran and *Gloor* Disclose Each Element of Claim 1

1 [pre] A method for reducing template cross-over error introduced during sequencing workflow, the method comprising the steps of

[a] **incorporating** at least two members of a plurality of identifier sequences into template nucleic acids obtained from at least two different samples, wherein said two members constitute a **distinct pair** associated with said template;

[b] **combining** said nucleic acid templates into a single sample;

[c] **amplifying** said nucleic acid templates on a surface of a flow cell thereby to form clusters,

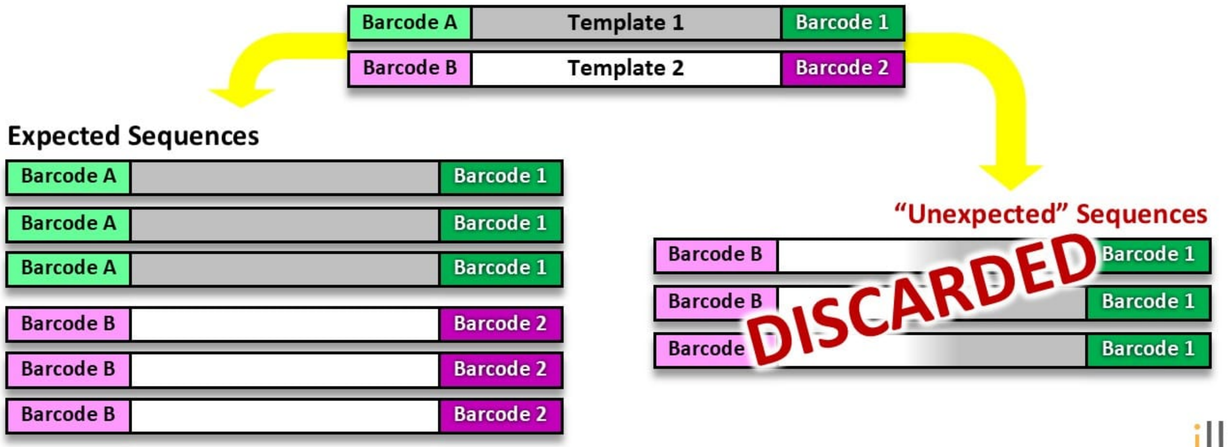
[d] at least one of which comprises a chimeric sequence comprising a combination of identifier sequences that are different than any of said distinct pairs;

[e] sequencing amplicons obtained in said amplifying step; and

[f] **discarding** sequence reads obtained in said sequencing step that contain said **combination of identifier sequences that are different than any of said distinct pairs**, thereby to reduce cross-over error.

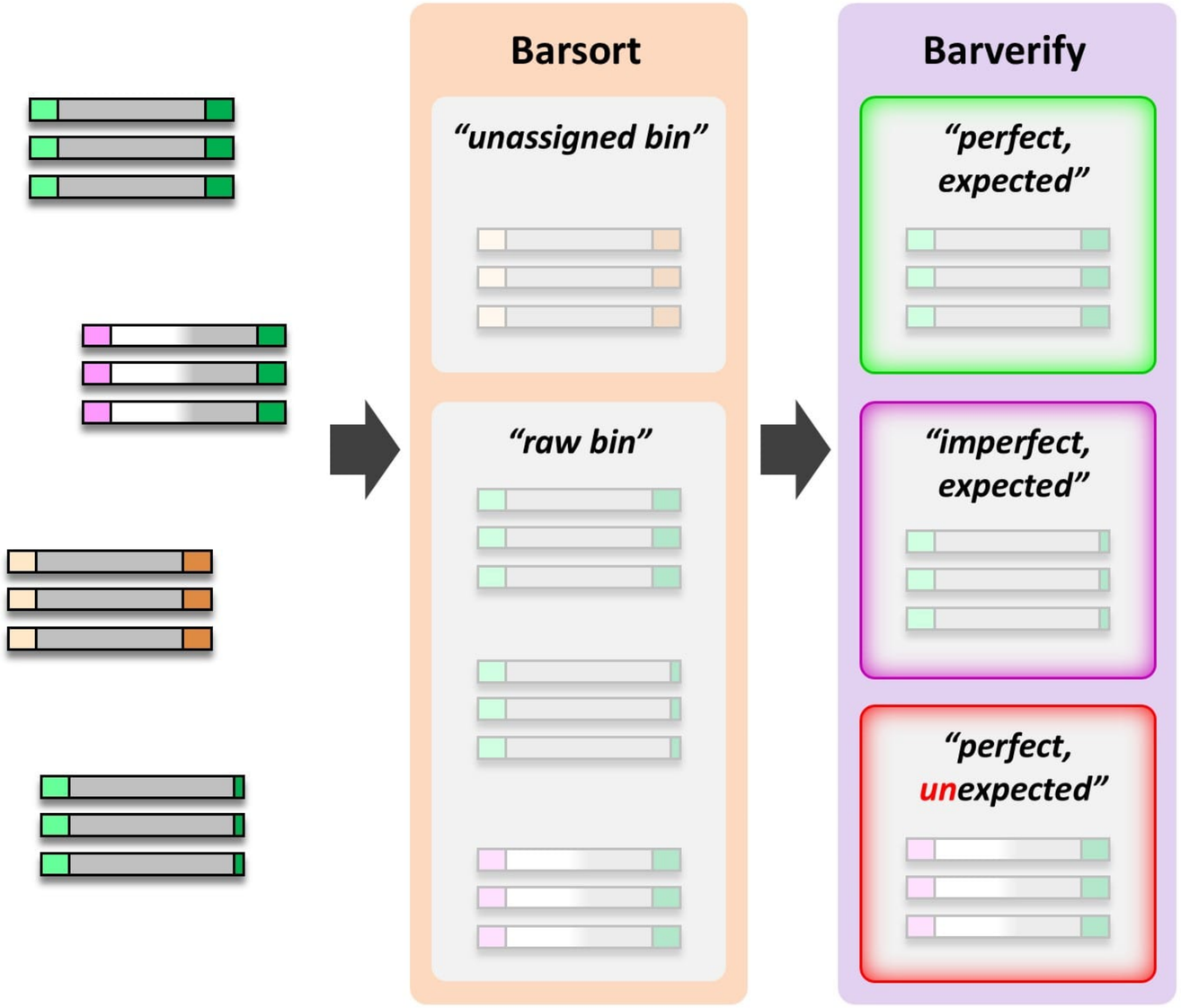
Parameswaran **(EX1004, 8)**

The false-discovery rate, defined by the percentage of 'perfect and unexpected' occurrences, was observed to decrease to <0.005% (Table 3), if three bases of the linker were used in addition to the 10-base barcode to assign sample identities. Thus a search motif of 13 nt



illumina

Parameswaran and *Gloor* Disclose Each Element of Claim 1



D. SEARCH MOTIF = Barcode (10 nt) + Cloning-Linker (3 nt)

	454-66	454-67	454-68	454-69	454-70	454-71	454-72
454-66	13566	0	0	1	0	0	0
454-67	0	7459	0	3	0	0	0
454-68	0	0	13260	0	0	0	0
454-69	0	0	0	6518	0	0	0
454-70	0	0	0	0	9484	0	0
454-71	0	0	0	0	0	9397	0
454-72	0	0	0	0	0	0	12302
			*	*	*		
454-90	0	0	0	0	0	0	0
(close)	369	167	210	122	237	154	317
(no match)	37	26	76	48	28	20	48

illumina

Parameswaran and *Gloor* Disclose Each Element of Claim 1

1 [pre] A method for reducing template cross-over error introduced during sequencing workflow, the method comprising the steps of

[a] **incorporating** at least two members of a plurality of identifier sequences into template nucleic acids obtained from at least two different samples, wherein said two members constitute a **distinct pair** associated with said template;

[b] **combining** said nucleic acid templates into a single sample;

[c] **amplifying** said nucleic acid templates on a surface of a flow cell thereby to form clusters,

[d] at least one of which comprises a chimeric sequence comprising a combination of identifier sequences that are different than any of said distinct pairs;

[e] sequencing amplicons obtained in said amplifying step; and

[f] **discarding** sequence reads obtained in said sequencing step that contain said **combination of identifier sequences that are different than any of said distinct pairs**, thereby to reduce cross-over error.

Parameswaran (EX1004, 29)

D. SEARCH MOTIF = Barcode (10 nt) + Cloning-Linker (3 nt)

	454-66	454-67	454-68	454-69	454-70	454-71	454-72	454-73	454-74	454-75	454-76	454-77
454-66	13566	0	0	1	0	0	0	0	0	0	0	0
454-67	0	7459	0	3	0	0	0	0	0	0	0	0
454-68	0	0	13260	0	0	0	0	0	0	0	0	0
454-69	0	0	0	6518	0	0	0	0	0	0	0	0
454-70	0	0	0	0	9484	0	0	0	0	0	0	0
454-71	0	0	0	0	0	9397	0	0	0	0	0	0
454-72	0	0	0	0	0	0	12302	0	0	0	0	0
454-73	0	0	0	0	0	0	0	7273	0	0	0	0
454-74	0	0	0	0	0	0	0	0	13923	0	0	0
454-75	0	0	0	0	0	0	0	0	1	15060	0	0
454-76	0	0	0	0	0	0	0	0	1	0	10266	0
454-77	0	0	0	0	0	0	0	0	0	0	0	9602

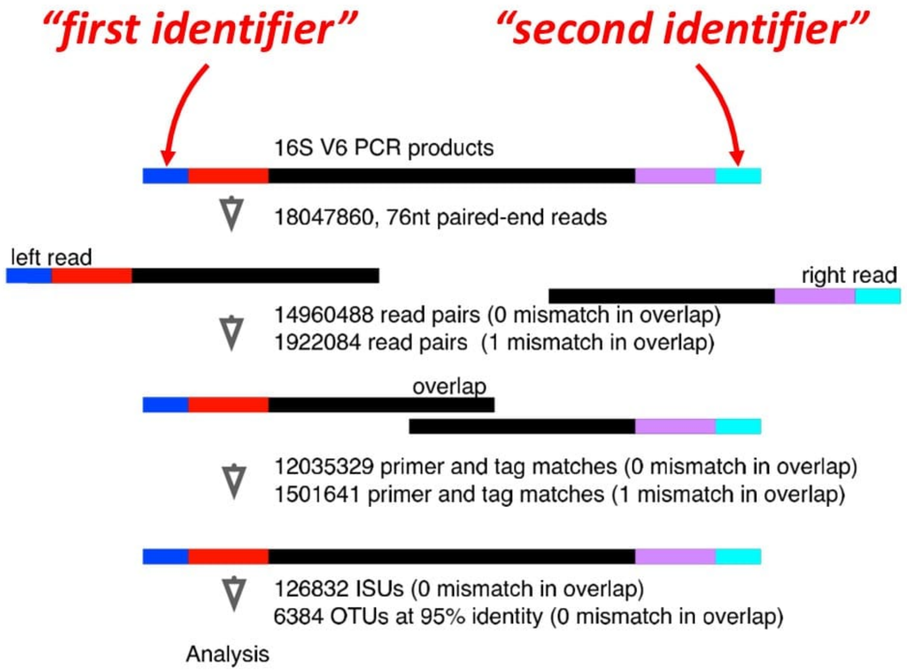
"perfect, unexpected occurrence" **cannot be assigned to any one sample**

"perfect, expected occurrence" **can be assigned to samples**

illumina

Parameswaran and *Gloor* Disclose Each Element of Claim 1

Gloor's Sequencing Workflow (EX1005, 7)



Gloor's Dual-Indexing Method (EX1005, 1)

We developed a low-cost, high-throughput microbiome profiling method that uses combinatorial sequence tags attached to PCR primers that amplify the rRNA V6 region. Amplified PCR products are sequenced using an Illumina paired-end protocol to generate millions of overlapping reads. Combinatorial

Gloor's Chimeric Sequences (EX1005, 8)

The occurrence of chimeric sequences was examined using UCHIME, a part of the UCLUST package (REF, http://www.drive5.com/uclust/). Chimeric

illumina

A POSA Would Have Been Motivated to Combine *Parameswaran* and *Gloor*

Illumina Sequencers Were Faster, More Efficient, and Less Expensive (Pet., 36-37)

Kircher **(EX1010, 10)**

	Throughput	Length	Quality	Costs
454/Roche	750 Mb/day	400 nt	10^{-3}–10^{-4}	~20$/Mb
Illumina	5,000 Mb/day	100 nt	10^{-2}–10^{-3}	~0.50$/Mb

To Reduce Sequencing Error, Including Chimeras Formed During Multiplex Sequencing (Pet., 37-38)

Gloor **(EX1005, 2, 8)**

Importantly, we found that the Illumina sequencing method has an exceedingly low error rate and that the majority of errors arise during the PCR amplification step. We argue that the error profile has profound implications for

* * *

of the UCLUST package (REF, http://www.drive5.com/uclust/). Chimeric sequences can arise during PCR [21] or because of overlapping spots on the solid support when imaged during DNA sequencing [6]. The default settings of

DyStar Textilfarben GmbH & Co. v. C.H. Patrick Co., **464 F.3d 1356, 1368 (Fed. Cir. 2006)**

Motivation to combine exists when "the combination of references results in a product or process that is more desirable," for example because it is "cheaper" or "more efficient."

illumina

A POSA Would Have Been Motivated to Combine *Parameswaran* and *Gloor*

Parameswaran's Recommendation (Pet., 38-40)

Parameswaran (EX1004, 2, 8)

runs. These barcodes may also be compatible with other sequencing platforms. Optimally, however the precise

* * *

accelerated pace. Three major high-throughput sequencing platforms are in use today: the Genome Sequencers from Roche/454 Life Sciences [GS-20 or GS-FLX; (1)], the 1G Analyzer from Illumina/Solexa (2) and the SOLiD System from Applied Biosystems (http://solid.appliedbiosystems.com). Comparison across the three platforms

Acknowledgement in the '852 Patent (Pet., 40-41)

'852 Patent (EX1001, 20:37-46)

The invention is not intended to be limited to use of the sequencing method outlined above, as essentially any sequencing methodology which relies on successive incorporation of nucleotides into a polynucleotide chain can be used. Suitable alternative techniques include, for example, the Genome Sequencers from Roche/454 Life Sciences (Margulies et al. (2005) Nature, 437:376-380; U.S. Pat. Nos. 6,274,320; 6,258,568; 6,210,891), and the SOLiD system from Applied Biosystems (solid.appliedbiosystems.com), and the sequencer from Ion Torrent (www.iontorrent.com).

KSR Int'l Co. v. Teleflex, Inc., 550 U.S. 398, 417 (2007)

"[I]f a technique has been used to improve one device, and a person of ordinary skill in the art would recognize that it would improve similar devices in the same way, using the technique is obvious unless its actual application is beyond his or her skill."

illumina

A POSA Would Have Had a Reasonable Expectation of Success

Interchangeability Between Roche Pyrosequencing and Illumina Sequencing (Pet., 41-43)

Meyer (EX1015, 1)



Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing

Matthias Meyer* and Martin Kircher

Illumina Exhibit 1015, Page 1 of 25

al. 2005). By exchanging adapter sequences, removing and shortening several reaction steps, and introducing an amplification scheme, the protocol has been redesigned for rapid preparation of Illumina multiplex sequencing libraries using a 96-well plate format. In the example shown in Figure

Maughan (EX1038, 1)



Single-Nucleotide Polymorphism Genotyping in Mapping Populations via Genomic Reduction and Next-Generation Sequencing: Proof of Concept

Peter J. Maughan,* Scott M. Yourstone, Robert J. Byers, Scott M. Smith, and Joshua A. Udall

Illumina Exhibit 1038, Page 1 of 13

We report the results of a proof-of-concept experiment that validates the use of multiplex identifier (MID) barcodes and next-generation sequencing (454-pyrosequencing and Illumina GAIIe) to simultaneously discover and genotype single nucleotide polymorphisms (SNPs) from mapping populations using pooled genomic reduction libraries. The genomic reduction library

Increase in Illumina Sequencing Read Length (Pet., 44)

New Kit Helps Paint Truest Picture of Genome With Longer Reads and Enhanced Data Quality

SAN DIEGO, Jul 26, 2010 (BUSINESS WIRE) --

Illumina, Inc. (NASDAQ:ILMN) today announced the launch of its TruSeq SBS Sequencing kit that will enable high-quality paired end reads of 150 base pairs (bp), increase the output of high-quality base

illumina

A POSA Would Have Had a Reasonable Expectation of Success

Barcode Design Was Routine, As Acknowledged in the '852 Patent (Pet., 43-44)

'852 Patent **(EX1001, 3:44-48, 8:65-9:4)**

Numerous techniques are known for detecting sequence and for identifying barcodes and some are exemplified below. However, the exact means for detecting and compiling sequence data does not affect the function of the invention described herein.

* * *

Exemplary methods for designing sets of barcode sequences and other methods for attaching barcode sequences are shown in U.S. Pat. Nos. 6,138,077; 6,352,828;



Petitioner's Expert, Dr. Spellman (EX1003, ¶172)

"Like pyrosequencing, Illumina sequencing uses PCR to amplify template nucleic acids before sequencing, and PCR amplification was a step known to result in chimeric sequences and reduce sequencing accuracy."

"Because *Parameswaran*'s unique dual-indexing method was not limited to any sequencing method or platform, a POSA would have reasonably expected it to reduce cross-over error on an Illumina sequencer."

Medtronic, Inc. v. Marital Deduction Trust

IPR2014-00100, Paper 46 at 25 (PTAB Mar. 24, 2015)

A POSA would have had a reasonable expectation of success in "substitution of a similar component for another in a similar device to yield a result one would have predicted from reading the references."

illumina

Contemporaneous Researchers Praised *Parameswaran*

Fuller (EX1059, 1-3, Table 1)

nature biotech

The challenges of sequencing by synthesis

Carl W Fuller[1], Lyle R Middendorf[2], Steven A Benner[3], George M Church[4], Timothy Harris[5], Xiaohua Huang[6], Stevan B Jovanovich[7], John R Nelson[8], Jeffery A Schloss[9], David C Schwartz[10] & Dmitri V Vezenov[11]

* * *

Widespread clinical applications will likely require streamlined, robust sample preparation techniques for supporting typical clinical loads. A 2-day sample preparation time by itself may restrict clinical applications and the ultrahigh throughput of gigabases of data per run must be matched to clinical needs, which may be for only thousands of bases of sequence per patient. One option would be an effective technology to judiciously add sequence tags to the fragments of several different samples, then combine and process the DNA samples (encoding and binning) together, separating the resulting sequences by their tags in the computer analysis phase[60]. This reduces the inefficiencies of ensemble-based SBS architectures (e.g., when abundant DNA is available to be sequenced as is the case when the patient is at hand). Any approach that directly

* * *

60. Parameswaran, P. *et al.* A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Res.* **35**, e130 (2007).

* * *

NATURE BIOTECHNOLOGY VOLUME 27 NUMBER 11 NOVEMBER 2009

© 2009 Nature America, Inc. All rights reserved.

Illumina Exhibit 1059, Page 1 of 11
Illumina v. Molecular Loop Biosciences - IPR2024-00964

Figure 1 Schemes for SBS. Approaches include both single-molecule and ensemble methods requiring features populated with identical copies of DNA fragments. Ensemble methods monitor synthesis either using labeled substrates or detecting products of synthesis (pyrophosphate). Detection can be discontinuous (sequential) when labels can remain fixed for hours while a microscope scans up to hundreds of millions of sequence features[26,32]. It

* * *

26. Bentley, D.R. *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53–59 (2008).

illumina

Patent Owner Relies on Its Erroneous Construction

Patent Owner Response & Sur-Reply (POR, 29-36; Sur-Reply, 8-15)

"Chimeric sequences arising in Illumina sequencing were not a known problem."

* * *

Chimeric sequences **created** "during cluster amplification" was an "unrecognized" problem.



Petitioner's Expert, Dr. Spellman (EX1003, ¶122)

"Cross-over error" means: "**sequences of chimeric molecules made from template nucleic acids derived from different samples.**"

Patent Owner's Expert, Dr. Adalsteinsson (EX2001, ¶24)



"Cross-over error" means: "**error from sequences of chimeric molecules created during cluster amplification and made from template nucleic acids derived from different samples.**"

Chimeric Molecules "Created During Cluster Amplification" Were a Known Problem

Gloor **(EX1005, 8)**

The occurrence of chimeric sequences was examined using UCHIME, a part of the UCLUST package (REF, http://www.drive5.com/uclust/). Chimeric sequences can arise during PCR [21] or because of overlapping spots on the solid support when imaged during DNA sequencing [6]. The default settings of

'852 Patent **(EX1001, 6:25-29)**

primers at discrete locations on a flow cell. Without being limited by any particular theory or mechanism of action, the data in FIG. **3** suggest that the nucleic acid from sample A and the nucleic acid from sample B landed within a cluster radius on the solid support. When the amplification process



Petitioner's Expert, Dr. Spellman **(EX1054, ¶48)**

Incomplete PCR "on the surface of a flow cell would have been particularly problematic when **different templates were seeded nearby on the flow cell** . . . Incompletely extended templates were known to create an ideal opportunity for the formation of chimeric molecules."

illumina

Chimeric Molecules "Created During Cluster Amplification" Were a Known Problem

Campbell (EX1057, 2)

ARTICLES

nature genetics

Identification of somatically acquired rearrangements

Rearrangements

There were ~1.2 million reads from NCI-H2171 and 770,000 from NCI-H1770 in which mapping of the two ends to the genome suggested a possible rearrangement. Most of these were presumably

* * *

Using this strategy, we identified a total of 325 rearrangements in NCI-H2171 (**Table 1, Fig. 2a, Supplementary Note** and **Supplementary Table 2** online) and 84 in NCI-H1770 (Table 1, Fig. 2b,

be more frequent than previously thought in epithelial cancers as well[2], a prediction substantiated by recent discoveries of an *EML4-ALK* fusion in non-small-cell lung cancer[3] and *ETS* fusion genes in prostate cancer[4,5].

