United States Court of Appeals for the Federal Circuit

ARCHERDX, LLC, THE GENERAL HOSPITAL CORPORATION, dba Massachusetts General Hospital,

Plaintiffs-Cross-Appellants,

– v. –

QIAGEN SCIENCES, LLC, QIAGEN LLC, fka QIAGEN, Inc., QIAGEN BEVERLY, LLC, fka QIAGEN Beverly, Inc., QIAGEN GAITHERSBURG, LLC, fka QIAGEN Gaithersburg, Inc., QIAGEN GMBH, QIAGEN N.V., JONATHAN ARNOLD,

Defendants-Appellants.

On Appeal from the United States District Court for the District of Delaware in No. 1:18-cv-01019-MN, Honorable Maryellen Noreika, Judge

BRIEF FOR DEFENDANTS-APPELLANTS

BRIAN C. CANNON
QUINN EMANUEL URQUHART & SULLIVAN, LLP
555 Twin Dolphin Drive, 5th Floor
Redwood Shores, California 94065
(650) 801-5000
briancannon@quinnemanuel.com

DAVID BILSKER
ANDREW E. NARAVAGE
QUINN EMANUEL URQUHART & SULLIVAN, LLP
50 California Street, 22nd Floor
San Francisco, California 94111
(415) 875-6600
davidbilsker@quinnemanuel.com
andrewnaravage@quinnemanuel.com

Counsel for Defendants-Appellants

DECEMBER 15, 2023

CP COUNSEL PRESS (800) 4-APPEAL • (713211)

EXHIBIT # Metzker 2207 12/03/2025 exhibitsticker.com

PATENTS CLAIMS AT ISSUE

This appeal concerns claims 16, 17, and 19 of U.S. Patent No. 10,017,810 and claims 1, 5, and 19 of U.S. Patent No. 10,450,597. Claim 16 of the ’810 patent and claim 1 of the ’597 patent are exemplary:

’810 patent, Claim 16:

> A method for preparing a nucleic acid for sequencing, the method comprising:
>
> (i) ligating a universal oligonucleotide tail adaptor that comprises a first ligatable duplex end and a second unpaired end to a nucleic acid comprising a known target nucleotide sequence to produce a ligation product, the universal oligonucleotide tail adaptor comprising an amplification strand and a blocking strand, wherein a 3′ duplex portion of the amplification strand and a 5′ duplex portion of the blocking strand are substantially complementary and form the first ligatable duplex end;
>
> (ii) amplifying the ligation product using a first target-specific primer that specifically anneals to the known target nucleotide sequence and a first adaptor primer having a nucleotide sequence identical to a first portion of the amplification strand; and
>
> (iii) amplifying an amplification product of (ii) using a second target-specific primer that specifically anneals to the amplification product of (ii) and a second adaptor primer having a nucleotide sequence identical to a second portion of the amplification strand, wherein ligating in step (i) comprises performing an overhang ligation reaction, and wherein the universal oligonucleotide tail adaptor further comprises a barcode portion.

’597 patent, Claim 1:

> A method of preparing nucleic acids for analysis, the method comprising:
>
> (a) contacting a first nucleic acid template comprising a sequence of a first strand of a double-stranded target nucleic acid with a

complementary target-specific primer that comprises a target-specific hybridization sequence, under conditions to promote template-specific hybridization and extension of the target-specific primer;

(b) contacting a second nucleic acid template comprising a sequence of a second strand that is complementary to the sequence of the first strand of the double-stranded target nucleic acid with a plurality of different primers that share a common sequence that is 5′ to different hybridization sequences, under conditions to promote template-specific hybridization and extension of at least one of the plurality of different primers, wherein the different hybridization sequences have different 3′ ends, and wherein each primer of the plurality of different primers does not anneal to the same sequence of the double-stranded target nucleic acid as any other primer of the plurality of different primers,

wherein, following (a) and (b), an extension product is generated to contain both a sequence that is characteristic of the target-specific primer and a sequence that is characteristic of the at least one of the plurality of different primers; and

(c) subjecting the extension product to an amplification reaction comprising successive rounds of polymerase extension of i) a tail primer that comprises a 3′ sequence that specifically anneals to the complement of the common sequence and that comprises a 5′ tail sequence, and ii) a primer that specifically anneals to the complement of the target-specific hybridization sequence.