The relatively late discovery of commonly rearranged genes associated with cancer in solid tumors reflects the difficulty of systematically characterizing rearrangements in highly rearranged cancer genomes. G-banded cytogenetics, spectral karyotyping and FISH

out genuinely genome-wide screening for point mutations, copy number changes and rearrangements on a single platform. Here, we use a massively parallel technology—incorporating a procedure in which short reads are generated from both ends of millions of DNA fragments up to 500 bp in size—to systematically detect and characterize rearrangements. Using this strategy, we explore and compare patterns of somatic and germline rearrangement in human cancer genomes.

[1]Wellcome Trust Sanger Institute, Hinxton CB10 1SA, UK. [2]Hutchison/Medical Research Council Research Centre and Department of Pathology, University of Cambridge, Hills Road, Cambridge CB2 0XZ, UK. [3]Department of Neurosurgery, University of Cambridge, Hills Road, Cambridge CB2 2QQ, UK. [4]These authors contributed equally to this work. Correspondence should be addressed to M.R.S. (mrs@sanger.ac.uk) or P.A.F. (paf@sanger.ac.uk).

Received 8 January; accepted 27 February; published online 27 April 2008; doi:10.1038/ng.128

722 VOLUME 40 | NUMBER 6 | JUNE 2008 NATURE GENETICS

Illumina Exhibit 1057, Page 1 of 8
Illumina v. Molecular Loop Biosciences - IPR2024-00964



Petitioner's Expert, Dr. Spellman (EX1054, ¶¶41-46)

"*Campbell* demonstrates that . . . **the amount of artificial chimeric sequences generated during the Illumina sequencing workflow was high** (nearly 1.2 million chimeric sequences)."

"A POSA would have understood that **many of these chimeric sequences arose** from chimeric molecules created **during cluster amplification** on Illumina's flow cells."

illumina

Chimeric Molecules "Created During Cluster Amplification" Were a Known Problem

Patent Owner Sur-Reply (Sur-Reply, 9)

"Campbell never once mentions a chimera or a chimeric sequence."



Petitioner's Expert, Dr. Spellman (EX2015, 66:3-8)

Q. What's the difference between a rearrangement and a chimeric sequence?

A. Technically nothing.

Patent Owner Sur-Reply (Sur-Reply, 10)

"Petitioner's read of Campbell as describing rampant chimera errors on Illumina platforms also conflicts with a linchpin of Petitioner's obviousness rationale: that Illumina sequencers were much better and more accurate, allegedly motivating their use in any sequencing application."



Petitioner's Expert, Dr. Spellman (EX2015, 60:13-16)

Q. In December 2010, sequencing accuracy of the Illumina machines was extremely high. That's your opinion, right?

A. On individual-base[] determinations, yes.

illumina

EX1054, ¶¶41-46; EX2015, 66:3-8, 60:13-16; Reply, 7-10; Sur-reply, 9-10.

Chimeric Molecules "Created During Cluster Amplification" Were a Known Problem

Patent Owner Sur-Reply (Sur-Reply, 14)

EX2004 (published in 2011) "remains the earliest evidence of anyone other than the inventors recognizing the problem of cross-over error on Illumina sequencers."

Kircher (EX2004, 2)

Published online 21 October 2011 *Nucleic Acids Research, 2012, Vol. 40, No. 1 e3*
doi:10.1093/nar/gkr771

Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform

Martin Kircher*, Susanna Sawyer and Matthias Meyer*

* * *

Unfortunately, PCR can produce chimeras by recombining different templates molecules (a process often referred to as 'jumping PCR') (17–20). Consequently, multiplex capture may introduce significant levels of sample cross-contamination.

Ref. 17 = EX1018 (Meyerhans)	**Ref. 18 = EX1019 (Paabo)**
Ref. 19 = EX1020 (Odelberg)	**Ref. 20 = EX1017 (Lahr)**

Meyer (EX1015, 1, 9)

© 2010 Cold Spring Harbor Laboratory Press Vol. 2010, Issue 6, June

* * *

Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing

Matthias Meyer[1] and Martin Kircher

* * *

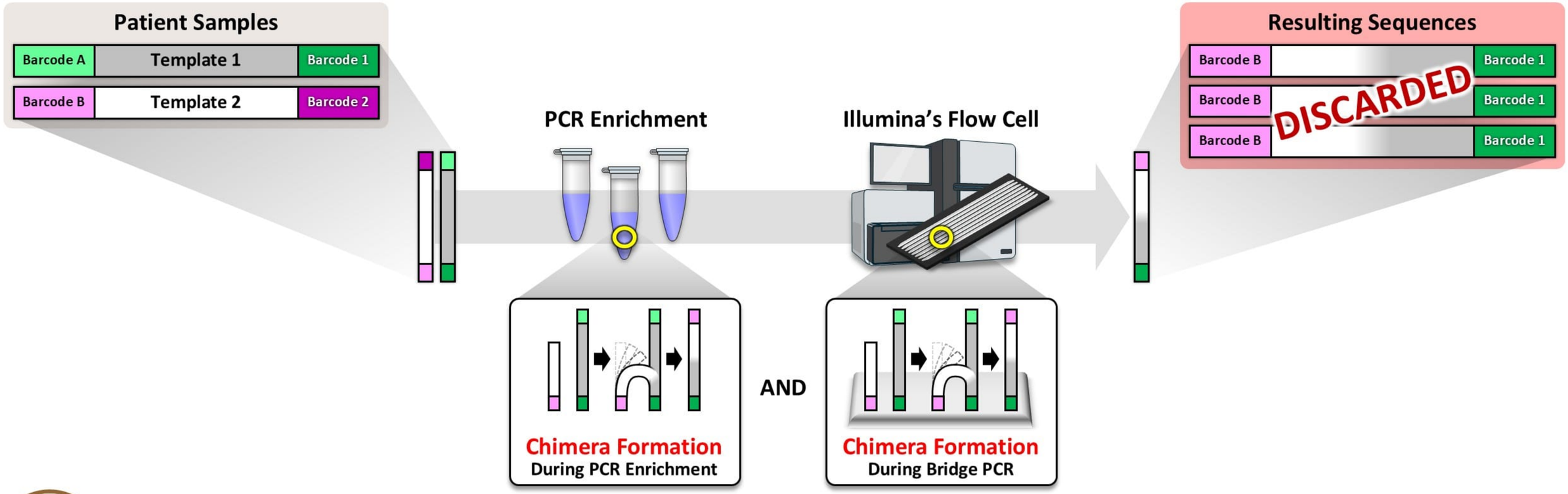
recommended to carry the blank library all the way through target capture and/or sequencing. To avoid cross-contamination of samples through jumping PCR (Meyerhans et al. 1990), pools of indexed libraries should be amplified with a minimum number of PCR cycles or sequenced without amplification if possible.



Petitioner's Expert, Dr. Spellman (EX1054, ¶112)

There is "**no scientific basis** for asserting that chimeric molecules would not form during cluster amplification"

By Discarding All Inter-Sample Chimeras, *Parameswaran's* Unique Dual Indexing Successfully Reduces Cross-Over



KSR Int'l Co. v. Teleflex Inc.,
550 U.S. 398, 419-20 (2007)

It is error to "look only to the problem the patentee was trying to solve. . . . Under the correct analysis, ***any need or problem known*** in the field of endeavor at the time of invention and addressed by the patent can provide a reason for combining the elements in the manner claimed."

Patent Owner Fails to Read *Parameswaran* From the Perspective of a POSA

Patent Owner Response (POR, 31-32)

"Parameswaran's mismatched barcode motifs were **not cross-over error**. . . . Parameswaran instead posits three potential sources of its mismatched barcode motifs, **none** of which **are cross-over error**."

Donner Tech., LLC v. Pro Stage Gear, LLC
979 F.3d 1353, 1360 (Fed. Cir. 2020)

Obviousness analysis "must be carried out from the vantage point of a [POSA]"

***Gloor* (EX1005, 8)**

The occurrence of chimeric sequences was examined using UCHIME, a part of the UCLUST package (REF, http://www.drive5.com/uclust/). Chimeric sequences can arise during PCR [21] or because of overlapping spots on the solid support when imaged during DNA sequencing [6]. The default settings of



Petitioner's Expert, Dr. Spellman (EX1054, ¶70)

A POSA would have understood that "sequences with mismatched barcode motifs would have included sequences from chimeric molecules created between templates from two different samples."

illumina

Parameswaran "Discard[ed]" Chimeric Sequences

	454-66	454-67	454-68	454-69	454-70	454-71	454-72	454-73	454-74	454-75	454-76	454-77	454-78	454-79	454-80
454-66	13566	0	0	1	0	0	0	0	0	0	0	0	0	0	0
454-67	0	7459	0	3	0	0	0	0	0	0	0	0	0	0	0
454-68	0	0	13260	0	0	0	0	0	0	0	0	0	0	0	0
454-69	0	0	0	6518	0	0	0	0	0	0	0	0	0	0	0
454-70	0	0	0	0	9484	0	0	0	0	0	0	0	0	0	0
454-71	0	0	0	0	0	9397	0	0	0	0	0	0	0	0	
454-72	0	0	0	0	0	0	12302	0	0	0	0	0	0	0	
454-73	0	0	0	0	0	0	0	7273	0	0	0	0	0	0	
454-74	0	0	0	0	0	0	0	0	13923	0	0	0	0	0	
454-75	0	0	0	0	0	0	0	0	1	15060	0	0	0	0	0
454-76	0	0	0	0	0	0	0	0	1	0	10266	0	0	0	0
454-77	0	0	0	0	0	0	0	0	0	0	0	9602	0	0	0
454-78	0	0	0	0	0	0	0	0	0	0	0	0	12082	0	0
454-79	0	0	0	0	0	0	0	0	0	0	0	0	0	10349	0
454-80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100
454-81	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
454-82	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
454-83	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
454-84	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
454-85	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
454-86	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
454-87	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
454-88	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
454-89	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
454-90	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(close)	369	167	210	122	237	154	317	172	268	280	209	215	299	319	24
(no match)	37	26	76	48	28	20	48	29	64	46	42	36	78	59	60

Patent Owner Sur-Reply (Sur-Reply, 22)

"Petitioner asserts that Parameswaran 'greyed out' certain cells of data to indicate they were not used. But Parameswaran contains no such explanation, and Petitioner cites to none."



Petitioner's Expert, Dr. Spellman (EX2015, 109:13-110:1)

Q. Can you show me where it talks about grey boxes?

A. I will point you to the abstract, the very last sentence, "The false discovery rate as measured by the percentage of sequences with unexpected perfect pairings on unmatched forward/reverse barcodes was estimated to be less than 0.005 percent." That is exactly the grey cells in this table....

Parameswaran (EX1004, 8)

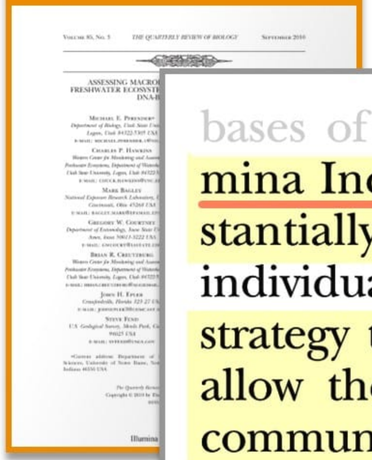
with reasonable accuracy into the correct bin. For our purposes, all sequences with a match to one of the two expected sequence motifs were considered as true members of the corresponding dataset. Run I (0.13%) and run II

There Were Multiple Reasons to Combine *Parameswaran* and *Gloor*

Patent Owner Response (POR, 21)

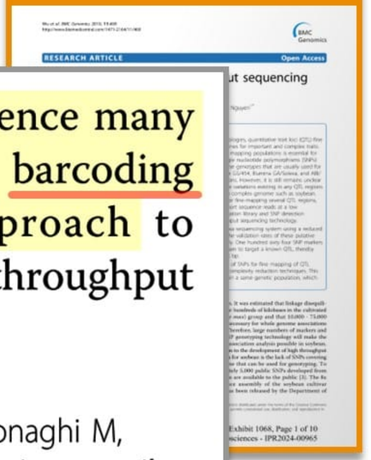
A POSA "would not have thought that Parameswaran's barcoding methods would have utility in Gloor's Illumina sequencers because Illumina sequencers did not have the same needs of pyrosequencing."

Pfrender **(EX1062, 11)**



bases of sequence. Other platforms (Illumina Inc. and ABIs SOLiD) generate substantially more bases in total, but shorter individual sequences. Applying a tagging strategy to uniquely identify samples would allow the combination of multiple whole community assays in a single run, thereby reducing the per sample cost (Meyer et al. 2007, 2008; Parameswaran et al. 2007).

Wu **(EX1068, 7)**



coding technique. It is not a problem to sequence many samples in a Solexa instrument run if using a barcoding technique. It is becoming a standard approach to increase the number of samples run on high-throughput instruments [19].

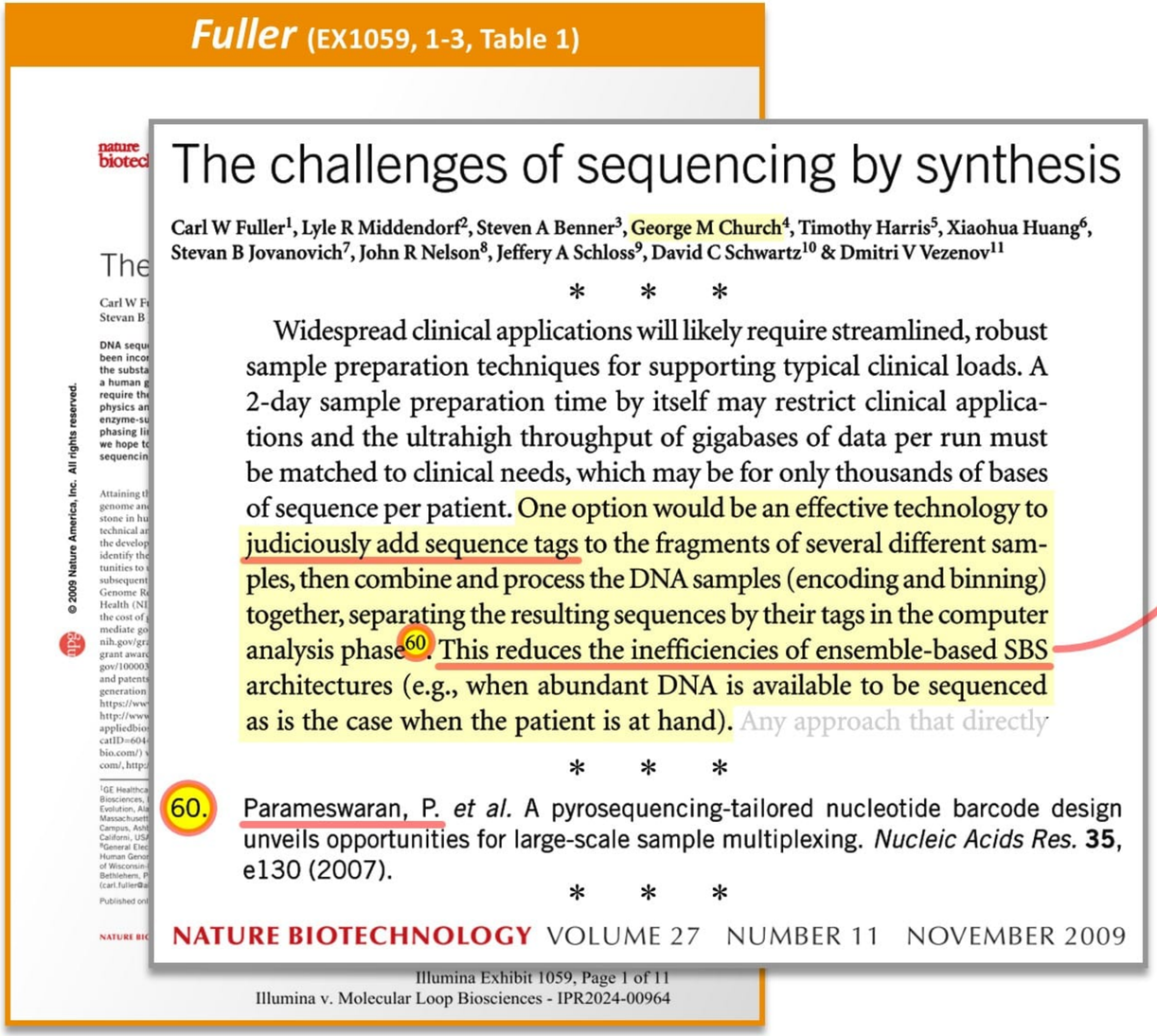
* * *

19. Parameswaran P, Jalili R, Tao L, Shokralla S, Gharizadeh B, Ronaghi M, Fire AZ: **A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing.** *Nucleic Acids Res* 2007, **35**:e130.

illumina

There Were Multiple Reasons to Combine *Parameswaran* and *Gloor*

Fuller (EX1059, 1-3, Table 1)



nature biotech

The challenges of sequencing by synthesis

Carl W Fuller[1], Lyle R Middendorf[2], Steven A Benner[3], George M Church[4], Timothy Harris[5], Xiaohua Huang[6], Stevan B Jovanovich[7], John R Nelson[8], Jeffery A Schloss[9], David C Schwartz[10] & Dmitri V Vezenov[11]

* * *

Widespread clinical applications will likely require streamlined, robust sample preparation techniques for supporting typical clinical loads. A 2-day sample preparation time by itself may restrict clinical applications and the ultrahigh throughput of gigabases of data per run must be matched to clinical needs, which may be for only thousands of bases of sequence per patient. One option would be an effective technology to judiciously add sequence tags to the fragments of several different samples, then combine and process the DNA samples (encoding and binning) together, separating the resulting sequences by their tags in the computer analysis phase[60]. This reduces the inefficiencies of ensemble-based SBS architectures (e.g., when abundant DNA is available to be sequenced as is the case when the patient is at hand). Any approach that directly

* * *

60. Parameswaran, P. *et al.* A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Res.* **35**, e130 (2007).

* * *

NATURE BIOTECHNOLOGY VOLUME 27 NUMBER 11 NOVEMBER 2009

Illumina Exhibit 1059, Page 1 of 11
Illumina v. Molecular Loop Biosciences - IPR2024-00964

Figure 1 Schemes for SBS. Approaches include both single-molecule and ensemble methods requiring features populated with identical copies of DNA fragments. Ensemble methods monitor synthesis either using labeled substrates or detecting products of synthesis (pyrophosphate). Detection can be discontinuous (sequential) when labels can remain fixed for hours while a microscope scans up to hundreds of millions of sequence features[26,32]. It

* * *

26. Bentley, D.R. *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53–59 (2008).

illumina

POSA Is Capable of Choosing Appropriate Barcoding Scheme

Patent Owner Sur-Reply (Sur-Reply, 17-19)

Parameswaran and *Gloor* "pitch [combinatorial dual-indexing] as much more than a mere alternative or choice."



Petitioner's Expert, Dr. Spellman (EX1054, ¶¶77, 109)

A "POSA would have understood that **sample tracking and sequence fidelity are critical in many situations**," including cancer research, ancient DNA research, and molecular diagnosis

Patent Owner's Expert, Dr. Adalsteinsson (EX1055, 113:16-19)



Q. Do you agree that some applications of sample multiplexing require highly accurate genotyping and that others do not?

A. Yes. . . .

Elekta Ltd. v. ZAP Surgical Sys., Inc., 81 F.4th 1368, 1374 (Fed. Cir. 2023)

"The inquiry into the existence of a motivation to combine assumes that a skilled artisan is a person of ordinary creativity with common sense, common wisdom, and common knowledge."

Bayer Pharma AG v. Watson Lab'ys, Inc., 874 F.3d 1316, 1328 (Fed. Cir. 2017)

"obviousness does not require that the motivation be the *best* option, only that it be a *suitable* option from which the prior art did not teach away"

EX1061, 1; EX1063, 1; EX1056, 19; EX1003, ¶¶160-161, EX1054, ¶¶75-78, 102-103, 109-110; EX1055, 113:16-114:10; Pet., 37-38; Reply 19-20, 22-23; Sur-Reply, 17-19.

Gloor Was Publicly Available Before December 2010

Gloor (EX1005, 1)



Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products

Gregory B. Gloor [1,*], Ruben Hummelen [2,3], Jean M. Macklaim [1,2], Russell J. Dickson [1] Andrew D. Fernandes [1,4], Roderick MacPhee [2,5], Gregor Reid [1,2,5,6]



DECLARATION OF GREGORY B. GLOOR, PH.D.

* * *

REPLY DECLARATION OF GREGORY B. GLOOR, PH.D.

* * *

I, Gregory B. Gloor, Ph.D., was asked whether my article titled "*Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products*" ("*Gloor*") was accessible to the public from the arXiv repository on July 29, 2010. This declaration provides my understanding that my article was accessible to the public as of July 29, 2010.

EX1002, 1

illumina

Gloor Was Publicly Available Before December 2010



"[I]t was my familiarity with **the publication policy** of arXiv.org, which ensured speedy public availability and ease of accessibility, that motivated me to submit *Gloor* to arXiv.org."

EX1046, ¶11

EXS1047-49

http://arxiv.org/help/submit Go OCT JUN SEP

947 captures
27 Jun 2001 - 13 May 2025

◀ 2009 **23** 2010 ▶ 2011

* * *

Availability of submissions and announcement order

* * *

Submissions to arXiv are normally announced on a daily schedule, with no activity scheduled on weekends (check current local time at the main site).

Example: A submission made on Monday before 16:00 (US Eastern time, EDT/EST) will be publicly available on the web after 20:00 (US Eastern time, EDT/EST) on that day, and will appear in the Tuesday AM regular subject mailing.

'852 Patent **(EX1001)**

Patent No.: US 11,041,852 B2

* * *

Provisional application No. 61/426,817, filed on Dec. 23, 2010.

July | August | September | October | November | December

EXS1047-49; EX1046, ¶¶11-13; Pet., 12; Reply, 24-28.