UNITED STATES COURT OF APPEALS
FOR THE FEDERAL CIRCUIT

CERTIFICATE OF INTEREST

Case Number 23-2350, 23-2379

Short Case Caption ArcherDX, LLC v. QIAGEN Sciences, LLC

Filing Party/Entity QIAGEN Sciences, LLC; QIAGEN LLC; QIAGEN Beverly, LLC; QIAGEN Gaithersburg, LLC; QIAGEN GmbH; QIAGEN N.V.; Jonathan Arnold

Instructions:

1. Complete each section of the form and select none or N/A if appropriate.
2. Please enter only one item per box; attach additional pages as needed, and check the box to indicate such pages are attached.
3. In answering Sections 2 and 3, be specific as to which represented entities the answers apply; lack of specificity may result in non-compliance.
4. Please do not duplicate entries within Section 5.
5. Counsel must file an amended Certificate of Interest within seven days after any information on this form changes. Fed. Cir. R. 47.4(c).

I certify the following information and any attached sheets are accurate and complete to the best of my knowledge.

Date: 12/15/2023

Signature: /s/ Brian C. Cannon

Name: Brian C. Cannon

1. Represented Entities. Fed. Cir. R. 47.4(a)(1).	**2. Real Party in Interest.** Fed. Cir. R. 47.4(a)(2).	**3. Parent Corporations and Stockholders.** Fed. Cir. R. 47.4(a)(3).
Provide the full names of all entities represented by undersigned counsel in this case.	Provide the full names of all real parties in interest for the entities. Do not list the real parties if they are the same as the entities. ☑ None/Not Applicable	Provide the full names of all parent corporations for the entities and all publicly held companies that own 10% or more stock in the entities. ☐ None/Not Applicable
QIAGEN Sciences, LLC	N/A	QIAGEN N. V.
QIAGEN LLC	N/A	QIAGEN N. V.
QIAGEN Beverly, LLC	N/A	QIAGEN N. V.
QIAGEN Gaithersburg, LLC	N/A	QIAGEN N. V.
QIAGEN GmbH	N/A	QIAGEN N. V.
QIAGEN N. V.	N/A	BlackRock, Inc.; Sun Life Financial, Inc. (MFS)
Jonathan Arnold	N/A	N/A

☐ Additional pages attached

4. Legal Representatives. List all law firms, partners, and associates that (a) appeared for the entities in the originating court or agency or (b) are expected to appear in this court for the entities. Do not include those who have already entered an appearance in this court. Fed. Cir. R. 47.4(a)(4).

☐ None/Not Applicable ☐ Additional pages attached

Anne S. Toker Quinn Emanuel Urquhart & Sullivan LLP	ames E. Baker Quinn Emanuel Urquhart & Sullivan LLP	Anastasia M. Fernands Quinn Emanuel Urquhart & Sullivan LLP
Jeffrey C. Wu Quinn Emanuel Urquhart & Sullivan LLP	David E. Moore Potter Anderson & Corroon LLP	Bindu A. Palapura Potter Anderson & Corroon LLP
Stephanie E. O'Byrne formerly Potter Anderson & Corroon LLP		

5. Related Cases. Other than the originating case(s) for this case, are there related or prior cases that meet the criteria under Fed. Cir. R. 47.5(a)?

☑ Yes (file separate notice; see below) ☐ No ☐ N/A (amicus/movant)

If yes, concurrently file a separate Notice of Related Case Information that complies with Fed. Cir. R. 47.5(b). **Please do not duplicate information.** This separate Notice must only be filed with the first Certificate of Interest or, subsequently, if information changes during the pendency of the appeal. Fed. Cir. R. 47.5(b).

6. Organizational Victims and Bankruptcy Cases. Provide any information required under Fed. R. App. P. 26.1(b) (organizational victims in criminal cases) and 26.1(c) (bankruptcy case debtors and trustees). Fed. Cir. R. 47.4(a)(6).