Gloor Was Publicly Available Before December 2010



"I submitted *Gloor* to arXiv.org on **July 28, 2010** at 17:17:22 ET." EX1002, ¶16

EX1002, Ex. B

Submission history

From: Greg Gloor Dr [view email]
[v1] Wed, 28 Jul 2010 21:17:22 UTC (899 KB)

EX1002, Ex. F

From: **arXiv Technical Support** jira@arxiv-org.atlassian.net
Subject: AH-93225 Date of release of https://arxiv.org/abs/1007.5075
Date: March 21, 2024 at 12:41 PM
To: ggloor@uwo.ca

* * *

Dear Greg,

Thank you for reaching out. https://arxiv.org/abs/1007.5075 was submitted: Wed, 28 Jul 2010 17:17:22 EST

'852 Patent **(EX1001)**

Patent No.: **US 11,041,852 B2**

* * *

Provisional application No. 61/426,817, filed on Dec. 23, 2010.

July | August | September | October | November | December

EX1002; Pet., 12; Reply, 24-28.

Gloor Was Publicly Available Before December 2010



"*Gloor* was made publicly available at [arxiv.org link] on **July 29, 2010**." EX1002, ¶17

EX1002, Ex. E

From: e-prints@arxiv.org
Subject: arXiv New submission -> 1007.5075 in q-bio.GN from ggloor@uwo.ca
Date: July 29, 2010 at 8:03 PM
To: ggloor@uwo.ca

* * *

Abstract will appear in today's mailing as:

\\
arXiv:1007.5075
From: Greg Gloor Dr <ggloor@uwo.ca>
Date: Wed, 28 Jul 2010 21:17:22 GMT (900kb)

Title: Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products
Authors: Gregory B. Gloor, Ruben Hummelen, Jean M. Macklaim, Russell J. Dickson, Andrew D. Fernandes, Roderick MacPhee, Gregor Reid

EX1051

Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products

Authors: Gregory B. Gloor, Ruben Hummelen, Jean M. Macklaim, Russell J. Dickson, Andrew D. Fernandes, Roderick MacPhee, Gregor Reid

* * *

Submitted 28 July, 2010; **originally announced** July 2010.

Correspondence on Mar. 31, 2025 (EX1050)

From: **arXiv Technical Support** jira@arxiv-org.atlassian.net
Subject: AH-93225 Date of release of https://arxiv.org/abs/1007.5075
Date: March 31, 2025 at 10:27 AM
To: ggloor@uwo.ca

* * *

Dear Greg,

I looked at this again and wanted to clarify. The paper would have appeared online on 2010-07-30 sometime after midnight or 1AM UTC.

* * *

Regards,

Jim

'852 Patent (EX1001)

Patent No.: **US 11,041,852 B2**

* * *

Provisional application No. 61/426,817, filed on Dec. 23, 2010.

July | August | September | October | November | December

EX1002; EX1050; EX1051; EX1046, ¶¶14-16; Pet., 12; Reply, 24-28.

Gloor Was Publicly Available Before December 2010



"I presented a poster ... between **August 31 and September 2, 2010**" with "a complete citation to [*Gloor*]." EX1002, ¶18

EX1002, Ex. G

- Human Microbiome Research Conference
  August 31-September 2, 2010
  Chase Park Plaza, St Louis, Missouri

EX1002, Ex. H

References

Hummelen, R., Fernandes, A.D., Macklaim, J.M., Dickson, R.J., Changalucha, J., Gloor, G.B., and Reid, G. (2010). Deep sequencing of the vaginal microbiota of women with HIV. PLoS One *5, e12078.*

Gloor, G.B., Hummelen, R., Macklaim, J.M., Dickson, R.J., Fernandes, A.D., MacPhee, R., and Reid, G. (2010). Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products. ArXiv e-Prints. arXiv:1007.5075v1 [q-bio.GN]

'852 Patent **(EX1001)**

Patent No.: US 11,041,852 B2

* * *

Provisional application No. 61/426,817, filed on Dec. 23, 2010.

July | August | September | October | November | December

EX1002; Pet., 12; Reply, 24-28.

Gloor Was Publicly Available Before December 2010



"*Gloor* was the subject matter of a discussion in a public forum called mothur.org as of **October 2010**" and "I was a participant in the discussion."

EX1046, ¶17

EX1052

U **umberar7** Oct 2010

Hi Pat,
there is another paper where they developed their own pipeline to analyze paired end . It was published in PLoS ONE in august 2010 vol.5 by R. Hummelen. One of the authors was Gregory Gloor (I believe he developed the code) and he is in the process of publishing a methodology paper on analyzing paired ends.
You can find the paper at http://arxiv.org/abs/1007.5075. And yes the biggest problem is matching

'852 Patent **(EX1001)**

Patent No.: US 11,041,852 B2

* * *

Provisional application No. 61/426,817, filed on Dec. 23, 2010.

July | August | September | October | November | December

EX1052; EX1046, ¶¶17-18; Pet., 12; Reply, 24-28.

Gloor Was Publicly Available Before December 2010



"In **November 2010**, following the publication of the 'methodology paper' in PLoS ONE on **October 26, 2010**," "I posted a response in the discussion."

EX1046, ¶18

EX1052

G ggloor Nov 2010

We have now published the methods paper and have an automated pipeline to take paired-end Illumina reads all the way from the reads to the overlap to the ISU and OTU sequences. Feel free to contact me or my team members for more information

EX1024

OPEN ACCESS Freely available online PLoS one

Microbiome Profiling by Illumina Sequencing of Combinatorial Sequence-Tagged PCR Products

Gregory B. Gloor[1]*, Ruben Hummelen[2,3], Jean M. Macklaim[1,2], Russell J. Dickson[1], Andrew D. Fernandes[1,4], Roderick MacPhee[2,5], Gregor Reid[1,2,5,6]

* * *

October 2010 | Volume 5 | Issue 10 | e15406

'852 Patent **(EX1001)**

Patent No.: US 11,041,852 B2

* * *

Provisional application No. 61/426,817, filed on Dec. 23, 2010.

July | August | September | October | November | December

EX1024; EX1052; EX1046, ¶¶17-18; Pet., 12; Reply, 24-28.

Gloor Was Publicly Available Before December 2010

Correspondence on Feb. 27, 2024 (EX1039)

From: arXiv Technical Support <jira@arxiv-org.atlassian.net>
Sent: Tuesday, February 27, 2024 11:21 AM
To: Cheek, Douglas
Subject: AH-90140 Request to Speak re arXiv.org

* * *

Hi Douglas,

Thank you for the additional information. The paper was made public on arxiv.org at 8PM Eastern on 7/29/2010.

Regards,

Stephanie

Gloor Was Publicly Available Before December 2010

Correspondence on Feb. 27, 2024 (EX1039)

From:
Sent:
To:
Subject:

Hi Douglas,

Thank you for the a
7/29/2010.

Regards,

Stephanie

Correspondence on Mar. 21, 2024 (EX1002, Ex. F)

From: **arXiv Technical Support** jira@arxiv-org.atlassian.net
Subject: AH-93225 Date of release of https://arxiv.org/abs/1007.5075
Date: March 21, 2024 at 12:41 PM
To: ggloor@uwo.ca

* * *

Dear Greg,

* * *

It would have been announced on Thursday, July 29, 2010.

Regards,
arXiv Support

Gloor Was Publicly Available Before December 2010

Correspondence on Feb. 17, 2025 (EX1043)

From: arXiv Technical Support <jira@arxiv-org.atlassian.net>
Sent: Monday, February 17, 2025 11:39 AM
To: Collins-Chase, Charles
Subject: AH-129177 Publication Information (Attn: Stephanie Orphan)

* * *

Hi Charles,

For the paper https://doi.org/10.48550/arXiv.1007.5075 it was posted on 30 July 2010 shortly after midnight UTC (8PM ET).

Unfortunately arXiv as a service does not provide affidavits.

Regards,

Jim Entwood
arXiv Head of Content & User Support

illumina

Gloor Was Publicly Available Before December 2010

Correspondence on Mar. 31, 2025 (EX1050)

From: **arXiv Technical Support** jira@arxiv-org.atlassian.net
Subject: AH-93225 Date of release of https://arxiv.org/abs/1007.5075
Date: March 31, 2025 at 10:27 AM
To: ggloor@uwo.ca

* * *

Dear Greg,

I looked at this again and wanted to clarify. The paper would have appeared online on 2010-07-30 sometime after midnight or 1AM UTC.

* * *

Regards,

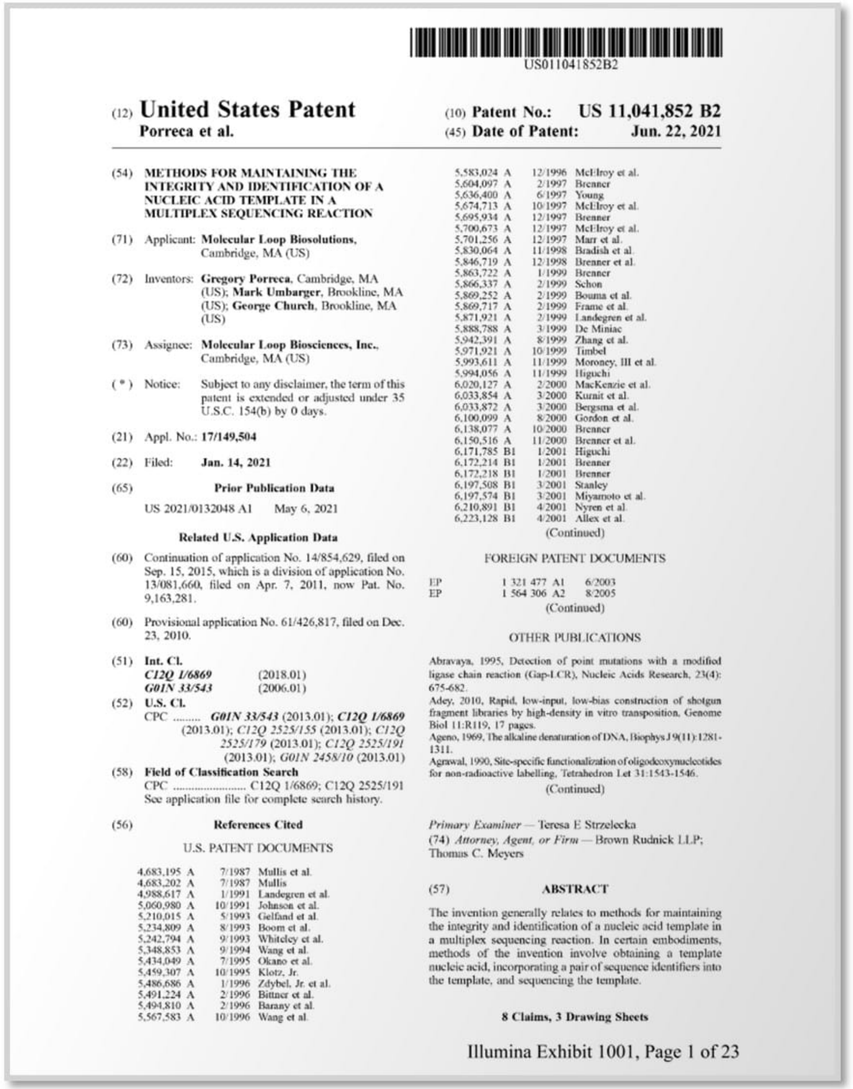
Jim

***CQV Co. v. Merck Pat. GmbH*, 130 F.4th 1344, 1350 (Fed. Cir. 2025)**

Public availability assessment should "not . . . overstate the required degree of certainty. . . . [A] party with the burden of persuasion must convince the factfinder to the degree of certainty required by the standard (here, preponderance of the evidence), and no further."

illumina

Ground 2: *Parameswaran* and *Bentley*



US011041852B2

(12) **United States Patent**
Porreca et al.

(10) **Patent No.: US 11,041,852 B2**
(45) **Date of Patent: Jun. 22, 2021**

(54) **METHODS FOR MAINTAINING THE INTEGRITY AND IDENTIFICATION OF A NUCLEIC ACID TEMPLATE IN A MULTIPLEX SEQUENCING REACTION**

(71) Applicant: **Molecular Loop Biosolutions**, Cambridge, MA (US)

(72) Inventors: **Gregory Porreca**, Cambridge, MA (US); **Mark Umbarger**, Brookline, MA (US); **George Church**, Brookline, MA (US)

(73) Assignee: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **17/149,504**

(22) Filed: **Jan. 14, 2021**

(65) **Prior Publication Data**

US 2021/0132048 A1 May 6, 2021

Related U.S. Application Data

(60) Continuation of application No. 14/854,629, filed on Sep. 15, 2015, which is a division of application No. 13/081,660, filed on Apr. 7, 2011, now Pat. No. 9,163,281.

(60) Provisional application No. 61/426,817, filed on Dec. 23, 2010.

(51) **Int. Cl.**
C12Q 1/6869 (2018.01)
G01N 33/543 (2006.01)

(52) **U.S. Cl.**
CPC ***G01N 33/543*** (2013.01); ***C12Q 1/6869*** (2013.01); *C12Q 2525/155* (2013.01); *C12Q 2525/179* (2013.01); *C12Q 2525/191* (2013.01); *G01N 2458/10* (2013.01)

(58) **Field of Classification Search**
CPC C12Q 1/6869; C12Q 2525/191
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,683,195	A	7/1987	Mullis et al.
4,683,202	A	7/1987	Mullis
4,988,617	A	1/1991	Landegren et al.
5,060,980	A	10/1991	Johnson et al.
5,210,015	A	5/1993	Gelfand et al.
5,234,809	A	8/1993	Boom et al.
5,242,794	A	9/1993	Whiteley et al.
5,348,853	A	9/1994	Wang et al.
5,434,049	A	7/1995	Okano et al.
5,459,307	A	10/1995	Klotz, Jr.
5,486,686	A	1/1996	Zdybel, Jr. et al.
5,491,224	A	2/1996	Bittner et al.
5,494,810	A	2/1996	Barany et al.
5,567,583	A	10/1996	Wang et al.
5,583,024	A	12/1996	McElroy et al.
5,604,097	A	2/1997	Brenner
5,636,400	A	6/1997	Young
5,674,713	A	10/1997	McElroy et al.
5,695,934	A	12/1997	Brenner
5,700,673	A	12/1997	McElroy et al.
5,701,256	A	12/1997	Marr et al.
5,830,064	A	11/1998	Bradish et al.
5,846,719	A	12/1998	Brenner et al.
5,863,722	A	1/1999	Brenner
5,866,337	A	2/1999	Schon
5,869,252	A	2/1999	Bouma et al.
5,869,717	A	2/1999	Frame et al.
5,871,921	A	2/1999	Landegren et al.
5,888,788	A	3/1999	De Mintac
5,942,391	A	8/1999	Zhang et al.
5,971,921	A	10/1999	Timbel
5,993,611	A	11/1999	Moroney, III et al.
5,994,056	A	11/1999	Higuchi
6,020,127	A	2/2000	MacKenzie et al.
6,033,854	A	3/2000	Kurait et al.
6,033,872	A	3/2000	Bergsma et al.
6,100,099	A	8/2000	Gordon et al.
6,138,077	A	10/2000	Brenner
6,150,516	A	11/2000	Brenner et al.
6,171,785	B1	1/2001	Higuchi
6,172,214	B1	1/2001	Brenner
6,172,218	B1	1/2001	Brenner
6,197,508	B1	3/2001	Stanley
6,197,574	B1	3/2001	Miyamoto et al.
6,210,891	B1	4/2001	Nyren et al.
6,223,128	B1	4/2001	Allex et al.

(Continued)

FOREIGN PATENT DOCUMENTS

EP	1 321 477	A1	6/2003
EP	1 564 306	A2	8/2005

(Continued)

OTHER PUBLICATIONS

Abravaya, 1995, Detection of point mutations with a modified ligase chain reaction (Gap-LCR), Nucleic Acids Research, 23(4): 675-682.
Adey, 2010, Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition, Genome Biol 11:R119, 17 pages.
Ageno, 1969, The alkaline denaturation of DNA, Biophys J 9(11):1281-1311.
Agrawal, 1990, Site-specific functionalization of oligodeoxynucleotides for non-radioactive labelling, Tetrahedron Let 31:1543-1546.

(Continued)

Primary Examiner — Teresa E Strzelecka
(74) *Attorney, Agent, or Firm* — Brown Rudnick LLP; Thomas C. Meyers

(57) **ABSTRACT**

The invention generally relates to methods for maintaining the integrity and identification of a nucleic acid template in a multiplex sequencing reaction. In certain embodiments, methods of the invention involve obtaining a template nucleic acid, incorporating a pair of sequence identifiers into the template, and sequencing the template.

8 Claims, 3 Drawing Sheets

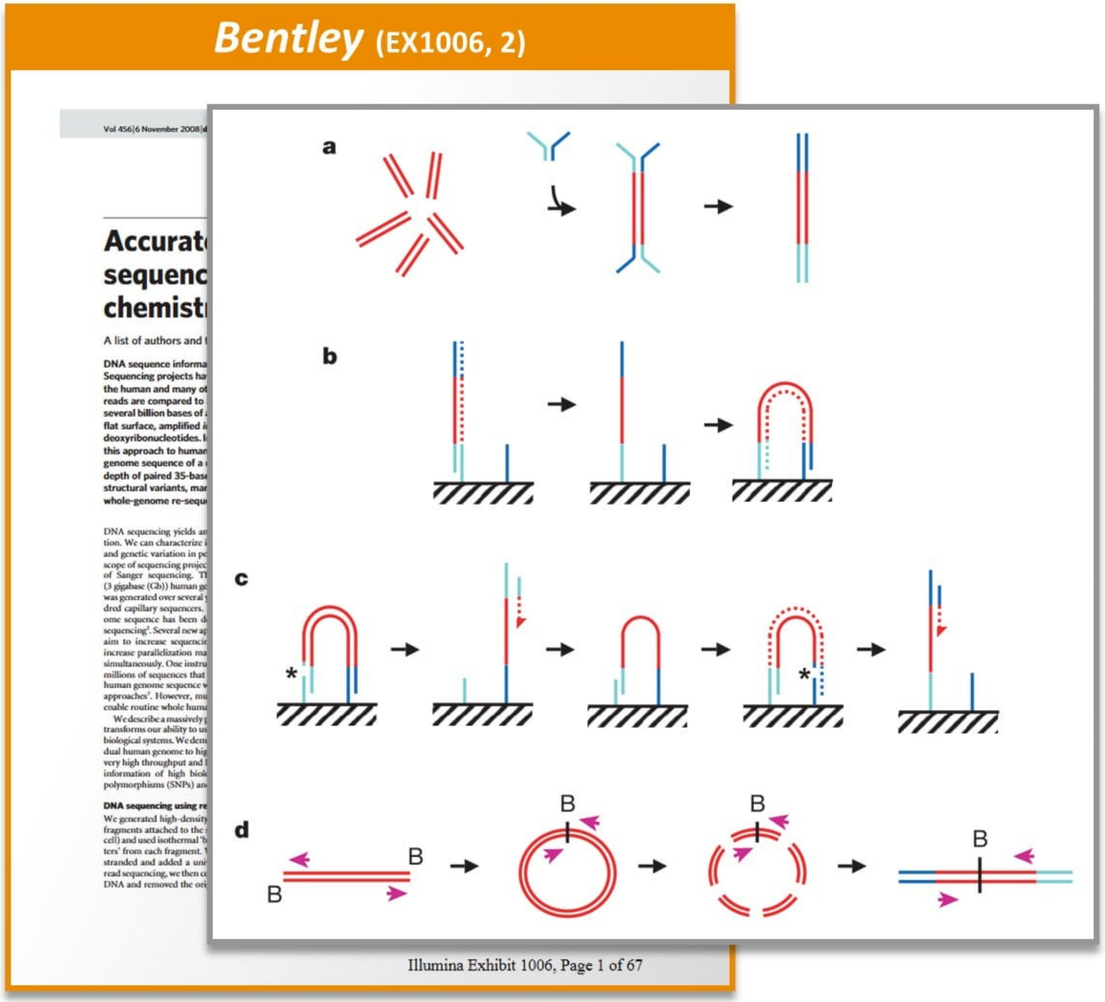
Illumina Exhibit 1001, Page 1 of 23

Ground	References	Basis	Claims
1	*Parameswaran* (EX1004) and *Gloor* (EX1005)	§103	1-8
2	*Parameswaran* (EX1004) and *Bentley* (EX1006)	§103	1-8

Claims Would Have Been Obvious Over *Parameswaran* and *Bentley*

1[c] amplifying said nucleic acid templates on a surface of a flow cell thereby to form clusters,

Bentley **(EX1006, 2)**



Motivation to Combine

Illumina Sequencers Were Faster, More Efficient, and Less Expensive (Pet., 36-37, 55-56)

To Reduce Sequencing Error, Including Chimeras Formed During Multiplex Sequencing (Pet., 37-38, 56)

***Parameswaran*'s Recommendation** (Pet., 38-41, 56)

Reasonable Expectation of Success

Interchangeability Between Roche Pyrosequencing and Illumina Sequencing (Pet., 41-44, 56-57)

Increase in Illumina Sequencing Read Length (Pet., 44-45, 57-58)

illumina

EX1006, 2; EX1003, ¶¶156-172, 195-222; Pet., 36-45, 51-61.

Patent Owner Fails to Rebut Obviousness

Patent Owner's Expert, Dr. Adalsteinsson (EX2001, ¶149)



"Bentley does not have any discussion of multiplexing. Instead, Bentley performed whole-genome sequencing for a sample using Illumina sequencers"

Donner Tech., LLC v. Pro Stage Gear, LLC
979 F.3d 1353, 1360 (Fed. Cir. 2020)

Obviousness analysis "must be carried out from the vantage point of a [POSA]"

Patent Owner's Expert, Dr. Adalsteinsson (EX1055, 73:4-7)



Q. So you agree that by 2010, it was known that an Illumina sequencer could be us[ed] for multiplex sequencing; correct?

A. Yes, I agree.

'200 Patent (IPR2024-00965)

Challenged Claims – '200 Patent

1 [pre]	A method for **validating** the sequence of a nucleic acid analyte of interest in a multiplex sequencing reaction, the method comprising
[a]	**detecting** the presence of two or more identifier sequences that are uniquely associated with the nucleic acid analyte of interest, wherein at least a first identifier sequence is incorporated into a 5′ portion of the nucleic acid analyte of interest and wherein at least a second identifier sequence is incorporated into a 3′ portion of the nucleic acid analyte of interest, and wherein the first and second identifier sequences have four or more nucleotides and are different;
[b]	**sequencing** the first identifier sequence, the nucleic acid analyte of interest and the second identifier sequence; and
[c]	**validating** the sequence of the nucleic acid analyte of interest by **analyzing both identifiers** and **excluding the sequences** of those nucleic acid analytes of interest **containing only one identifier or an incorrect pair of identifiers** from sequence analysis of the multiplex sequencing reaction.

No Reference to "Cross-Over Error" (Reply, 9)

Individual Identifiers Can Be Used with Different Samples (Reply, 2-4)

No Reference to a "Flow Cell" (Reply, 1, 9)

Excludes Sequences With "Only One Identifier" *or* an "Incorrect Pair of Identifiers" (Reply, 2-3, 9)

illumina

Grounds for IPR2024-00965 ('200 Patent)

Gloor (EX1005)



Illumina Exhibit 1005, Page 1 of 28

Ground 1

Gloor
anticipates all claims

Parameswaran (EX1004)



Illumina Exhibit 1004, Page 1 of 31

Ground 2

Parameswaran
anticipates all claims

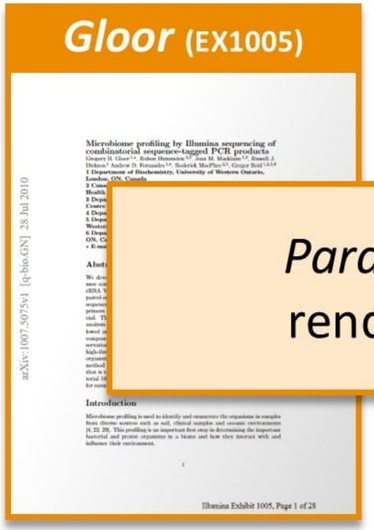
Parameswaran (EX1004)



Illumina Exhibit 1004, Page 1 of 31

+

Gloor (EX1005)



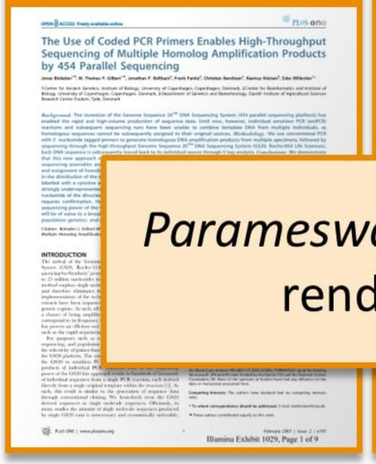
Illumina Exhibit 1005, Page 1 of 28

Ground 3

Parameswaran and *Gloor*
render all claims **obvious**

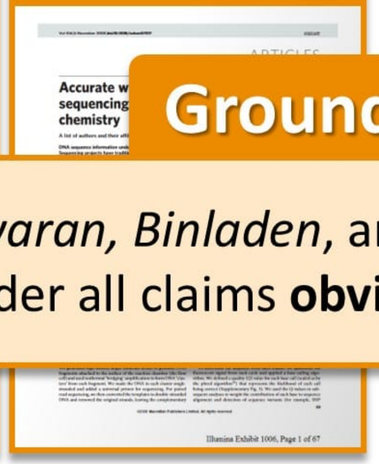
+

Binladen (EX1029)



Illumina Exhibit 1029, Page 1 of 9

Bentley (EX1006)



Illumina Exhibit 1006, Page 1 of 67

Grounds 4 & 5

Parameswaran, *Binladen*, and *Bentley*
render all claims **obvious**

illumina

Claim-Construction: "Two or More Identifier Sequences That Are Uniquely Associated"

1 [pre]	A method for validating the sequence of a nucleic acid analyte of interest in a multiplex sequencing reaction, the method comprising
[a]	**detecting** the presence of two or more identifier sequences that are uniquely associated with the nucleic acid analyte of interest, wherein at least a first identifier sequence is incorporated into a 5′ portion of the nucleic acid analyte of interest and wherein at least a second identifier sequence is incorporated into a 3′ portion of the nucleic acid analyte of interest, and wherein the first and second identifier sequences have four or more nucleotides and are different;
[b]	**sequencing** the first identifier sequence, the nucleic acid analyte of interest and the second identifier sequence; and
[c]	**validating** the sequence of the nucleic acid analyte of interest by **analyzing both identifiers** and **excluding the sequences** of those nucleic acid analytes of interest **containing only one identifier or an incorrect pair of identifiers** from sequence analysis of the multiplex sequencing reaction.

Does Not State That Each Identifier Sequence Must Be Uniquely Associated (Reply, 2-4)

'200 Patent **(EX1001, 2:50-53)**

sequenced, or portions thereof. While the pair of barcodes associated with any given template is unique, individual barcode sequences can be used with multiple templates as long as no pair is used across two different samples. While



Petitioner's Expert, Dr. Spellman (EX1054, ¶16)

This term "should be given its plain and ordinary meaning, i.e., **each pair of two or more identifier sequences must be uniquely associated** with the nucleic acid analyte of interest."

illumina

Patent Owner's Narrow Construction Is Contrary to the Evidence

1 [pre] A method for validating the sequence of a nucleic acid analyte of interest in a multiplex sequencing reaction, the method comprising

[a] **detecting** the presence of two or more identifier sequences that are uniquely associated with the nucleic acid analyte of interest,
wherein at least a first identifier sequence is incorporated into a 5′ portion of the nucleic acid analyte of interest and wherein at least a second identifier sequence is incorporated into a 3′ portion of the nucleic acid analyte of interest, and
wherein the first and second identifier sequences have four or more nucleotides and are different;

[b] **sequencing** the first identifier sequence, the nucleic acid analyte of interest and the second identifier sequence; and

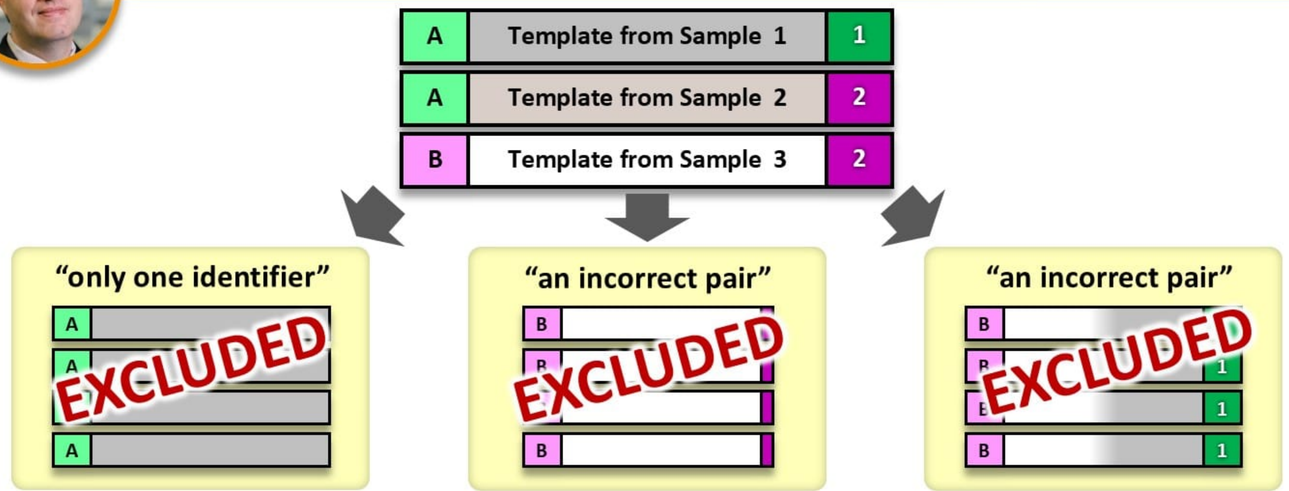
[c] **validating** the sequence of the nucleic acid analyte of interest by **analyzing both identifiers** and **excluding the sequences** of those nucleic acid analytes of interest **containing only one identifier or an incorrect pair of identifiers** from sequence analysis of the multiplex sequencing reaction.

Patent Owner Response (POR, 14-15)

"One would not be able to discern whether pairs of identifiers were 'incorrect' if identifier sequences were used across different nucleic acids of interest."



Petitioner's Expert, Dr. Spellman (EX1054, ¶¶7-16)

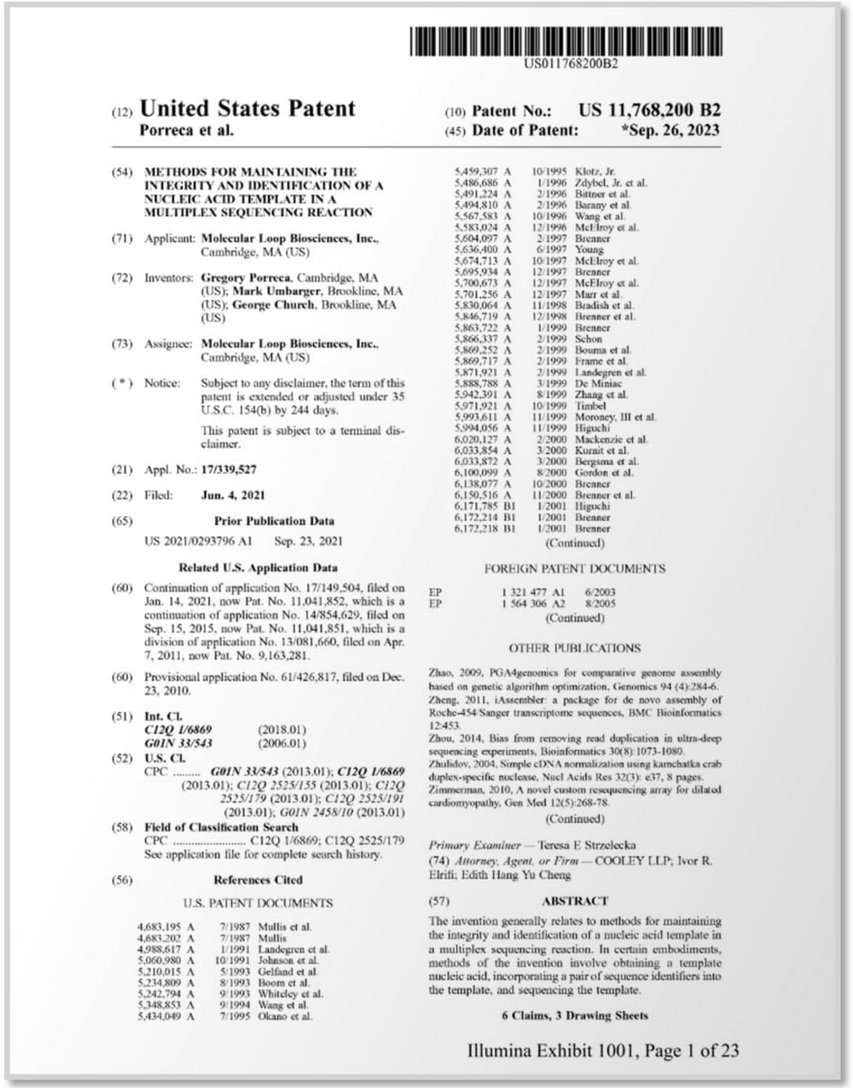


Recites Excluding Sequences With "Only One Identifier OR an Incorrect Pair of Identifiers" (Reply, 2-3)

illumina

EX1001 ('200 Pat.), cl. 1; EX1054, ¶¶7-16; EX2002, 14:20-16:1; POR, 14-15; Reply, 2-3.

Ground 1: *Gloor*



US011768200B2

(12) **United States Patent**
Porreca et al.

(10) **Patent No.: US 11,768,200 B2**
(45) **Date of Patent: *Sep. 26, 2023**

(54) **METHODS FOR MAINTAINING THE INTEGRITY AND IDENTIFICATION OF A NUCLEIC ACID TEMPLATE IN A MULTIPLEX SEQUENCING REACTION**

(71) Applicant: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(72) Inventors: **Gregory Porreca**, Cambridge, MA (US); **Mark Umbarger**, Brookline, MA (US); **George Church**, Brookline, MA (US)

(73) Assignee: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 244 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **17/339,527**

(22) Filed: **Jun. 4, 2021**

(65) **Prior Publication Data**

US 2021/0293796 A1 Sep. 23, 2021

Related U.S. Application Data

(60) Continuation of application No. 17/149,504, filed on Jan. 14, 2021, now Pat. No. 11,041,852, which is a continuation of application No. 14/854,629, filed on Sep. 15, 2015, now Pat. No. 11,041,851, which is a division of application No. 13/081,660, filed on Apr. 7, 2011, now Pat. No. 9,163,281.

(60) Provisional application No. 61/426,817, filed on Dec. 23, 2010.

(51) **Int. Cl.**
C12Q 1/6869 (2018.01)
G01N 33/543 (2006.01)

(52) **U.S. Cl.**
CPC ***G01N 33/543*** (2013.01); ***C12Q 1/6869*** (2013.01); *C12Q 2525/155* (2013.01); *C12Q 2525/179* (2013.01); *C12Q 2525/191* (2013.01); *G01N 2458/10* (2013.01)

(58) **Field of Classification Search**
CPC C12Q 1/6869; C12Q 2525/179
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,683,195 A 7/1987 Mullis et al.
4,683,202 A 7/1987 Mullis
4,988,617 A 1/1991 Landegren et al.
5,060,980 A 10/1991 Johnson et al.
5,210,015 A 5/1993 Gelfand et al.
5,234,809 A 8/1993 Boom et al.
5,242,794 A 9/1993 Whiteley et al.
5,348,853 A 9/1994 Wang et al.
5,434,049 A 7/1995 Okano et al.
5,459,307 A 10/1995 Klotz, Jr.
5,486,686 A 1/1996 Zdybel, Jr. et al.
5,491,224 A 2/1996 Bittner et al.
5,494,810 A 2/1996 Barany et al.
5,567,583 A 10/1996 Wang et al.
5,583,024 A 12/1996 McElroy et al.
5,604,097 A 2/1997 Brenner
5,636,400 A 6/1997 Young
5,674,713 A 10/1997 McElroy et al.
5,695,934 A 12/1997 Brenner
5,700,673 A 12/1997 McElroy et al.
5,701,256 A 12/1997 Marr et al.
5,830,064 A 11/1998 Bradish et al.
5,846,719 A 12/1998 Brenner et al.
5,863,722 A 1/1999 Brenner
5,866,337 A 2/1999 Schon
5,869,252 A 2/1999 Bouma et al.
5,869,717 A 2/1999 Frame et al.
5,871,921 A 2/1999 Landegren et al.
5,888,788 A 3/1999 De Miniac
5,942,391 A 8/1999 Zhang et al.
5,971,921 A 10/1999 Timbel
5,993,611 A 11/1999 Moroney, III et al.
5,994,056 A 11/1999 Higuchi
6,020,127 A 2/2000 Mackenzie et al.
6,033,854 A 3/2000 Kurait et al.
6,033,872 A 3/2000 Bergsma et al.
6,100,099 A 8/2000 Gordon et al.
6,138,077 A 10/2000 Brenner
6,150,516 A 11/2000 Brenner et al.
6,171,785 B1 1/2001 Higuchi
6,172,214 B1 1/2001 Brenner
6,172,218 B1 1/2001 Brenner
(Continued)

FOREIGN PATENT DOCUMENTS

EP 1 321 477 A1 6/2003
EP 1 564 306 A2 8/2005
(Continued)

OTHER PUBLICATIONS

Zhao, 2009, PGAAgenomics for comparative genome assembly based on genetic algorithm optimization, Genomics 94 (4):284-6.
Zheng, 2011, iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequences, BMC Bioinformatics 12:453.
Zhou, 2014, Bias from removing read duplication in ultra-deep sequencing experiments, Bioinformatics 30(8):1073-1080.
Zhulidov, 2004, Simple cDNA normalization using kamchatka crab duplex-specific nuclease, Nucl Acids Res 32(3): e37, 8 pages.
Zimmerman, 2010, A novel custom resequencing array for dilated cardiomyopathy, Gen Med 12(5):268-78.
(Continued)

Primary Examiner — Teresa E Strzelecka
(74) *Attorney, Agent, or Firm* — COOLEY LLP; Ivor R. Elrifi; Edith Hang Yu Cheng

(57) **ABSTRACT**

The invention generally relates to methods for maintaining the integrity and identification of a nucleic acid template in a multiplex sequencing reaction. In certain embodiments, methods of the invention involve obtaining a template nucleic acid, incorporating a pair of sequence identifiers into the template, and sequencing the template.

6 Claims, 3 Drawing Sheets

Illumina Exhibit 1001, Page 1 of 23

Ground	References	Basis	Claims
1	*Gloor* (EX1005)	§102	1-6
2	*Parameswaran* (EX1004)	§102	1-6
3	*Parameswaran* (EX1004) and *Gloor* (EX1005)	§103	1-6
4	*Parameswaran* (EX1004) and *Binladen* (EX1029)	§103	1-6
5	*Parameswaran* (EX1004), *Binladen* (EX1029), and *Bentley* (EX1006)	§103	6

illumina

Gloor Anticipates All Claims

1 [pre] A method for validating the sequence of a nucleic acid analyte of interest in a multiplex sequencing reaction, the method comprising

[a] **detecting** the presence of two or more identifier sequences that are uniquely associated with the nucleic acid analyte of interest,
wherein at least a first identifier sequence is incorporated into a 5′ portion of the nucleic acid analyte of interest and wherein at least a second identifier sequence is incorporated into a 3′ portion of the nucleic acid analyte of interest, and
wherein the first and second identifier sequences have four or more nucleotides and are different;

[b] **sequencing** the first identifier sequence, the nucleic acid analyte of interest and the second identifier sequence; and

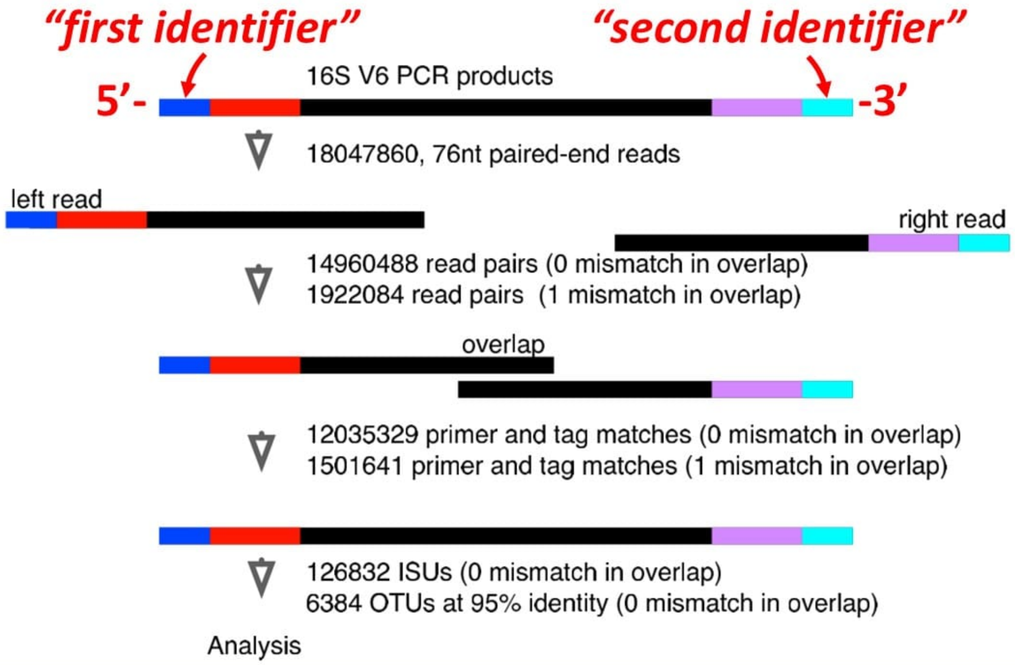
[c] validating the sequence of the nucleic acid analyte of interest by analyzing both identifiers and excluding the sequences of those nucleic acid analytes of interest containing only one identifier or an incorrect pair of identifiers from sequence analysis of the multiplex sequencing reaction.

Gloor **(EX1005, 6-7, Fig. 2)**

Table 2: **sequence tag and primer sequences**

L-tag	Name	R-tag	Name
catgcg	0-v6L	cgcatg	0-v6R
gcagt	1-v6L	actgc	1-v6R
tagct	2-v6L	agcta	2-v6R
gactgt	3-v6L	acagtc	3-v6R

* * *



illumina

Gloor Anticipates All Claims

1 [pre] A method for validating the sequence of a nucleic acid analyte of interest in a multiplex sequencing reaction, the method comprising

[a] detecting the presence of two or more identifier sequences that are uniquely associated with the nucleic acid analyte of interest, wherein at least a first identifier sequence is incorporated into a 5′ portion of the nucleic acid analyte of interest and wherein at least a second identifier sequence is incorporated into a 3′ portion of the nucleic acid analyte of interest, and wherein the first and second identifier sequences have four or more nucleotides and are different;

[b] sequencing the first identifier sequence, the nucleic acid analyte of interest and the second identifier sequence; and

[c] **validating** the sequence of the nucleic acid analyte of interest by **analyzing both identifiers** and **excluding the sequences** of those nucleic acid analytes of interest **containing only** one identifier or an incorrect pair of identifiers from sequence analysis of the multiplex sequencing reaction.