☑ None/Not Applicable ☐ Additional pages attached

TABLE OF CONTENTS

Page

TABLE OF AUTHORITIES iii

STATEMENT OF RELATED CASES vi

INTRODUCTION 1

JURISDICTIONAL STATEMENT 5

STATEMENT OF THE ISSUES 5

STATEMENT OF THE CASE 5

A. Technology At Issue 5

B. The Accused QIAGEN Products 10

C. The Asserted '810 Patent 13

D. The Asserted '597 Patent 18

E. The District Court Case 21

SUMMARY OF THE ARGUMENT 24

STANDARDS OF REVIEW 27

ARGUMENT 28

I. THE ACCUSED QIAGEN PRODUCTS DO NOT INFRINGE THE '810 PATENT 28

A. The Accused Products Do Not Use A "Second Target-Specific Primer" 28

B. QIAGEN's Sample Index Primer Cannot Be A "Second Target-Specific Primer" Under The DOE 31

C. If Non-Nested SIP Can Be Expanded Under The DOE To Be A "Second Target-Specific Primer" Then The Claim Fails For Lack Of Written Description 39

D. The District Court Erred With Respect To The "Sequencing Primer" Requirement Of The Claims 41

II. THE ACCUSED QIAGEN PRODUCTS DO NOT INFRINGE THE '597 PATENT 44

A. QIAGEN's Accused "Forward Primer" Cannot Be A "Target-Specific Primer" 44

B. Insufficient Evidence Established That Sequences In The First And Second Strands Of The "Double-Stranded Target Nucleic Acid" Interacted With The Accused Primers To Satisfy Steps (a) And (b) Of Claim 1 49

III. THE DISTRICT COURT ERRED IN PERMITTING DAMAGES FOR FOREIGN INFRINGEMENT 57

A. There Is No Legal Basis For Plaintiffs To Recover Damages For The Practice of Method Claims In Foreign Countries 58

B. Even If *CMU* Remains Good Law, It Does Not Support The Foreign Damages Jury Instruction 65

C. Even If The Jury Instruction Was Correct, Insufficient Evidence Supported The Foreign Damages Verdict 66

CONCLUSION 67

Tecan Ex. 2207, Page 8 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

TABLE OF AUTHORITIES

Cases

Abitron Austria GmbH v. Hetronic Int'l, Inc.,
600 U.S. 412 (2023)..... 64

Amgen Inc. v. Sandoz Inc.,
923 F.3d 1023 (Fed. Cir.), *reh'g granted, opinion modified*,
776 F. App'x 707 (Fed. Cir. 2019)..... 37

Biogen MA Inc. v. EMD Serono, Inc.,
976 F.3d 1326 (Fed. Cir. 2020)..... 27

Bicon, Inc. v. Straumann Co.,
441 F.3d 945 (Fed. Cir. 2006)..... 54-55

Ariad Pharms. v. Eli Lilly,
598 F.3d 1336 (Fed. Cir. 2010)..... 39

Cardiac Pacemakers, Inc. v. St. Jude Medical, Inc.,
576 F.3d 1348 (Fed. Cir. 2009)..... 59, 62-64

Carnegie Mellon Univ. v. Hoffmann-La Roche Inc.,
541 F.3d 1115 (Fed. Cir. 2008)..... 32

Carnegie Mellon v. Marvell,
807 F.3d 1283 (Fed. Cir. 2015).....26-27, 59-60

Chicago Bd. Options Exch., Inc. v. Int'l Sec. Exch., LLC,
677 F.3d 1361 (Fed. Cir. 2012)..... 30

CommScope Techs. LLC v. Dali Wireless Inc.,
10 F.4th 1289 (Fed. Cir. 2021)..... 52

CVI/Beta Ventures, Inc. v. Tura LP,
112 F.3d 1146, 42 USPQ2d 1577 (Fed.Cir.1997)..... 56

Ecolab, Inc. v. Envirochem, Inc.,
264 F.3d 1358 (Fed. Cir. 2001)..... 30

Energizer Holdings, Inc. v. Int'l Trade Comm'n,
435 F.3d 1366 (Fed. Cir. 2006)..... 53