Gloor (EX1005, 6-7)

Inspection of sequences with incorrect sequence tags showed that the single largest contributor to the difference between the number of reads with proper primer sequences and the number of reads with proper sequence tag sequences was an N-1 truncation of the sequence tag, which presumably arose during the primer synthesis. The next largest class of sequence tag error was complete lack of the left or right end sequence tag. Together, these classes account for slightly more than half of the missing reads. The remaining missing reads are composed of a large number of classes of sequence tag sequences each containing small numbers of errors including additional 5' bases, misincorporated bases or difficult to classify errors that presumably arose during the PCR amplification.

Patent Owner's Expert, Dr. Adalsteinsson (EX1055, 77:10-12)



Q. So these are sequences of nucleic acids containing only one identifier sequence; correct?

A. I believe so.



Petitioner's Expert, Dr. Spellman (EX1054, ¶18)

Gloor "excluded sequences that contained truncated tags, tags with additional bases, or tags with misincorporated bases, i.e., sequences containing an incorrect pair of identifiers."

Patent Owner's Arguments on Claim 5 Should Be Rejected

5 The method of claim 4, wherein t [sic] barcode sequences are **designed to have no homopolymer regions**.

Gloor (EX1005, 6)

Table 2: **sequence tag and primer sequences**

L-tag	Name	R-tag	Name
catgcg	0-v6L	cgcatg	0-v6R
gcagt	1-v6L	actgc	1-v6R
tagct	2-v6L	agcta	2-v6R
gactgt	3-v6L	acagtc	3-v6R

Patent Owner Response (POR, 15)

"Gloor's discussion of its sequence tag design contains ***no discussion*** of homopolymer regions."

Patent Owner's Expert, Dr. Adalsteinsson (EX1055, 79:18-80:1)



Q. Why don't you take a minute and confirm that the barcode sequences used in Gloor do not include two of the same bases used in a row.

A. I recall looking at this in my analysis and believe that is right, that none of those have two or more in a row. . . .

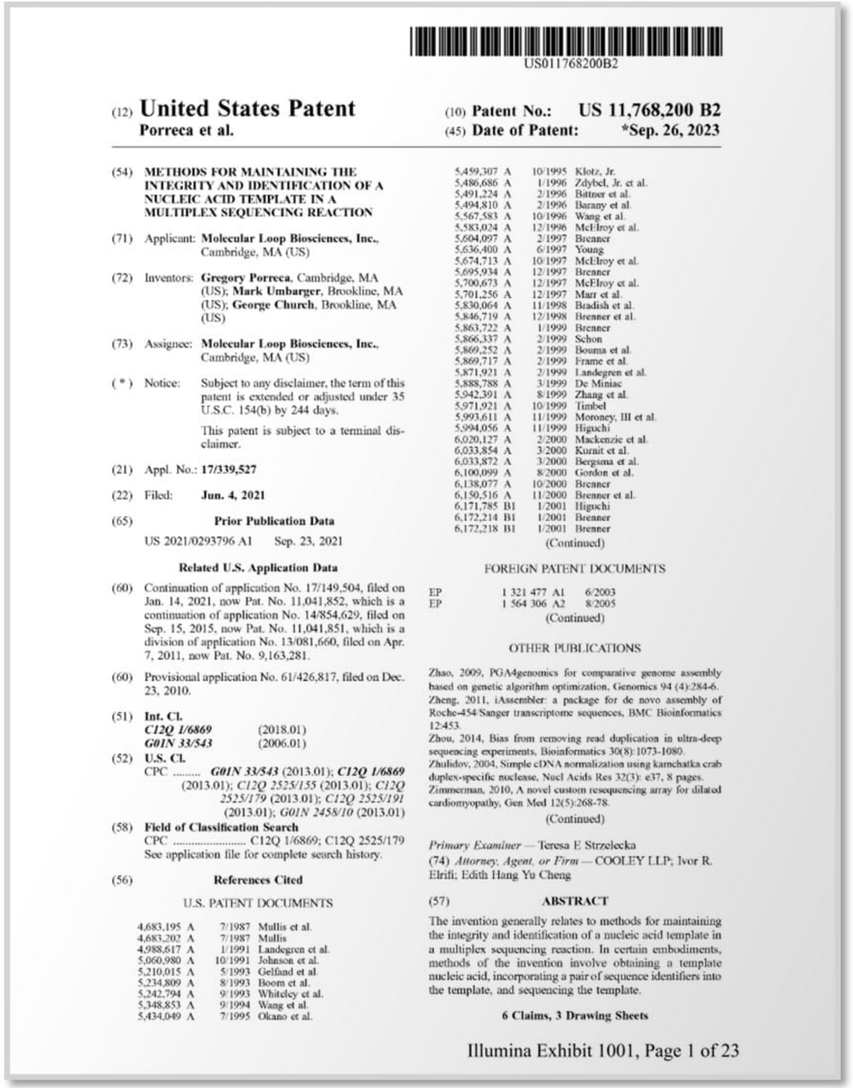


Petitioner's Expert, Dr. Spellman (EX1054, ¶24)

"The chance of all the barcodes in *Gloor* having no homopolymer region simply by chance is incredibly small, about 2.4 X 10^{-15}, which is about 2 in 1 quadrillion (or about 1 million times lower odds than winning the Powerball lottery)."

illumina

Ground 2: *Parameswaran*



US011768200B2

(12) **United States Patent**
Porreca et al.

(10) **Patent No.: US 11,768,200 B2**
(45) **Date of Patent: *Sep. 26, 2023**

(54) **METHODS FOR MAINTAINING THE INTEGRITY AND IDENTIFICATION OF A NUCLEIC ACID TEMPLATE IN A MULTIPLEX SEQUENCING REACTION**

(71) Applicant: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(72) Inventors: **Gregory Porreca**, Cambridge, MA (US); **Mark Umbarger**, Brookline, MA (US); **George Church**, Brookline, MA (US)

(73) Assignee: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 244 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **17/339,527**

(22) Filed: **Jun. 4, 2021**

(65) **Prior Publication Data**

US 2021/0293796 A1 Sep. 23, 2021

Related U.S. Application Data

(60) Continuation of application No. 17/149,504, filed on Jan. 14, 2021, now Pat. No. 11,041,852, which is a continuation of application No. 14/854,629, filed on Sep. 15, 2015, now Pat. No. 11,041,851, which is a division of application No. 13/081,660, filed on Apr. 7, 2011, now Pat. No. 9,163,281.

(60) Provisional application No. 61/426,817, filed on Dec. 23, 2010.

(51) **Int. Cl.**
C12Q 1/6869 (2018.01)
G01N 33/543 (2006.01)

(52) **U.S. Cl.**
CPC ***G01N 33/543*** (2013.01); ***C12Q 1/6869*** (2013.01); *C12Q 2525/155* (2013.01); *C12Q 2525/179* (2013.01); *C12Q 2525/191* (2013.01); *G01N 2458/10* (2013.01)

(58) **Field of Classification Search**
CPC C12Q 1/6869; C12Q 2525/179
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,683,195 A 7/1987 Mullis et al.
4,683,202 A 7/1987 Mullis
4,988,617 A 1/1991 Landegren et al.
5,060,980 A 10/1991 Johnson et al.
5,210,015 A 5/1993 Gelfand et al.
5,234,809 A 8/1993 Boom et al.
5,242,794 A 9/1993 Whiteley et al.
5,348,853 A 9/1994 Wang et al.
5,434,049 A 7/1995 Okano et al.
5,459,307 A 10/1995 Klotz, Jr.
5,486,686 A 1/1996 Zdybel, Jr. et al.
5,491,224 A 2/1996 Bittner et al.
5,494,810 A 2/1996 Barany et al.
5,567,583 A 10/1996 Wang et al.
5,583,024 A 12/1996 McElroy et al.
5,604,097 A 2/1997 Brenner
5,636,400 A 6/1997 Young
5,674,713 A 10/1997 McElroy et al.
5,695,934 A 12/1997 Brenner
5,700,673 A 12/1997 McElroy et al.
5,701,256 A 12/1997 Marr et al.
5,830,064 A 11/1998 Bradish et al.
5,846,719 A 12/1998 Brenner et al.
5,863,722 A 1/1999 Brenner
5,866,337 A 2/1999 Schon
5,869,252 A 2/1999 Bouma et al.
5,869,717 A 2/1999 Frame et al.
5,871,921 A 2/1999 Landegren et al.
5,888,788 A 3/1999 De Miniac
5,942,391 A 8/1999 Zhang et al.
5,971,921 A 10/1999 Timbel
5,993,611 A 11/1999 Moroney, III et al.
5,994,056 A 11/1999 Higuchi
6,020,127 A 2/2000 Mackenzie et al.
6,033,854 A 3/2000 Kurait et al.
6,033,872 A 3/2000 Bergsma et al.
6,100,099 A 8/2000 Gordon et al.
6,138,077 A 10/2000 Brenner
6,150,516 A 11/2000 Brenner et al.
6,171,785 B1 1/2001 Higuchi
6,172,214 B1 1/2001 Brenner
6,172,218 B1 1/2001 Brenner
(Continued)

FOREIGN PATENT DOCUMENTS

EP 1 321 477 A1 6/2003
EP 1 564 306 A2 8/2005
(Continued)

OTHER PUBLICATIONS

Zhao, 2009, PGAAgenomics for comparative genome assembly based on genetic algorithm optimization, Genomics 94 (4):284-6.
Zheng, 2011, iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequences, BMC Bioinformatics 12:453.
Zhou, 2014, Bias from removing read duplication in ultra-deep sequencing experiments, Bioinformatics 30(8):1073-1080.
Zhulidov, 2004, Simple cDNA normalization using kamchatka crab duplex-specific nuclease, Nucl Acids Res 32(3): e37, 8 pages.
Zimmerman, 2010, A novel custom resequencing array for dilated cardiomyopathy, Gen Med 12(5):268-78.
(Continued)

Primary Examiner — Teresa E Strzelecka
(74) *Attorney, Agent, or Firm* — COOLEY LLP; Ivor R. Elrifi; Edith Hang Yu Cheng

(57) **ABSTRACT**

The invention generally relates to methods for maintaining the integrity and identification of a nucleic acid template in a multiplex sequencing reaction. In certain embodiments, methods of the invention involve obtaining a template nucleic acid, incorporating a pair of sequence identifiers into the template, and sequencing the template.

6 Claims, 3 Drawing Sheets

Illumina Exhibit 1001, Page 1 of 23

Ground	References	Basis	Claims
1	*Gloor* (EX1005)	§102	1-6
2	*Parameswaran* (EX1004)	§102	1-6
3	*Parameswaran* (EX1004) and *Gloor* (EX1005)	§103	1-6
4	*Parameswaran* (EX1004) and *Binladen* (EX1029)	§103	1-6
5	*Parameswaran* (EX1004), *Binladen* (EX1029), and *Bentley* (EX1006)	§103	6

Parameswaran Anticipates All Claims

1 [pre] A method for validating the sequence of a nucleic acid analyte of interest in a multiplex sequencing reaction, the method comprising

[a] **detecting** the presence of two or more identifier sequences that are uniquely associated with the nucleic acid analyte of interest,
wherein at least a first identifier sequence is incorporated into a 5′ portion of the nucleic acid analyte of interest and wherein at least a second identifier sequence is incorporated into a 3′ portion of the nucleic acid analyte of interest, and
wherein the first and second identifier sequences have four or more nucleotides and are different;

[b] **sequencing** the first identifier sequence, the nucleic acid analyte of interest and the second identifier sequence; and

[c] validating the sequence of the nucleic acid analyte of interest by analyzing both identifiers and excluding the sequences of those nucleic acid analytes of interest containing only one identifier or an incorrect pair of identifiers from sequence analysis of the multiplex sequencing reaction.

Parameswaran (EX1004, 3, 5, 7)

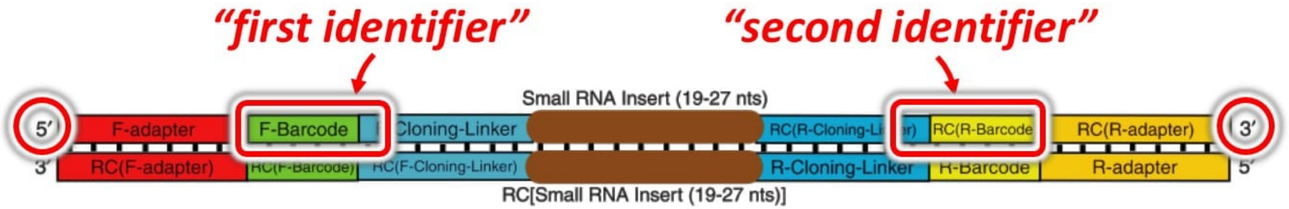
Table 2. List of possible barcodes

Forward barcodes	Reverse barcodes
AGCCTAAGCT	AGCTTAGGCT
AGTTCAAGTC	AGTCCAGGTC
ACTTGAACTG	ACTGGACCTG
ACGGTAACGT	ACGTTACCGT
ATCCGAATCG	ATCGGATTCG

* * *

dent datasets. These barcode + linker sequences are referred to herein as 'search motifs' or 'barcode motifs'.

* * *



Parameswaran Anticipates All Claims

1 [pre] A method for validating the sequence of a nucleic acid analyte of interest in a multiplex sequencing reaction, the method comprising

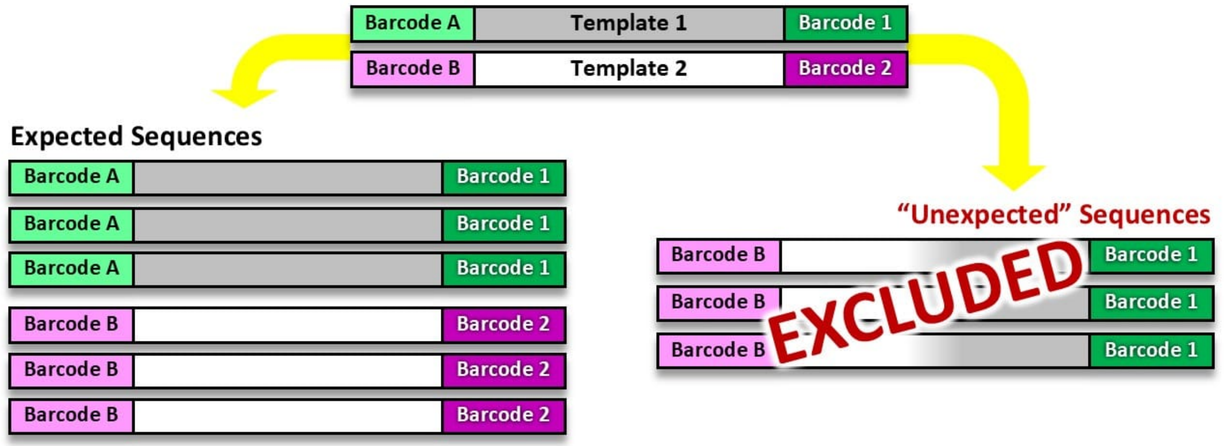
[a] detecting the presence of two or more identifier sequences that are uniquely associated with the nucleic acid analyte of interest,
wherein at least a first identifier sequence is incorporated into a 5′ portion of the nucleic acid analyte of interest and wherein at least a second identifier sequence is incorporated into a 3′ portion of the nucleic acid analyte of interest, and
wherein the first and second identifier sequences have four or more nucleotides and are different;

[b] sequencing the first identifier sequence, the nucleic acid analyte of interest and the second identifier sequence; and

[c] **validating** the sequence of the nucleic acid analyte of interest by **analyzing both identifiers** and **excluding the sequences** of those nucleic acid analytes of interest **containing only one identifier or an incorrect pair of identifiers** from sequence analysis of the multiplex sequencing reaction.

Parameswaran **(EX1004, 8)**

The false-discovery rate, defined by the percentage of 'perfect and unexpected' occurrences, was observed to decrease to <0.005% (Table 3), if three bases of the linker were used in addition to the 10-base barcode to assign sample identities. Thus a search motif of 13 nt



illumina

Patent Owner's Arguments on *Parameswaran* Should Be Rejected

1 [pre] A method for validating the sequence of a nucleic acid analyte of interest in a multiplex sequencing reaction, the method comprising

[a] detecting the presence of two or more identifier sequences that are uniquely associated with the nucleic acid analyte of interest, wherein at least a first identifier sequence is incorporated into a 5′ portion of the nucleic acid analyte of interest and wherein at least a second identifier sequence is incorporated into a 3′ portion of the nucleic acid analyte of interest, and wherein the first and second identifier sequences have four or more nucleotides and are different;

[b] sequencing the first identifier sequence, the nucleic acid analyte of interest and the second identifier sequence; and

[c] **validating** the sequence of the nucleic acid analyte of interest by **analyzing both identifiers** and **excluding the sequences** of those nucleic acid analytes of interest **containing only one identifier or an incorrect pair of identifiers** from sequence analysis of the multiplex sequencing reaction.

Patent Owner Response (POR, 23)

"Parameswaran includes the sequences in the dataset that Petitioner alleges Parameswaran somehow excluded." (Citing EX1004, 8).

***Parameswaran* (EX1004, 8)**

with reasonable accuracy into the correct bin. For our purposes, all sequences with a match to one of the two expected sequence motifs were considered as true members of the corresponding dataset. Run I (0.13%) and run II



Petitioner's Expert, Dr. Spellman (EX1054, ¶27)

This refers to "***expected*** sequence motifs," "not sequences with mismatched barcode motifs."

illumina

Parameswaran Anticipates All Claims

1 [pre] A method for validating the sequence of a nucleic acid analyte of interest in a multiplex sequencing reaction, the method comprising

[a] detecting the presence of two or more identifier sequences that are uniquely associated with the nucleic acid analyte of interest,
wherein at least a first identifier sequence is incorporated into a 5′ portion of the nucleic acid analyte of interest and wherein at least a second identifier sequence is incorporated into a 3′ portion of the nucleic acid analyte of interest, and
wherein the first and second identifier sequences have four or more nucleotides and are different;

[b] sequencing the first identifier sequence, the nucleic acid analyte of interest and the second identifier sequence; and

[c] **validating** the sequence of the nucleic acid analyte of interest by **analyzing both identifiers** and **excluding the sequences** of those nucleic acid analytes of interest **containing only one identifier** or **an incorrect pair of identifiers** from sequence analysis of the multiplex sequencing reaction.

Parameswaran (EX1004, 29)

D. SEARCH MOTIF = Barcode (10 nt) + Cloning-Linker (3 nt)

	454-66	454-67	454-68	454-69	454-70	454-71	454-72	454-73	454-74	454-75	454-76	454-77
454-66	13566	0	0	1	0	0	0	0	0	0	0	0
454-67	0	7459	0	3	0	0	0	0	0	0	0	0
454-68	0	0	13260	0	0	0	0	0	0	0	0	0
454-69	0	0	0	6518	0	0	0	0	0	0	0	0
454-70	0	0	0	0	9484	0	0	0	0	0	0	0
454-71	0	0	0	0	0	9397	0	0	0	0	0	0
454-72	0	0	0	0	0	0	12302	0	0	0	0	0
454-73	0	0	0	0	0	0	0	7273	0	0	0	0



Petitioner's Expert, Dr. Spellman (EX2015, 109:13-21)

Q. Can you show me where it talks about grey boxes?

A. I will point you to the abstract, the very last sentence, "The false discovery rate as measured by the percentage of sequences with unexpected perfect pairings . . . " That is exactly the grey cells in this table. . . .

Illumina

Patent Owner's Arguments on Claim 5 Should Be Rejected

5 The method of claim 4, wherein t [sic] barcode sequences are **designed to have no homopolymer regions**.

Patent Owner Response (POR, 25)

"Parameswaran specifically designed ***every one*** of its barcodes to ***include*** two homopolymer regions."

Parameswaran **(EX1004, 4)**

reads used per barcode. The barcode design (Table 1) facilitated unambiguous assignment, with care taken to avoid homopolymeric runs that are particularly problematic in pyrosequencing technologies. All forward

In re Applied Materials, Inc.,
692 F.3d 1289, 1298 (Fed. Cir. 2012)

"A reference must be considered for everything that it teaches, not simply the described invention or a preferred embodiment."

illumina

Patent Owner's Arguments on Claim 6 Should Be Rejected

6 The method of claim 1, wherein the nucleic acid analyte of interest is **on the surface of a flow cell**.

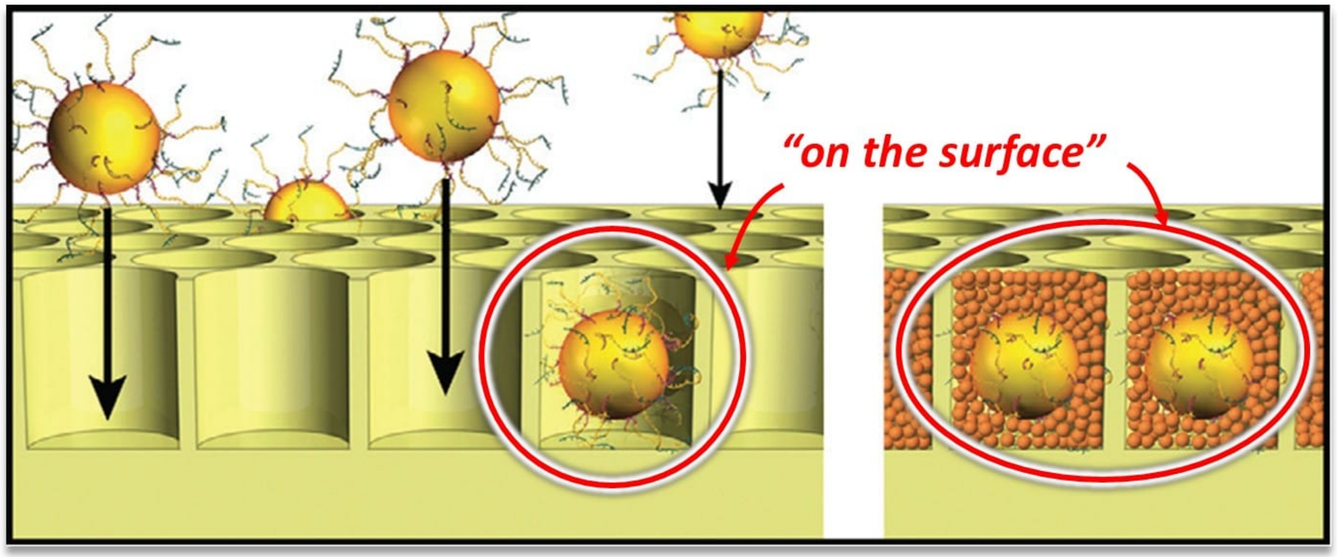
Patent Owner Response (POR, 26)

"In Roche 454 pyrosequencing used in Parameswaran, the analyte of interest is not ***on the surface*** of the flow cell, but suspended above the flow cell."



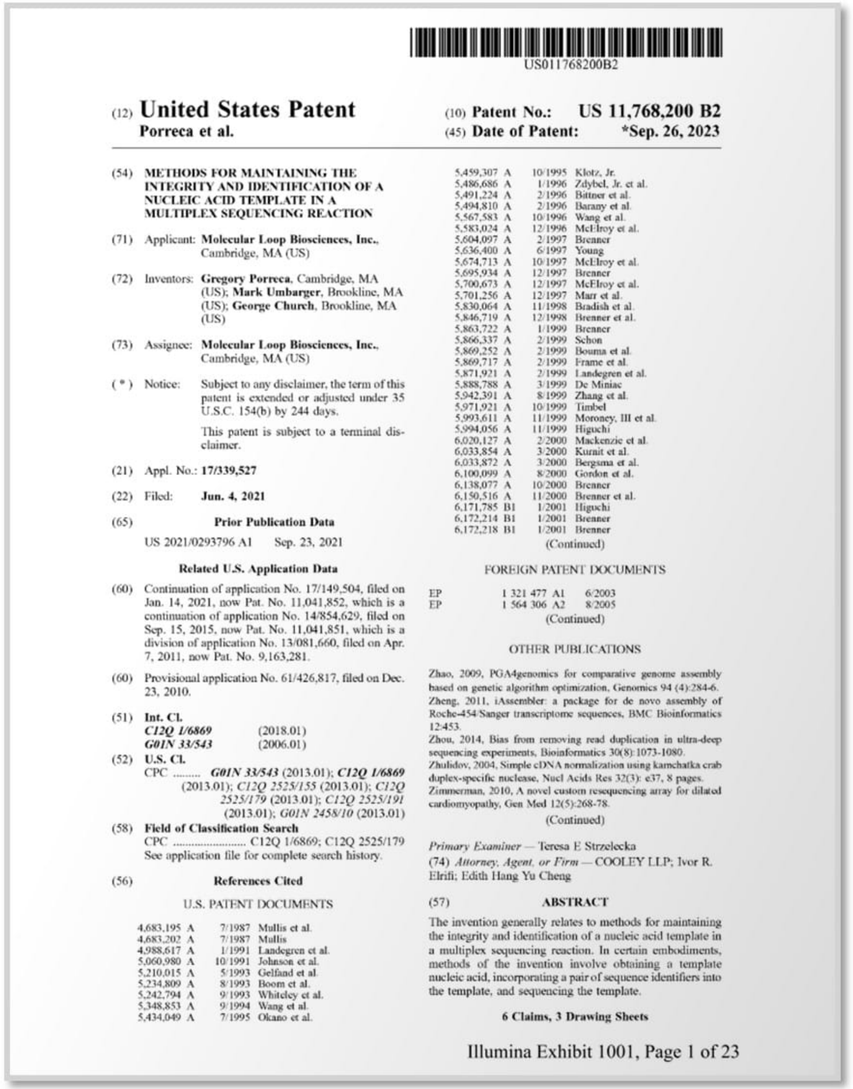
Petitioner's Expert, Dr. Spellman (EX1054, ¶39; EX1028, 8)

Claim 6 "does ***not*** require the analyte of interest to be attached to the surface of a flow cell. . . . It only requires the analyte to be 'on the surface of a flow cell.'"



illumina

Ground 3: *Parameswaran* and *Gloor*



US011768200B2

(12) **United States Patent**
Porreca et al.

(10) **Patent No.: US 11,768,200 B2**
(45) **Date of Patent: *Sep. 26, 2023**

(54) **METHODS FOR MAINTAINING THE INTEGRITY AND IDENTIFICATION OF A NUCLEIC ACID TEMPLATE IN A MULTIPLEX SEQUENCING REACTION**

(71) Applicant: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(72) Inventors: **Gregory Porreca**, Cambridge, MA (US); **Mark Umbarger**, Brookline, MA (US); **George Church**, Brookline, MA (US)

(73) Assignee: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 244 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **17/339,527**

(22) Filed: **Jun. 4, 2021**

(65) **Prior Publication Data**

US 2021/0293796 A1 Sep. 23, 2021

Related U.S. Application Data

(60) Continuation of application No. 17/149,504, filed on Jan. 14, 2021, now Pat. No. 11,041,852, which is a continuation of application No. 14/854,629, filed on Sep. 15, 2015, now Pat. No. 11,041,851, which is a division of application No. 13/081,660, filed on Apr. 7, 2011, now Pat. No. 9,163,281.

(60) Provisional application No. 61/426,817, filed on Dec. 23, 2010.

(51) **Int. Cl.**
C12Q 1/6869 (2018.01)
G01N 33/543 (2006.01)

(52) **U.S. Cl.**
CPC ***G01N 33/543*** (2013.01); ***C12Q 1/6869*** (2013.01); *C12Q 2525/155* (2013.01); *C12Q 2525/179* (2013.01); *C12Q 2525/191* (2013.01); *G01N 2458/10* (2013.01)

(58) **Field of Classification Search**
CPC C12Q 1/6869; C12Q 2525/179
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,683,195 A 7/1987 Mullis et al.
4,683,202 A 7/1987 Mullis
4,988,617 A 1/1991 Landegren et al.
5,060,980 A 10/1991 Johnson et al.
5,210,015 A 5/1993 Gelfand et al.
5,234,809 A 8/1993 Boom et al.
5,242,794 A 9/1993 Whiteley et al.
5,348,853 A 9/1994 Wang et al.
5,434,049 A 7/1995 Okano et al.
5,459,307 A 10/1995 Klotz, Jr.
5,486,686 A 1/1996 Zdybel, Jr. et al.
5,491,224 A 2/1996 Bittner et al.
5,494,810 A 2/1996 Barany et al.
5,567,583 A 10/1996 Wang et al.
5,583,024 A 12/1996 McElroy et al.
5,604,097 A 2/1997 Brenner
5,636,400 A 6/1997 Young
5,674,713 A 10/1997 McElroy et al.
5,695,934 A 12/1997 Brenner
5,700,673 A 12/1997 McElroy et al.
5,701,256 A 12/1997 Marr et al.
5,830,064 A 11/1998 Bradish et al.
5,846,719 A 12/1998 Brenner et al.
5,863,722 A 1/1999 Brenner
5,866,337 A 2/1999 Schon
5,869,252 A 2/1999 Bouma et al.
5,869,717 A 2/1999 Frame et al.
5,871,921 A 2/1999 Landegren et al.
5,888,788 A 3/1999 De Mintac
5,942,391 A 8/1999 Zhang et al.
5,971,921 A 10/1999 Timbel
5,993,611 A 11/1999 Moroney, III et al.
5,994,056 A 11/1999 Higuchi
6,020,127 A 2/2000 Mackenzie et al.
6,033,854 A 3/2000 Kurait et al.
6,033,872 A 3/2000 Bergsma et al.
6,100,099 A 8/2000 Gordon et al.
6,138,077 A 10/2000 Brenner
6,150,516 A 11/2000 Brenner et al.
6,171,785 B1 1/2001 Higuchi
6,172,214 B1 1/2001 Brenner
6,172,218 B1 1/2001 Brenner
(Continued)

FOREIGN PATENT DOCUMENTS

EP 1 321 477 A1 6/2003
EP 1 564 306 A2 8/2005
(Continued)

OTHER PUBLICATIONS

Zhao, 2009, PGAAgenomics for comparative genome assembly based on genetic algorithm optimization, Genomics 94 (4):284-6.
Zheng, 2011, iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequences, BMC Bioinformatics 12:453.
Zhou, 2014, Bias from removing read duplication in ultra-deep sequencing experiments, Bioinformatics 30(8):1073-1080.
Zhulidov, 2004, Simple cDNA normalization using kamchatka crab duplex-specific nuclease, Nucl Acids Res 32(3): e37, 8 pages.
Zimmerman, 2010, A novel custom resequencing array for dilated cardiomyopathy, Gen Med 12(5):268-78.
(Continued)

Primary Examiner — Teresa E Strzelecka
(74) *Attorney, Agent, or Firm* — COOLEY LLP; Ivor R. Elrifi; Edith Hang Yu Cheng

(57) **ABSTRACT**

The invention generally relates to methods for maintaining the integrity and identification of a nucleic acid template in a multiplex sequencing reaction. In certain embodiments, methods of the invention involve obtaining a template nucleic acid, incorporating a pair of sequence identifiers into the template, and sequencing the template.

6 Claims, 3 Drawing Sheets

Illumina Exhibit 1001, Page 1 of 23

Ground	References	Basis	Claims
1	*Gloor* (EX1005)	§102	1-6
2	*Parameswaran* (EX1004)	§102	1-6
3	***Parameswaran* (EX1004) and *Gloor* (EX1005)**	**§103**	**1-6**
4	*Parameswaran* (EX1004) and *Binladen* (EX1029)	§103	1-6
5	*Parameswaran* (EX1004), *Binladen* (EX1029), and *Bentley* (EX1006)	§103	6

Parameswaran and *Gloor* Render All Claims Obvious

Parameswaran
(EX1004)



Published online 11 October 2007 Nucleic Acids Research, 2007, Vol. 35, No. 19 e130
doi:10.1093/nar/gkm760

A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing

Poornima Parameswaran[1], Roxana Jalili[2], Li Tao[2], Shadi Shokralla[2], Baback Gharizadeh[2], Mostafa Ronaghi[2] and Andrew Z. Fire[2,*]

[1]Department of Microbiology and Immunology, [2]Department of Pathology and Department of Genetics, Stanford University School of Medicine, [3]Stanford Genome Technology Center and [4]Department of Biological Sciences, Stanford University, Stanford, CA-94305, USA

Received July 30, 2007; Revised September 10, 2007; Accepted September 11, 2007

ABSTRACT

Multiplexed high-throughput pyrosequencing is currently limited in complexity (number of samples sequenced in parallel), and in capacity (number of sequences obtained per sample). Physical-space segregation of the sequencing platform into a fixed number of channels allows limited multiplexing, but obscures available sequencing space. To overcome these limitations, we have devised a novel barcoding approach to allow for pooling and sequencing of DNA from independent samples, and to facilitate subsequent segregation of sequencing capacity. Forty-eight forward-reverse barcode pairs are described: each forward and each reverse barcode unique with respect to at least 4 nt positions. With improved read lengths of pyrosequencers, combinations of forward and reverse barcodes may be used to sequence from as many as n^2 independent libraries for each set of 'n' forward and 'n' reverse barcodes, for each defined set of cloning-linkers. In two pilot series of barcoded sequencing using the GS20 Sequencer (454/Roche), we found that over 95.8% of obtained sequences could be assigned to 25 independent, uniquely barcoded libraries based on the presence of either a perfect forward or a perfect reverse barcode. The false-discovery rate, as measured by the percentage of sequences with unexpected perfect pairings of unmatched forward and reverse barcodes, was estimated to be <0.005%.

INTRODUCTION

With the advent of high-throughput sequencing, characterization of the nucleic acid world is proceeding at an accelerated pace. Three major high-throughput sequencing platforms are in use today: the Genome Sequencers from Roche/454 Life Sciences (GS-20 or GS-FLX; (1)), the 1G Analyzer from Illumina/Solexa (2) and the SOLiD System from Applied Biosystems (http://solid.appliedbiosystems.com). Comparison across the three platforms reveals a trade-off between average sequence read length and the number of DNA molecules that are sequenced. At present, the Solexa and SOLiD systems provide many more sequence reads, but render much shorter read lengths than the 454/Roche Genome Sequencers. This makes the 454/Roche platform appealing for use with barcoding technology, as the enhanced read length facilitates the unambiguous identification of both complex barcodes and sequences of interest.

The 454/Roche Genome Sequencers are called pyrosequencers because their sequencing technology is based on the detection of pyrophosphates released during DNA synthesis (3). A few sequencing runs using 454/Roche's pyrosequencing platform can generate sufficient coverage for assembling entire microbial genomes (4,5), for the discovery, identification and quantitation of small RNAs (6,7), and for the detection of rare variations in cancers (8), among many other applications. For analysis of multiple libraries, the currently available 454/Roche pyrosequencers can accommodate a maximum of only 16 independent samples, which have to be physically separated using manifolds on the sequencing medium. These separation manifolds occlude wells on the

*To whom correspondence should be addressed. Tel: +1 650 723 2885; Fax: +1 650 724 9650; Email: allez@stanford.edu

© 2007 The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Illumina Exhibit 1004, Page 1 of 31

+

Gloor
(EX1005)



arXiv:1007.5075v1 [q-bio.GN] 28 Jul 2010

Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products

Gregory B. Gloor[1,*], Ruben Hummelen[2,3], Jean M. Macklaim[1,2], Russell J. Dickson[4] Andrew D. Fernandes[1,4], Roderick MacPhee[2,3], Gregor Reid[1,2,3,5,6]

1 Department of Biochemistry, University of Western Ontario, London, ON, Canada
2 Canadian Research & Development Centre for Probiotics, Lawson Health Research Institute, London, Canada
3 Department of Public Health, Erasmus MC, University Medical Centre Rotterdam, The Netherlands
4 Department Applied Mathematics
5 Department of Microbiology & Immunology, University of Western Ontario, London, ON, Canada
6 Department of Surgery, University of Western Ontario, London, ON, Canada
* E-mail: ggloor@uwo.ca

Abstract

We developed a low-cost, high-throughput microbiome profiling method that uses combinatorial sequence tags attached to PCR primers that amplify the rRNA V6 region. Amplified PCR products are sequenced using an Illumina paired-end protocol to generate millions of overlapping reads. Combinatorial sequence tagging can be used to examine hundreds of samples with far fewer primers than is required when sequence tags are incorporated at only a single end. The number of reads generated permitted saturating or near-saturating analysis of samples of the vaginal microbiome. The large number of reads allowed an in-depth analysis of errors, and we found that PCR-induced errors composed the vast majority of non-organism derived species variants, an observation that has significant implications for sequence clustering of similar high-throughput data. We show that the short reads are sufficient to assign organisms to the genus or species level in most cases. We suggest that this method will be useful for the deep sequencing of any short nucleotide region that is taxonomically informative; those include the V3, V5 regions of the bacterial 16S rRNA genes and the eukaryote V9 region that is gaining popularity for sampling protist diversity.

Introduction

Microbiome profiling is used to identify and enumerate the organisms in samples from diverse sources such as soil, clinical samples and oceanic environments [4, 22, 29]. This profiling is an important first step in determining the important bacterial and protist organisms in a biome and how they interact with and influence their environment.

1

Illumina Exhibit 1005, Page 1 of 28

Motivation to Combine

To Reduce Technical Errors (e.g., "Only One Identifier") and Chimeras (e.g., "Incorrect Pair of Identifiers") Formed During Multiplex Sequencing (Pet., 48-49)

Illumina Sequencers Were Faster, More Efficient, and Less Expensive (Pet., 50)

***Parameswaran*'s Recommendation** (Pet., 50-51)

Reasonable Expectation of Success

Interchangeability Between Roche Pyrosequencing and Illumina Sequencing (Pet., 51-53)

illumina

Patent Owner Fails to Rebut Obviousness

1 [pre]	A method for validating the sequence of a nucleic acid analyte of interest in a multiplex sequencing reaction, the method comprising
[a]	**detecting** the presence of two or more identifier sequences that are uniquely associated with the nucleic acid analyte of interest, wherein at least a first identifier sequence is incorporated into a 5′ portion of the nucleic acid analyte of interest and wherein at least a second identifier sequence is incorporated into a 3′ portion of the nucleic acid analyte of interest, and wherein the first and second identifier sequences have four or more nucleotides and are different;
[b]	**sequencing** the first identifier sequence, the nucleic acid analyte of interest and the second identifier sequence; and
[c]	**validating** the sequence of the nucleic acid analyte of interest by **analyzing both identifiers** and **excluding the sequences** of those nucleic acid analytes of interest **containing only one identifier or an incorrect pair of identifiers** from sequence analysis of the multiplex sequencing reaction.

No Reference to "Cross-Over Error" (Reply, 9)

Patent Owner Response (POR, 36)

"Petitioner overstates the awareness of problems arising from chimeric sequences, such as cross-over error, prior to the invention."

Excludes Sequences With "Only One Identifier" *or* an "Incorrect Pair of Identifiers" (Reply, 2-3, 9)

illumina

EX1001 ('200 Pat.), cl. 1; EX1054, ¶¶10-11, 40-42; POR, 36-42; Reply, 2-3, 9.

Patent Owner Fails to Rebut Obviousness

1 [pre]	A method for validating the sequence of a nucleic acid analyte of interest in a multiplex sequencing reaction, the method comprising
[a]	**detecting** the presence of two or more identifier sequences that are uniquely associated with the nucleic acid analyte of interest, wherein at least a first identifier sequence is incorporated into a 5′ portion of the nucleic acid analyte of interest and wherein at least a second identifier sequence is incorporated into a 3′ portion of the nucleic acid analyte of interest, and wherein the first and second identifier sequences have four or more nucleotides and are different;
[b]	**sequencing** the first identifier sequence, the nucleic acid analyte of interest and the second identifier sequence; and
[c]	**validating** the sequence of the nucleic acid analyte of interest by **analyzing both identifiers** and **excluding the sequences** of those nucleic acid analytes of interest **containing only one identifier or an incorrect pair of identifiers** from sequence analysis of the multiplex sequencing reaction.

Not Limited to Unique Dual-Indexing (Reply, 2-4)

Patent Owner Response (POR, 44)

There "were downsides to simply adding [**unique dual-indexes**] to Gloor's sequencing with [combinatorial dual indexes] on **an Illumina sequencer**."



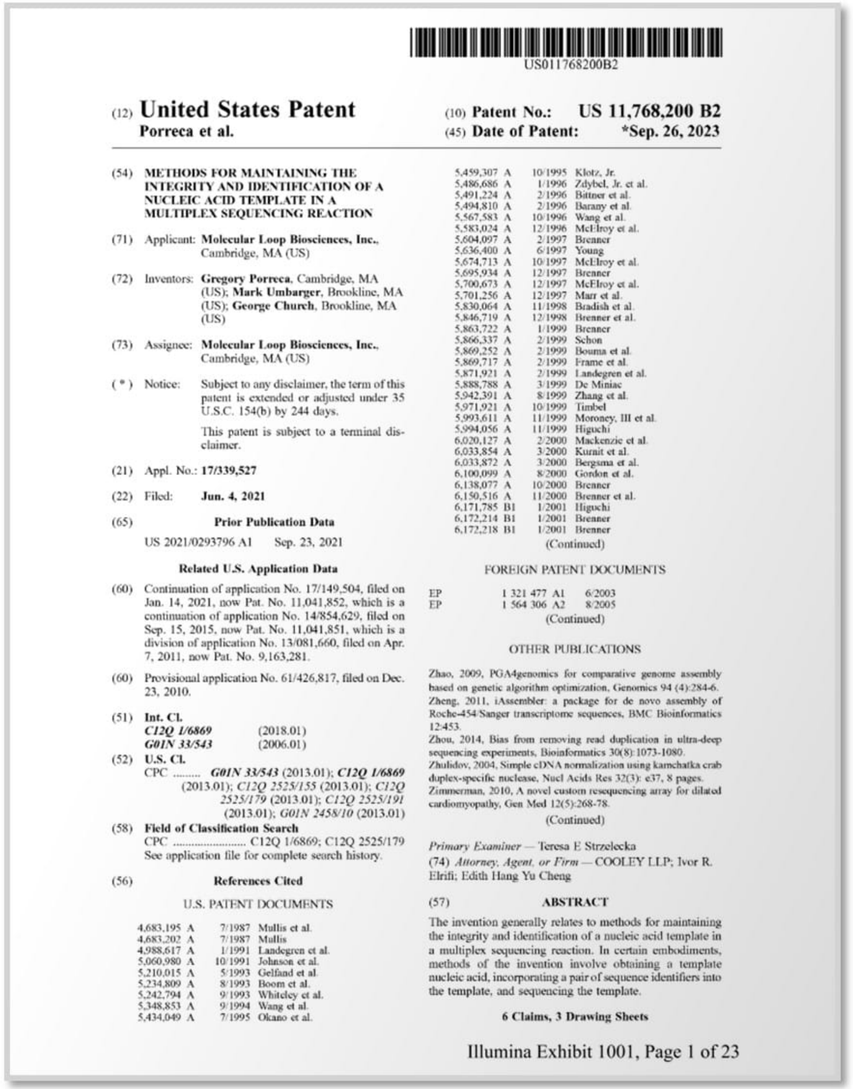
Petitioner's Expert, Dr. Spellman (EX1054, ¶92)

A "POSA would have understood that **sample tracking and sequence fidelity are critical in many situations**," including cancer research, ancient DNA research, and molecular diagnosis

illumina

EX1001 ('200 Pat.), cl. 1; EX1054, ¶¶9-16, 90-93, 104-105; POR, 44-45; Reply, 2-4, 18-19.

Ground 4: *Parameswaran* and *Binladen*



US011768200B2

(12) **United States Patent**
Porreca et al.