Tecan Ex. 2207, Page 9 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

Enzo Biochem, Inc. v. Applera Corp.,
599 F.3d 1325 (Fed. Cir. 2010) 6, 57

In re Fought,
941 F.3d 1175 (Fed. Cir. 2019) 55

Genentech, Inc. v. Wellcome Found. Ltd.,
29 F.3d 1555 (Fed. Cir. 1994) 34-35

Haemonetics Corp. v. Baxter Healthcare Corp.,
607 F.3d 776 (Fed. Cir. 2010) 55

Halo Elec. v. Pulse Elec.,
769 F.3d 1371 (Fed. Cir. 2014) 58, 65-66

Joy Tech. v. Flakt, Inc.,
6 F.3d 770 (Fed. Cir. 1993) 59, 62-63

Microsoft Corp. v. AT&T Corp.,
550 U.S. 437 (2007)..... 58, 64

Morrison v. Nat'l Australia Bank Ltd.,
561 U.S. 247 (2010)..... 64

Newell Companies, Inc. v. Kenney Mfg. Co.,
864 F.2d 757 (Fed. Cir. 1988) 62

NTP, Inc. v. Rsch. In Motion, Ltd.,
418 F.3d 1282 (Fed. Cir. 2005) 57-59

Nuvo v. Dr. Reddy's Lab.,
923 F.3d 1368 (Fed. Cir. 2019) 41

O2 Micro Int'l v. Beyond Innovation Tech.,
521 F.3d 1351 (Fed. Cir. 2008) 44

Phonometrics, Inc. v. Northern Telecom Inc.,
133 F.3d 1459, 45 USPQ2d 1421 (Fed.Cir.1998)..... 55

Power Integrations, Inc. v. Fairchild Semiconductor Int'l, Inc.,
843 F.3d 1315 (Fed. Cir. 2016) 32-33

Tecan Ex. 2207, Page 10 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

Rexnord Corp. v. Laitram Corp.,
274 F.3d 1336 (Fed. Cir. 2001) 55

Roche Molecular Sys., Inc. v. CEPHEID,
905 F.3d 1363 (Fed. Cir. 2018) 8

TomTom, Inc. v. Adolph,
790 F.3d 1315 (Fed. Cir. 2015) 27

Virnetx, Inc. v. Cisco Sys., Inc.,
767 F.3d 1308 (Fed. Cir. 2014) 38

VLSI Tech. v. Intel Corp.,
No. 2022-1906, slip op. (Fed. Cir. Dec. 4, 2023) 31-32

Warner-Jenkinson Co. v. Hilton Davis Chem. Co.,
520 U.S. 17 (1997) 32

WesternGeco LLC v. ION Geophysical,
138 S.Ct. 2129 (2018) 63

Statutes

28 U.S.C. § 1295(a)(1) 5

28 U.S.C. § 1331 5

28 U.S.C. § 1338 5

35 U.S.C. § 112 ..... 5, 25

35 U.S.C. § 271(a) 5, 26, 57-59, 63, 65-66

35 U.S.C. § 271(f) 63-64

Other Authorities

Fed. R. Civ. P. 68 21

STATEMENT OF RELATED CASES

Defendants-Appellants identify the following appeals as related pursuant to Federal Circuit Rule 47.5: *ArcherDX, LLC v. QIAGEN Sciences, LLC*, Case No. 23-2379 (Fed. Cir.).

Tecan Ex. 2207, Page 12 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

INTRODUCTION

This is an appeal from a judgment after jury trial in the District Court for the District of Delaware (Norieka, J.) awarding Plaintiffs damages for infringement of method claims in two patents plus supplemental damages and ongoing royalty. That judgment should be reversed because QIAGEN does not infringe either patent.

Doctors and researchers often seek to know whether a patient's DNA includes a particular DNA sequence such as a mutation associated with a disease. However, sequencing the entire genome to find a single target has always been expensive. Thus, from the early days of DNA sequencing, researchers developed target enrichment techniques to amplify the target sequences one was seeking to analyze, while at the same time minimizing or eliminating sequences that were not of interest. Plaintiffs' patents are both directed to such enrichment methods that create what are called "DNA libraries" that are placed on third-party instruments for sequencing.

Critical to this case is the structure of DNA and the design of DNA "primers"—short single-stranded stretches of DNA that anneal to other DNA sequences. As the claims set forth, there are many types of primers at issue such as target-specific primers, adaptor primers, sequencing primers, and tail primers. Each of these primers has a particular sequence and a very specific

Tecan Ex. 2207, Page 13 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

role in the methods and technology at issue. The sequence of a primer determines what other DNA it will anneal to. Primer sequences and function are critical to the inventions at issue. One type of primer and its role and sequence cannot be swapped in for another—and Plaintiffs' efforts to do that here lie at the heart of why this case should be reversed.

Amplifying a particular fragment of DNA frequently relies upon the polymerase chain reaction ("PCR"). In PCR, primers work in pairs, one for each complementary strand of double-stranded DNA. Primers are selected to flank, among other things, the region a scientist is interested in amplifying and studying. Primers "prime" the PCR reaction by being the initiating point for making copies of that region.

Using PCR, QIAGEN and others have applied enrichment techniques to prepare samples for sequencing on third-party instruments such as those sold by Illumina. PCR and enrichment techniques are in the prior art. Indeed, QIAGEN sold earlier versions of its enrichment kits years before the patents-in-suit were sought. Like many before it, the accused QIAGEN kits add DNA molecules (called "adaptors") to one end of each strand of fragmented DNA in a sample. QIAGEN then uses a primer called "Gene Specific Primer" (GSP)—carefully designed for maximum specificity—to anneal to a gene sequence under investigation and, when paired with what the inventors

Tecan Ex. 2207, Page 14 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

referred to as a "universal" primer that attaches to the adaptor sequences at the ends of the fragments, to amplify that gene sequence many times over when compared to sequences not of interest contained in the sample.

The patents at issue, however, approach enrichment using different primer designs and different steps. Plaintiffs' two patents have overlapping inventors and one (the '597 patent) is said to be an improvement over the earlier '810 patent.

For the '810 patent, the method requires two separate amplification steps with two different primer pairs—*first* with a "first target-specific primer" and its partner (a first adaptor primer), and then *second* with a "second target-specific primer" and its partner (a second adaptor primer). QIAGEN does not infringe the '810 patent because the QIAGEN primer accused of being the "second target-specific primer" is a "universal" primer that fails to satisfy the requirements for a second target-specific primer for multiple reasons. Indeed, the district court granted summary judgment of no literal infringement recognizing the absence of one of those requirements. The jury's verdict finding that requirement satisfied through the doctrine of equivalents, as well as the existence of one other requirement, was not supported by substantial evidence and should have been reversed.

Tecan Ex. 2207, Page 15 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

For the ’597 patent, Plaintiffs turn their infringement theory on its head. For the “target-specific primer” for the ’597 patent, Plaintiffs accuse the QIAGEN primer that they admitted was a “universal” primer when pursuing the ’810 patent. And Plaintiffs then accuse QIAGEN’s Gene Specific Primers of being the random “plurality of different primers” the ’597 patent uses in place of the ligated adaptors of the ’810 patent. Nothing in the ’597 patent allows this complete role reversal. Plaintiffs’ infringement theory also depends upon abandoning the law of antecedent basis and changing the molecule it accuses of being the “double-stranded target nucleic acid” with particular first and second strands throughout the claim in an attempt to find sequences which the accused primers, having very specific required sequences, could anneal to.

Not only did the district court err on infringement, but it permitted the jury to award damages for the foreign use of the accused kits. If this Court reaches damages in this appeal, damages for the foreign uses should be reversed for multiple reasons including because: foreign use of method claims is not infringement, the law does not recognize the sale of a product as the sale of a patented method, and Congress did not sanction extraterritorial application of the patent laws on the facts here.

Tecan Ex. 2207, Page 16 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

JURISDICTIONAL STATEMENT

The district court had jurisdiction pursuant to 28 U.S.C. §§ 1331 and 1338. This court has jurisdiction over QIAGEN's appeal of the district court's August 8, 2023 final judgment pursuant to 28 U.S.C. § 1295(a)(1). Appx73-77. QIAGEN filed a timely notice of appeal on August 28, 2023. Appx222.