(10) **Patent No.: US 11,768,200 B2**
(45) **Date of Patent: *Sep. 26, 2023**

(54) **METHODS FOR MAINTAINING THE INTEGRITY AND IDENTIFICATION OF A NUCLEIC ACID TEMPLATE IN A MULTIPLEX SEQUENCING REACTION**

(71) Applicant: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(72) Inventors: **Gregory Porreca**, Cambridge, MA (US); **Mark Umbarger**, Brookline, MA (US); **George Church**, Brookline, MA (US)

(73) Assignee: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 244 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **17/339,527**

(22) Filed: **Jun. 4, 2021**

(65) **Prior Publication Data**

US 2021/0293796 A1 Sep. 23, 2021

Related U.S. Application Data

(60) Continuation of application No. 17/149,504, filed on Jan. 14, 2021, now Pat. No. 11,041,852, which is a continuation of application No. 14/854,629, filed on Sep. 15, 2015, now Pat. No. 11,041,851, which is a division of application No. 13/081,660, filed on Apr. 7, 2011, now Pat. No. 9,163,281.

(60) Provisional application No. 61/426,817, filed on Dec. 23, 2010.

(51) **Int. Cl.**
C12Q 1/6869 (2018.01)
G01N 33/543 (2006.01)

(52) **U.S. Cl.**
CPC ***G01N 33/543*** (2013.01); ***C12Q 1/6869*** (2013.01); *C12Q 2525/155* (2013.01); *C12Q 2525/179* (2013.01); *C12Q 2525/191* (2013.01); *G01N 2458/10* (2013.01)

(58) **Field of Classification Search**
CPC C12Q 1/6869; C12Q 2525/179
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,683,195	A	7/1987	Mullis et al.
4,683,202	A	7/1987	Mullis
4,988,617	A	1/1991	Landegren et al.
5,060,980	A	10/1991	Johnson et al.
5,210,015	A	5/1993	Gelfand et al.
5,234,809	A	8/1993	Boom et al.
5,242,794	A	9/1993	Whiteley et al.
5,348,853	A	9/1994	Wang et al.
5,434,049	A	7/1995	Okano et al.
5,459,307	A	10/1995	Klotz, Jr.
5,486,686	A	1/1996	Zdybel, Jr. et al.
5,491,224	A	2/1996	Bittner et al.
5,494,810	A	2/1996	Barany et al.
5,567,583	A	10/1996	Wang et al.
5,583,024	A	12/1996	McElroy et al.
5,604,097	A	2/1997	Brenner
5,636,400	A	6/1997	Young
5,674,713	A	10/1997	McElroy et al.
5,695,934	A	12/1997	Brenner
5,700,673	A	12/1997	McElroy et al.
5,701,256	A	12/1997	Marr et al.
5,830,064	A	11/1998	Bradish et al.
5,846,719	A	12/1998	Brenner et al.
5,863,722	A	1/1999	Brenner
5,866,337	A	2/1999	Schon
5,869,252	A	2/1999	Bouma et al.
5,869,717	A	2/1999	Frame et al.
5,871,921	A	2/1999	Landegren et al.
5,888,788	A	3/1999	De Mintac
5,942,391	A	8/1999	Zhang et al.
5,971,921	A	10/1999	Timbel
5,993,611	A	11/1999	Moroney, III et al.
5,994,056	A	11/1999	Higuchi
6,020,127	A	2/2000	Mackenzie et al.
6,033,854	A	3/2000	Kurait et al.
6,033,872	A	3/2000	Bergsma et al.
6,100,099	A	8/2000	Gordon et al.
6,138,077	A	10/2000	Brenner
6,150,516	A	11/2000	Brenner et al.
6,171,785	B1	1/2001	Higuchi
6,172,214	B1	1/2001	Brenner
6,172,218	B1	1/2001	Brenner

(Continued)

FOREIGN PATENT DOCUMENTS

EP	1 321 477	A1	6/2003
EP	1 564 306	A2	8/2005

(Continued)

OTHER PUBLICATIONS

Zhao, 2009, PGAAgenomics for comparative genome assembly based on genetic algorithm optimization, Genomics 94 (4):284-6.
Zheng, 2011, iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequences, BMC Bioinformatics 12:453.
Zhou, 2014, Bias from removing read duplication in ultra-deep sequencing experiments, Bioinformatics 30(8):1073-1080.
Zhulidov, 2004, Simple cDNA normalization using kamchatka crab duplex-specific nuclease, Nucl Acids Res 32(3): e37, 8 pages.
Zimmerman, 2010, A novel custom resequencing array for dilated cardiomyopathy, Gen Med 12(5):268-78.

(Continued)

Primary Examiner — Teresa E Strzelecka
(74) *Attorney, Agent, or Firm* — COOLEY LLP; Ivor R. Elrifi; Edith Hang Yu Cheng

(57) **ABSTRACT**

The invention generally relates to methods for maintaining the integrity and identification of a nucleic acid template in a multiplex sequencing reaction. In certain embodiments, methods of the invention involve obtaining a template nucleic acid, incorporating a pair of sequence identifiers into the template, and sequencing the template.

6 Claims, 3 Drawing Sheets

Illumina Exhibit 1001, Page 1 of 23

Ground	References	Basis	Claims
1	*Gloor* (EX1005)	§102	1-6
2	*Parameswaran* (EX1004)	§102	1-6
3	*Parameswaran* (EX1004) and *Gloor* (EX1005)	§103	1-6
4	*Parameswaran* (EX1004) and *Binladen* (EX1029)	§103	1-6
5	*Parameswaran* (EX1004), *Binladen* (EX1029), and *Bentley* (EX1006)	§103	6

illumina

Parameswaran and *Binladen* Render All Claims Obvious

Binladen's Dual-Indexing (EX1029, 7)

5'Tag	Sequences from forward primer	Sequences from reverse primer	Total sequences	Expected sequence frequency	Deviation
		* * *			
4A*	45 (gcta)	60 (gtca)	105	183,16	−78,16
4B*	48 (tcag)	50 (tgac)	98	274,75	−176,75
4C*	114 (ctag)	100 (tagc)	214	274,75	−60,75

Binladen's Erroneous Sequences (EX1029, 5-6)

Initial assignment of the sequence data

As the correct association of tags and sequences is crucial to the approach, we followed very conservative criteria post sequencing in regards to acceptance of the sequence data. Initially, we discarded all sequence reads without an exact match to any of the primers used in the studies (Primer Mismatched Sequences).

* * *

primer sequences in general (n = 2); Those that were exact match to the core primer but lacked the 5′ tag sequence (n = 121); Those



Petitioner's Expert, Dr. Spellman (EX1054, ¶125)

"By excluding sequence reads lacking an identifier, *Binladen* necessarily excluded the sequences of analytes containing only one identifier."

Motivated, With a Reasonable Expectation of Success, To Reduce Errors Formed During Multiplex Sequencing by Excluding Sequences With One Identifier (*Binladen*) or an Incorrect Pair of Identifiers (*Parameswaran*) (Pet., 62-64)

Patent Owner's Arguments on *Binladen* Should Be Rejected

1 [pre]	A method for validating the sequence of a nucleic acid analyte of interest in a multiplex sequencing reaction, the method comprising
[a]	**detecting** the presence of two or more identifier sequences that are uniquely associated with the nucleic acid analyte of interest, wherein at least a first identifier sequence is incorporated into a 5′ portion of the nucleic acid analyte of interest and wherein at least a second identifier sequence is incorporated into a 3′ portion of the nucleic acid analyte of interest, and wherein the first and second identifier sequences have four or more nucleotides and are different;
[b]	**sequencing** the first identifier sequence, the nucleic acid analyte of interest and the second identifier sequence; and
[c]	**validating** the sequence of the nucleic acid analyte of interest by **analyzing both identifiers** and **excluding the sequences** of those nucleic acid analytes of interest **containing only one identifier or an incorrect pair of identifiers** from sequence analysis of the multiplex sequencing reaction.

Patent Owner Response (POR, 47)

"Binladen does not analyze ***both identifiers*** and exclude sequences as required by claim 1."



Petitioner's Expert, Dr. Spellman (EX1054, ¶123)

"Claim 1 of the '200 patent does not require that the two identifiers of a given analyte of interest must be sequenced in a single read or analyzed from a single read."

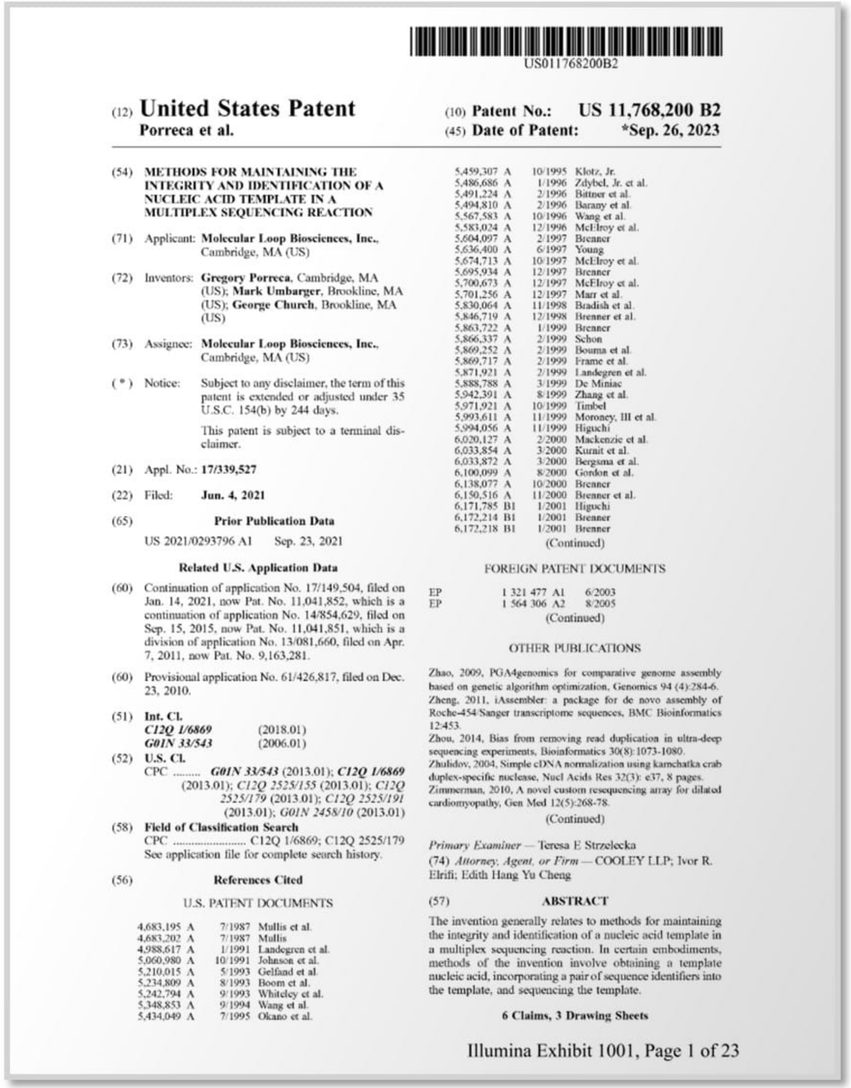
***Binladen* (EX1029, 3)**

details). The implication of this is that approximately 50% of each PCR product will be sequenced from the orientation of the forward primer, and 50% from the orientation of the reverse primer. Hence, this made it necessary to label both the forward

illumina

EX1001 ('200 Pat.), cl. 1; EX1029, 3, 5-7; EX1054, ¶¶120-126; POR, 47-49; Reply, 24.

Ground 5: *Parameswaran*, *Binladen*, and *Bentley*



US011768200B2

(12) **United States Patent**
Porreca et al.

(10) **Patent No.: US 11,768,200 B2**
(45) **Date of Patent: *Sep. 26, 2023**

(54) **METHODS FOR MAINTAINING THE INTEGRITY AND IDENTIFICATION OF A NUCLEIC ACID TEMPLATE IN A MULTIPLEX SEQUENCING REACTION**

(71) Applicant: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(72) Inventors: **Gregory Porreca**, Cambridge, MA (US); **Mark Umbarger**, Brookline, MA (US); **George Church**, Brookline, MA (US)

(73) Assignee: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 244 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **17/339,527**

(22) Filed: **Jun. 4, 2021**

(65) **Prior Publication Data**

US 2021/0293796 A1 Sep. 23, 2021

Related U.S. Application Data

(60) Continuation of application No. 17/149,504, filed on Jan. 14, 2021, now Pat. No. 11,041,852, which is a continuation of application No. 14/854,629, filed on Sep. 15, 2015, now Pat. No. 11,041,851, which is a division of application No. 13/081,660, filed on Apr. 7, 2011, now Pat. No. 9,163,281.

(60) Provisional application No. 61/426,817, filed on Dec. 23, 2010.

(51) **Int. Cl.**
C12Q 1/6869 (2018.01)
G01N 33/543 (2006.01)

(52) **U.S. Cl.**
CPC ***G01N 33/543*** (2013.01); ***C12Q 1/6869*** (2013.01); *C12Q 2525/155* (2013.01); *C12Q 2525/179* (2013.01); *C12Q 2525/191* (2013.01); *G01N 2458/10* (2013.01)

(58) **Field of Classification Search**
CPC C12Q 1/6869; C12Q 2525/179
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,683,195 A 7/1987 Mullis et al.
4,683,202 A 7/1987 Mullis
4,988,617 A 1/1991 Landegren et al.
5,060,980 A 10/1991 Johnson et al.
5,210,015 A 5/1993 Gelfand et al.
5,234,809 A 8/1993 Boom et al.
5,242,794 A 9/1993 Whiteley et al.
5,348,853 A 9/1994 Wang et al.
5,434,049 A 7/1995 Okano et al.
5,459,307 A 10/1995 Klotz, Jr.
5,486,686 A 1/1996 Zdybel, Jr. et al.
5,491,224 A 2/1996 Bittner et al.
5,494,810 A 2/1996 Barany et al.
5,567,583 A 10/1996 Wang et al.
5,583,024 A 12/1996 McElroy et al.
5,604,097 A 2/1997 Brenner
5,636,400 A 6/1997 Young
5,674,713 A 10/1997 McElroy et al.
5,695,934 A 12/1997 Brenner
5,700,673 A 12/1997 McElroy et al.
5,701,256 A 12/1997 Marr et al.
5,830,064 A 11/1998 Bradish et al.
5,846,719 A 12/1998 Brenner et al.
5,863,722 A 1/1999 Brenner
5,866,337 A 2/1999 Schon
5,869,252 A 2/1999 Bouma et al.
5,869,717 A 2/1999 Frame et al.
5,871,921 A 2/1999 Landegren et al.
5,888,788 A 3/1999 De Mintac
5,942,391 A 8/1999 Zhang et al.
5,971,921 A 10/1999 Timbel
5,993,611 A 11/1999 Moroney, III et al.
5,994,056 A 11/1999 Higuchi
6,020,127 A 2/2000 Mackenzie et al.
6,033,854 A 3/2000 Kurait et al.
6,033,872 A 3/2000 Bergsma et al.
6,100,099 A 8/2000 Gordon et al.
6,138,077 A 10/2000 Brenner
6,150,516 A 11/2000 Brenner et al.
6,171,785 B1 1/2001 Higuchi
6,172,214 B1 1/2001 Brenner
6,172,218 B1 1/2001 Brenner
(Continued)

FOREIGN PATENT DOCUMENTS

EP 1 321 477 A1 6/2003
EP 1 564 306 A2 8/2005
(Continued)

OTHER PUBLICATIONS

Zhao, 2009, PGAAgenomics for comparative genome assembly based on genetic algorithm optimization, Genomics 94 (4):284-6.
Zheng, 2011, iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequences, BMC Bioinformatics 12:453.
Zhou, 2014, Bias from removing read duplication in ultra-deep sequencing experiments, Bioinformatics 30(8):1073-1080.
Zhulidov, 2004, Simple cDNA normalization using kamchatka crab duplex-specific nuclease, Nucl Acids Res 32(3): e37, 8 pages.
Zimmerman, 2010, A novel custom resequencing array for dilated cardiomyopathy, Gen Med 12(5):268-78.
(Continued)

Primary Examiner — Teresa E Strzelecka
(74) *Attorney, Agent, or Firm* — COOLEY LLP; Ivor R. Elrifi; Edith Hang Yu Cheng

(57) **ABSTRACT**

The invention generally relates to methods for maintaining the integrity and identification of a nucleic acid template in a multiplex sequencing reaction. In certain embodiments, methods of the invention involve obtaining a template nucleic acid, incorporating a pair of sequence identifiers into the template, and sequencing the template.

6 Claims, 3 Drawing Sheets

Illumina Exhibit 1001, Page 1 of 23

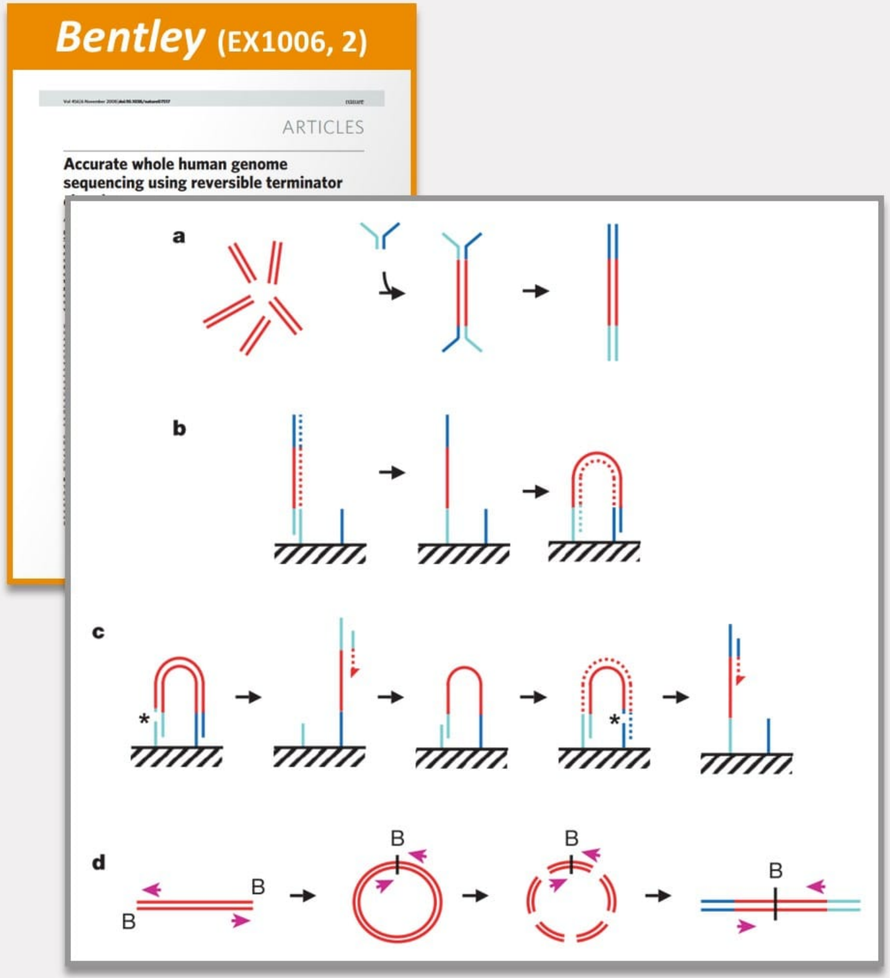
Ground	References	Basis	Claims
1	*Gloor* (EX1005)	§102	1-6
2	*Parameswaran* (EX1004)	§102	1-6
3	*Parameswaran* (EX1004) and *Gloor* (EX1005)	§103	1-6
4	*Parameswaran* (EX1004) and *Binladen* (EX1029)	§103	1-6
5	*Parameswaran* (EX1004), *Binladen* (EX1029), and *Bentley* (EX1006)	§103	6

illumina

Parameswaran, Binladen, and *Bentley* Render Claim 6 Obvious

6 The method of claim 1, wherein the nucleic acid analyte of interest is **on the surface of a flow cell**.

Bentley **(EX1006, 2)**





Petitioner's Expert, Dr. Spellman **(EX1003, ¶¶252-253)**

Claim 6 "does not require attaching the analyte to the surface of a flow cell, instead only requiring the nucleic acid analyte of interest to be 'on the surface of a flow cell.'"

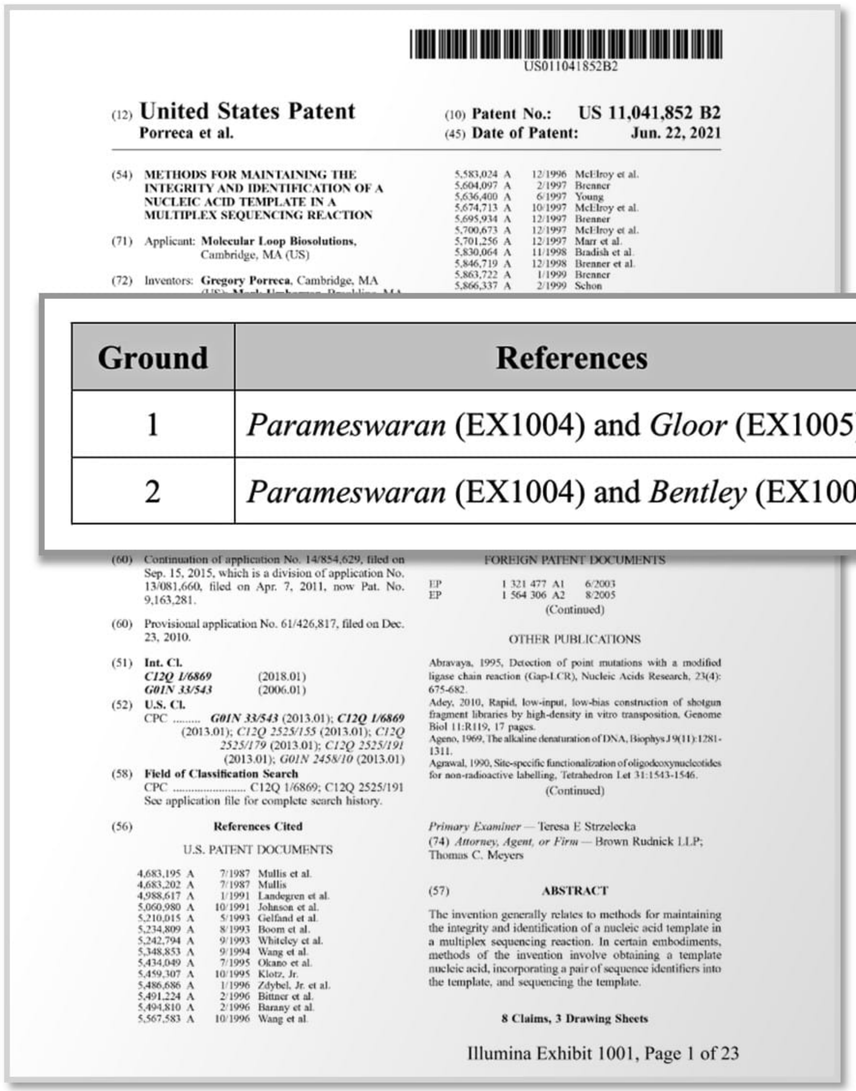
"Even if the claim required attaching analytes to the surface of a flow cell, *Bentley* discloses this limitation."