STATEMENT OF THE ISSUES

1. Whether the district court erred in denying QIAGEN's JMOL that the accused QIAGEN products do not infringe the '810 patent. And, in the alternative, if doctrine of equivalents infringement is permitted, whether the claims lack support under 35 U.S.C. § 112.

2. Whether the district court erred in denying QIAGEN's JMOL that the accused QIAGEN products do not infringe the '597 patent.

3. Whether the district court erred in permitting Plaintiffs to recover damages under 35 U.S.C. § 271(a) for foreign infringement of method claims where the accused products are sold and used outside the United States.

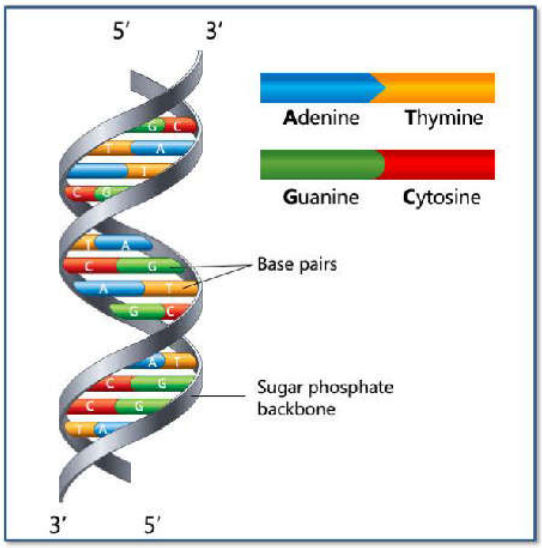
STATEMENT OF THE CASE

A. Technology At Issue

DNA Background: DNA is the double-stranded molecule that contains the genetic blueprint for the production of proteins. Appx504-05 at 154:7-158:8. DNA contains four nucleic acid bases abbreviated A, T, G and C. As one of the '810 patent inventors explained, the two strands of DNA—"top"

Tecan Ex. 2207, Page 17 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

and "bottom"—are "complementary" to each other. Appx505 at 158:11-19. A is complementary to T meaning it can bind or anneal to T. Similarly, G is complementary to C. The following illustration shows the relationship:



- **Genetic code** – A, C, G, T
- **Sequencing** – determining the order of the nucleotide bases in a DNA fragment

The importance of complementarity: The two strands of DNA are aligned. *Enzo Biochem, Inc. v. Applera Corp.*, 599 F.3d 1325, 1328 (Fed. Cir. 2010). One strand binds in a complementary fashion with the other. *Id.*

Sequencing: Researchers often seek to sequence DNA, i.e., determine the order of the base pairs, for many reasons, including looking for mutations which can cause disease. Appx505 at 160:12-16; Appx691 at 899:15-18. In order to sequence DNA on particular current-day sequencers, the sample must be prepared to contain specific adaptors that create what is called a "library" that is capable of being recognized and thus sequenced on instruments like an Illumina sequencer. Appx508 at 171:25-172:9.

Early recognition of enrichment techniques: Sequencing the whole genome was traditionally expensive and time consuming. First-generation sequencing methods such as Sanger sequencing took 13 years and $300 million to first sequence the human genome. Appx689 at 891:11-18. Because of the costs and time required, researchers developed techniques that could narrow down the universe of material to be sequenced to predominantly comprise regions of interest. By "enriching" a sample to make copies of the areas of interest—and not the entire genome—sequencing could be performed more efficiently.

Over time, sequencing methods improved, and next-generation sequencing ("NGS") greatly increased the efficiency of DNA sequencing. Appx689 at 891:11-18. But prior enrichment techniques still proved useful for increasing speed and reducing cost. Appx692 at 901:25-903:8.

The patents are directed to enrichment: The well-known cost saving benefit of sample enrichment for NGS is embraced in both the '810 and '597 patents-in-suit which identically state: "***Target enrichment*** prior to next-generation sequencing is more cost-effective than whole genome . . . and therefore more practical for broad implementation; both for research discovery and clinical applications." Appx90 col. 1:38-42; Appx137 col. 1:26-30 (emphasis added).