Motivated, With a Reasonable Expectation of Success, to Use Illumina Sequencers, Which Were Faster, More Efficient, and Less Expensive **(Pet., 66-67)**

illumina

EX1001 ('200 Pat.), cl. 6; EX1006, 2; EX1003, ¶¶252-257; EX1054, ¶¶137-139; Pet., 66-67; Reply, 26.

'852 and '200 Patents



US011041852B2

(12) **United States Patent**
Porreca et al.

(10) **Patent No.: US 11,041,852 B2**
(45) **Date of Patent: Jun. 22, 2021**

(54) **METHODS FOR MAINTAINING THE INTEGRITY AND IDENTIFICATION OF A NUCLEIC ACID TEMPLATE IN A MULTIPLEX SEQUENCING REACTION**

(71) Applicant: **Molecular Loop Biosolutions**, Cambridge, MA (US)

(72) Inventors: **Gregory Porreca**, Cambridge, MA

(60) Continuation of application No. 14/854,629, filed on Sep. 15, 2015, which is a division of application No. 13/081,660, filed on Apr. 7, 2011, now Pat. No. 9,163,281.

(60) Provisional application No. 61/426,817, filed on Dec. 23, 2010.

(51) **Int. Cl.**
C12Q 1/6869 (2018.01)
G01N 33/543 (2006.01)

(52) **U.S. Cl.**
CPC ***G01N 33/543*** (2013.01); ***C12Q 1/6869*** (2013.01); *C12Q 2525/155* (2013.01); *C12Q 2525/179* (2013.01); *C12Q 2525/191* (2013.01); *G01N 2458/10* (2013.01)

(58) **Field of Classification Search**
CPC C12Q 1/6869; C12Q 2525/191
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,683,195	A	7/1987	Mullis et al.
4,683,202	A	7/1987	Mullis
4,988,617	A	1/1991	Landegren et al.
5,060,980	A	10/1991	Johnson et al.
5,210,015	A	5/1993	Gelfand et al.
5,234,809	A	8/1993	Boom et al.
5,242,794	A	9/1993	Whiteley et al.
5,348,853	A	9/1994	Wang et al.
5,434,049	A	7/1995	Okano et al.
5,459,307	A	10/1995	Klotz, Jr.
5,486,686	A	1/1996	Zdybel, Jr. et al.
5,491,224	A	2/1996	Bittner et al.
5,494,810	A	2/1996	Barany et al.
5,567,583	A	10/1996	Wang et al.
5,583,024	A	12/1996	McElroy et al.
5,604,097	A	2/1997	Brenner
5,636,400	A	6/1997	Young
5,674,713	A	10/1997	McElroy et al.
5,695,934	A	12/1997	Brenner
5,700,673	A	12/1997	McElroy et al.
5,701,256	A	12/1997	Marr et al.
5,830,064	A	11/1998	Bradish et al.
5,846,719	A	12/1998	Brenner et al.
5,863,722	A	1/1999	Brenner
5,866,337	A	2/1999	Schon

FOREIGN PATENT DOCUMENTS

EP	1 321 477	A1	6/2003
EP	1 564 306	A2	8/2005

(Continued)

OTHER PUBLICATIONS

Abravaya, 1995, Detection of point mutations with a modified ligase chain reaction (Gap-LCR), Nucleic Acids Research, 23(4): 675-682.
Adey, 2010, Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition, Genome Biol 11:R119, 17 pages.
Ageno, 1969, The alkaline denaturation of DNA, Biophys J 9(11):1281-1311.
Agrawal, 1990, Site-specific functionalization of oligodeoxynucleotides for non-radioactive labelling, Tetrahedron Let 31:1543-1546.

(Continued)

Primary Examiner — Teresa E Strzelecka
(74) *Attorney, Agent, or Firm* — Brown Rudnick LLP; Thomas C. Meyers

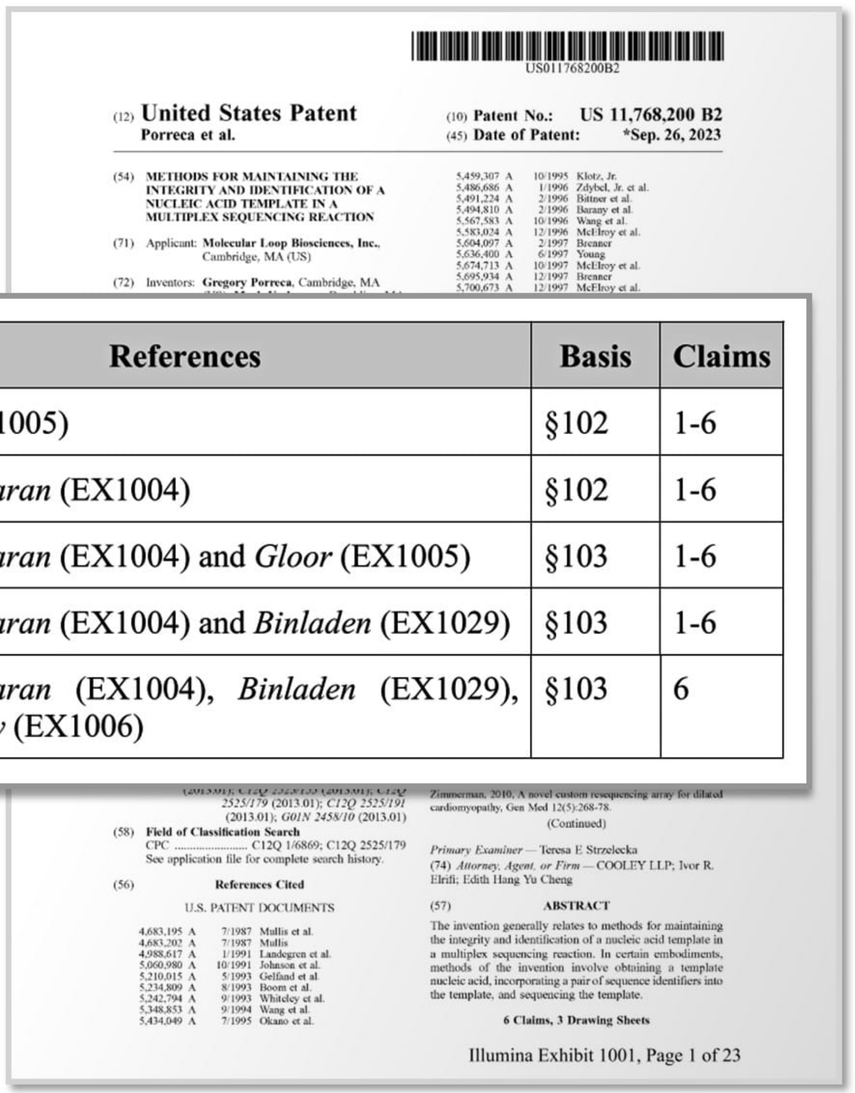
(57) **ABSTRACT**

The invention generally relates to methods for maintaining the integrity and identification of a nucleic acid template in a multiplex sequencing reaction. In certain embodiments, methods of the invention involve obtaining a template nucleic acid, incorporating a pair of sequence identifiers into the template, and sequencing the template.

8 Claims, 3 Drawing Sheets

Illumina Exhibit 1001, Page 1 of 23

Ground	References	Basis	Claims
1	*Parameswaran* (EX1004) and *Gloor* (EX1005)	§103	1-8
2	*Parameswaran* (EX1004) and *Bentley* (EX1006)	§103	1-8



US011768200B2

(12) **United States Patent**
Porreca et al.

(10) **Patent No.: US 11,768,200 B2**
(45) **Date of Patent: *Sep. 26, 2023**

(54) **METHODS FOR MAINTAINING THE INTEGRITY AND IDENTIFICATION OF A NUCLEIC ACID TEMPLATE IN A MULTIPLEX SEQUENCING REACTION**

(71) Applicant: **Molecular Loop Biosciences, Inc.**, Cambridge, MA (US)

(72) Inventors: **Gregory Porreca**, Cambridge, MA

2525/179 (2013.01); *C12Q 2525/191* (2013.01); *G01N 2458/10* (2013.01)

(58) **Field of Classification Search**
CPC C12Q 1/6869; C12Q 2525/179
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,683,195	A	7/1987	Mullis et al.
4,683,202	A	7/1987	Mullis
4,988,617	A	1/1991	Landegren et al.
5,060,980	A	10/1991	Johnson et al.
5,210,015	A	5/1993	Gelfand et al.
5,234,809	A	8/1993	Boom et al.
5,242,794	A	9/1993	Whiteley et al.
5,348,853	A	9/1994	Wang et al.
5,434,049	A	7/1995	Okano et al.
5,459,307	A	10/1995	Klotz, Jr.
5,486,686	A	1/1996	Zdybel, Jr. et al.
5,491,224	A	2/1996	Bittner et al.
5,494,810	A	2/1996	Barany et al.
5,567,583	A	10/1996	Wang et al.
5,583,024	A	12/1996	McElroy et al.
5,604,097	A	2/1997	Brenner
5,636,400	A	6/1997	Young
5,674,713	A	10/1997	McElroy et al.
5,695,934	A	12/1997	Brenner
5,700,673	A	12/1997	McElroy et al.

Zimmerman, 2010, A novel custom resequencing array for dilated cardiomyopathy, Gen Med 12(5):268-78.

(Continued)

Primary Examiner — Teresa E Strzelecka
(74) *Attorney, Agent, or Firm* — COOLEY LLP; Ivor R. Elrifi; Edith Hang Yu Cheng

(57) **ABSTRACT**

The invention generally relates to methods for maintaining the integrity and identification of a nucleic acid template in a multiplex sequencing reaction. In certain embodiments, methods of the invention involve obtaining a template nucleic acid, incorporating a pair of sequence identifiers into the template, and sequencing the template.

6 Claims, 3 Drawing Sheets

Illumina Exhibit 1001, Page 1 of 23

Ground	References	Basis	Claims
1	*Gloor* (EX1005)	§102	1-6
2	*Parameswaran* (EX1004)	§102	1-6
3	*Parameswaran* (EX1004) and *Gloor* (EX1005)	§103	1-6
4	*Parameswaran* (EX1004) and *Binladen* (EX1029)	§103	1-6
5	*Parameswaran* (EX1004), *Binladen* (EX1029), and *Bentley* (EX1006)	§103	6

Detailed Analysis

A POSA Would Have Understood that *Parameswaran*'s Sequences with Mismatched Barcode Motifs Included Cross-Over Error

Table 3. Summary of sequence distribution for both pyrosequencing runs as a function

Run number	Search motif	Perfect 5′ and 3′ Motifs	Mismatched 5′ and 3′ motifs	Perf impe 5′ m
I	Barcode (10 nt)	229 277	2349	2323
I	Barcode (10 nt) + Cloning-Linker (1 nt)	226 785	295	4339
I	Barcode (10 nt) + Cloning-Linker (2 nt)	226 108	4	5368
I	Barcode (10 nt) + Cloning-Linker (3 nt)	224 645	3	5800
I	Barcode (10 nt) + Cloning-Linker (4 nt)	223 744	3	6258
I	Barcode (10 nt) + Cloning-Linker (COMPLETE)	205 020	1	12 09
II	Barcode (10 nt)	258 929	5061	3138
II	Barcode (10 nt) + Cloning-Linker (1 nt)	257 998	17	6128
II	Barcode (10 nt) + Cloning-Linker (2 nt)	257 150	12	6587
II	Barcode (10 nt) + Cloning-Linker (3 nt)	255 695	11	7175
II	Barcode (10 nt) + Cloning-Linker (4 nt)	254 491	11	7818
II	Barcode (10 nt) + Cloning-Linker (COMPLETE)	226 041	11	14 6(

Parameswaran **(EX1004, 7)**

GS20's quality control filters. Since six of the forward barcodes are exact reverse complements of six reverse barcodes, we added a number of adjoining nucleotides in the Cloning-Linker (0–17) to the barcode query to reduce ambiguity while segregating the sequences into independent datasets. These barcode + linker sequences are referred to herein as 'search motifs' or 'barcode motifs'.



Petitioner's Expert, Dr. Spellman (EX1054, ¶69)

"To address this ambiguity, Parameswaran combined the 10-nt barcodes with three bases from the adjoining linker. This made each barcode motif unique to each sample, because neither the sequence nor its reverse complement was repeated across different samples. Importantly, while this configuration reduced the number of sequences with conflicting barcodes, it did ***not*** eliminate them."

illumina

There Were Multiple Reasons to Combine *Parameswaran* and *Gloor*

Patent Owner Response (POR, 22)

"Gloor ***warns against*** applying pyrosequencing approaches to Illumina sequencers."

Gloor (EX1005, 2)

rRNA regions. Similarly to others [8], we found that methods used to analyze pyrosequencing microbiome data were often unsuitable for reads generated by Illumina sequencing and we present a workflow that can be used for rapid and robust generation of the relative abundance of organisms in each sample.

Patent Owner Sur-Reply (Sur-Reply, 18)

"Parameswaran referred only to its individual barcode designs—and not its unique dual indexing approach as a whole—as potentially applicable to other platforms."

Parameswaran (EX1004, 2)

sequences) in run II. Discussed here are the modifications made for cloning and amplification protocols to handle highly complex nucleic acid pools, the considerations used in designing the barcodes, and the efficacy of the barcoding technology. We also speculate on the use of these pyrosequencer-tailored barcodes with other sequencing platforms.

illumina

Gloor Was Publicly Available Before December 2010



"Consistent with my knowledge and experience, the arXiv.org website provides information regarding its publication policy to its users. . . . This policy, updated in March 2010, stated that '**submission to arXiv are normally announced on a daily schedule**, with no activity scheduled on weekends.'" EX1046, ¶12

EX1048, EX1049

http://arxiv.org/help/submit Go

947 captu
27 Jun 2001

JUL SEP OCT
17
2009 2010 2011

* * *

Availability of submissions and announcement order

* * *

Submissions to arXiv are normally announced on a daily schedule, with no activity scheduled on weekends (check current local time at the main site).

Example: A submission made on Monday before 16:00 (US Eastern time, EDT/EST) will be publicly available on the web after 20:00 (US Eastern time, EDT/EST) on that day, and will appear in the Tuesday AM regular subject mailing.

'852 Patent **(EX1001)**

Patent No.: US 11,041,852 B2

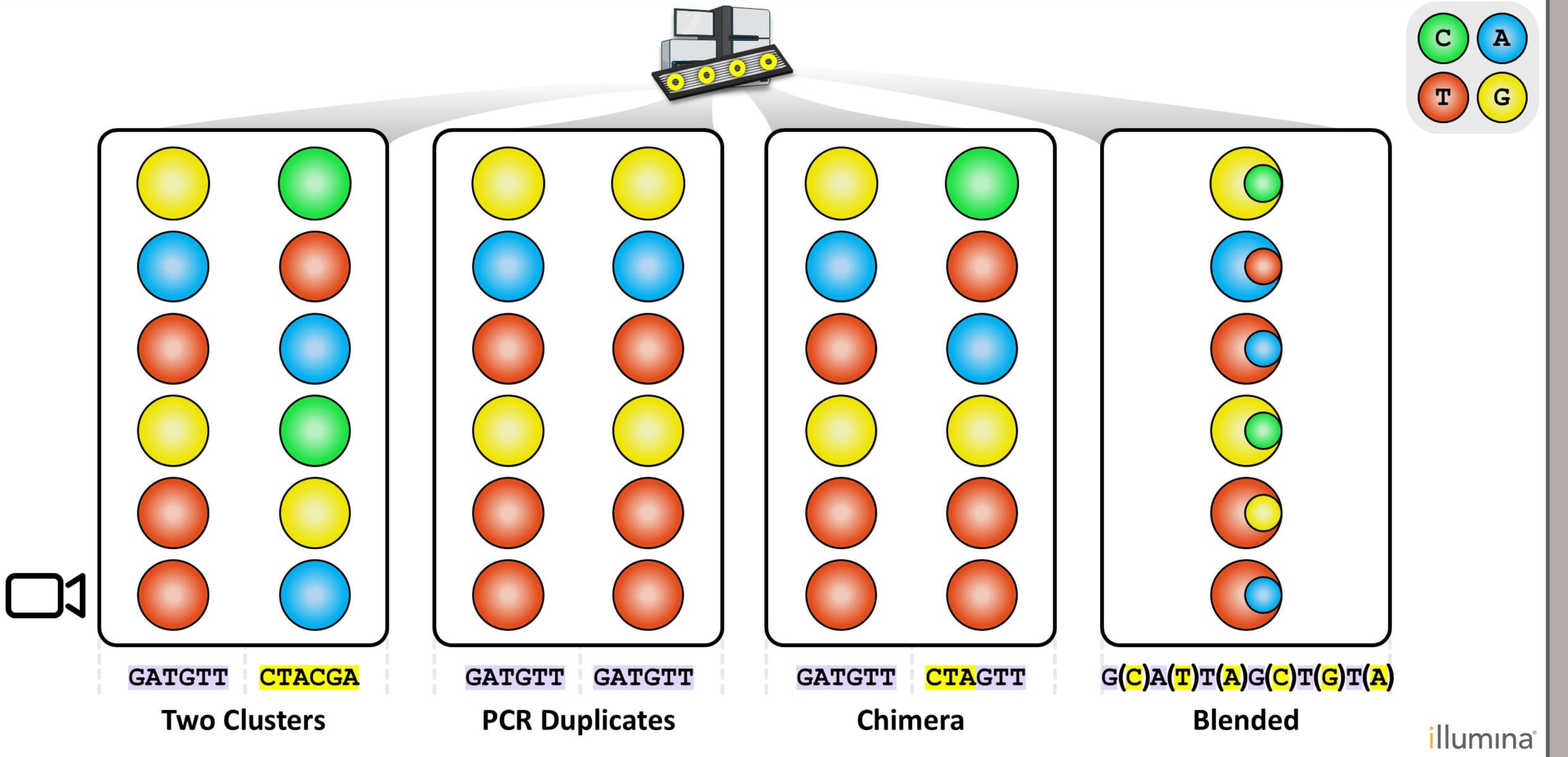
* * *

Provisional application No. 61/426,817, filed on Dec. 23, 2010.

July | August | September | October | November | December

EX1048; EX1049; EX1046, ¶¶11-13; Pet., 12; Reply, 24-28.

Illumina Sequencing



Illumina's Replies Reinforce Previously Asserted Arguments and Are Responsive

Patent Owner Allegation	Petition	Reply
"The Board should decline to even consider the new arguments and evidence regarding the alleged awareness of chimeras in Illumina sequencing." '852 Sur-Reply, 8	"As of December 2010, **a POSA would have known that *chimeric sequences*, including those resulting from cross-over errors**, were problematic for next-generation sequencing platforms that used PCR amplification, ***including Illumina sequencing*.**" '852 Pet., 38	"Patent Owner and Dr. Adalsteinsson failed to consider numerous prior-art references disclosing **the problem of *chimeric sequences* in *Illumina sequencing*.**" '852 Reply, 7
"Petitioner now asserts that the discarding step might be obvious depending upon the "purpose" of a particular application (Reply, 13-16), apparently abandoning its prior assertions of disclosure in the prior art." '852 Sur-Reply, 21	"*Parameswaran* states that '**for our *purposes***, all sequences with a match to one of the two expected sequence motifs were considered as true members of the corresponding dataset,' . . . . A POSA would have understood that ***these false discoveries could be excluded to increase accuracy*.**" '852 Pet., 35	"A POSA would have understood from *Parameswaran* that chimeric sequences could be kept or discarded, depending on the sequencing '***purpose***.' . . . For example, if sample tracking and sequence fidelity were important, ***discarding chimeras would have been obvious to avoid false positives*.**" '852 Reply, 14

Patent Owner's Arguments on *Gloor* Should Be Rejected

1 [pre] A method for validating the sequence of a nucleic acid analyte of interest in a multiplex sequencing reaction, the method comprising

[a] detecting the presence of two or more identifier sequences that are uniquely associated with the nucleic acid analyte of interest, wherein at least a first identifier sequence is incorporated into a 5′ portion of the nucleic acid analyte of interest and wherein at least a second identifier sequence is incorporated into a 3′ portion of the nucleic acid analyte of interest, and wherein the first and second identifier sequences have four or more nucleotides and are different;

[b] sequencing the first identifier sequence, the nucleic acid analyte of interest and the second identifier sequence; and

[c] validating the sequence of the nucleic acid analyte of interest by analyzing both identifiers and excluding the sequences of those nucleic acid analytes of interest **containing only one identifier or an incorrect pair of identifiers** from sequence analysis of the multiplex sequencing reaction.

Patent Owner Response (POR, 14)

"Gloor uses a **combinatorial** sequence tagging method, ***not*** a method where identifier sequences are uniquely associated with a nucleic acid template."

Claim Construction Dependent **(Reply, 2-4)**



Petitioner's Expert, Dr. Spellman (EX1054, ¶19)

Dr. Adalsteinsson does not "dispute that sequences containing **only one identifier were excluded in *Gloor***. . . . And he ignores that sequences containing incorrect pairs of identifiers, such as sequences with truncated tags, tags with additional bases, or tags with misincorporated bases, were also excluded in *Gloor*."

illumina