Tecan Ex. 2207, Page 19 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

Enriching the target, the DNA of interest: What researchers often look for is mutations in the DNA sequence. A mutation means a "change in the genetic information" and it may result in a protein that leads to disease. Appx692-93 at 903:20-904:4; Appx506 at 161:13-162:1. The stretch of DNA that may include a mutation in the code is often the target sequence to be enriched for pre-sequencing.

PCR using primer pairs to amplify DNA: Enrichment methods to create the DNA to be sequenced have frequently relied on PCR. Appx692 at 902:19-903:4. PCR allows one to make millions of copies of particular DNA sequences. This amplification is important because researchers are often working with a small amount of starting material. Appx575 at 436:7-20.

PCR starts with template DNA—a top and bottom strand. Appx508-09 at 172:17-174:6. PCR requires pieces of DNA called primers that initiate reactions that make copies of each strand of the double-stranded DNA molecule. *Id.*; *see also Roche Molecular Sys., Inc. v. CEPHEID*, 905 F.3d 1363, 1366 (Fed. Cir. 2018) (background to PCR). Heating the DNA results in the two strands of the DNA separating, which allows the primers to attach. The pair of primers act as the starting point to create a second strand that is the complement of the first strand. Appx509 at 173:2-174:11; *Roche*, 905 F.3d at 1366. Mediated by a polymerase protein, the first primer starts a

Tecan Ex. 2207, Page 20 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

reaction that adds one complementary nucleotide at a time in one direction (5′ to 3′, or "left" to "right") to the end of the first primer that has hybridized to that first strand ("extending") while the second primer starts a reaction that does the same thing in the opposite direction (3′ to 5′, or "right" to "left"). Appx509 at 173:2-174:11. Repeating cycles of PCR can result in millions of copies of DNA. By designing the primer pairs so that they anneal to and start the copying between certain sequence locations in the DNA, one can thereby enrich the sample (i.e., dictate the sequence regions that will be copied/amplified) to isolate those regions of sequence that one is interested in sequencing and analyzing *e.g.* the target sequence. Appx692-93 at 903:20-904:24.

Primers can serve many functions: Primers are generally named by their function, and where and how primer pairs attach to the stands of DNA is determined by their sequence, length and complementarity. Appx96 col. 13:58-61; Appx615 at 597:3-8. For example, in the accused process, one set of primers add sequences that allow the DNA that will be sequenced to attach to the relevant location on the sequencing instrument. Other primers initiate the sequencing reactions by adding one detectable complementary nucleotide at a time to the end of a primer that has annealed to the DNA strand being sequenced.

Tecan Ex. 2207, Page 21 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

B. The Accused QIAGEN Products

The accused QIAGEN product kits contain the reagents necessary to prepare a sample comprising DNA for next generation sequencing.[1] To use a QIAGEN product, the researcher first obtains genomic material (e.g. DNA) such as from a patient's tumor tissue. Appx579 at 453:10-18; Appx1045. Next, it is fragmented into smaller pieces. Appx589 at 495:15-24; Appx1046. Then, known artificial sequences called "adaptors" are added to the ends of each fragmented DNA like a tail. Appx579 at 453:19-24; Appx1046. This addition is by a "ligation" reaction. Appx579 at 453:19-24. The adaptor sequences are identical and common to every QIAGEN accused product regardless of what gene mutation the product is designed to find. Appx599 at 533:4-10. In general, these common adaptors contain the nucleotide sequences that allow the polynucleotides be attached to a sequencing flow cell in an Illumina instrument. Appx596 at 523:2-18.

All of the accused QIAGEN kits contain three types of primers:

1. forward ("FP") and universal ("UP") primers that are designed to anneal to the complement of the adaptor sequence and which

[1] The sample may contain RNA, which reflects what DNA has been expressed within the cell. In that case, the RNA is converted to cDNA but there is no fragmentation step. While necessary for users to understand, this difference is not directly relevant for the infringement issues on appeal. The QIAGEN kits for RNA were subject to a different damages calculation.

Tecan Ex. 2207, Page 22 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

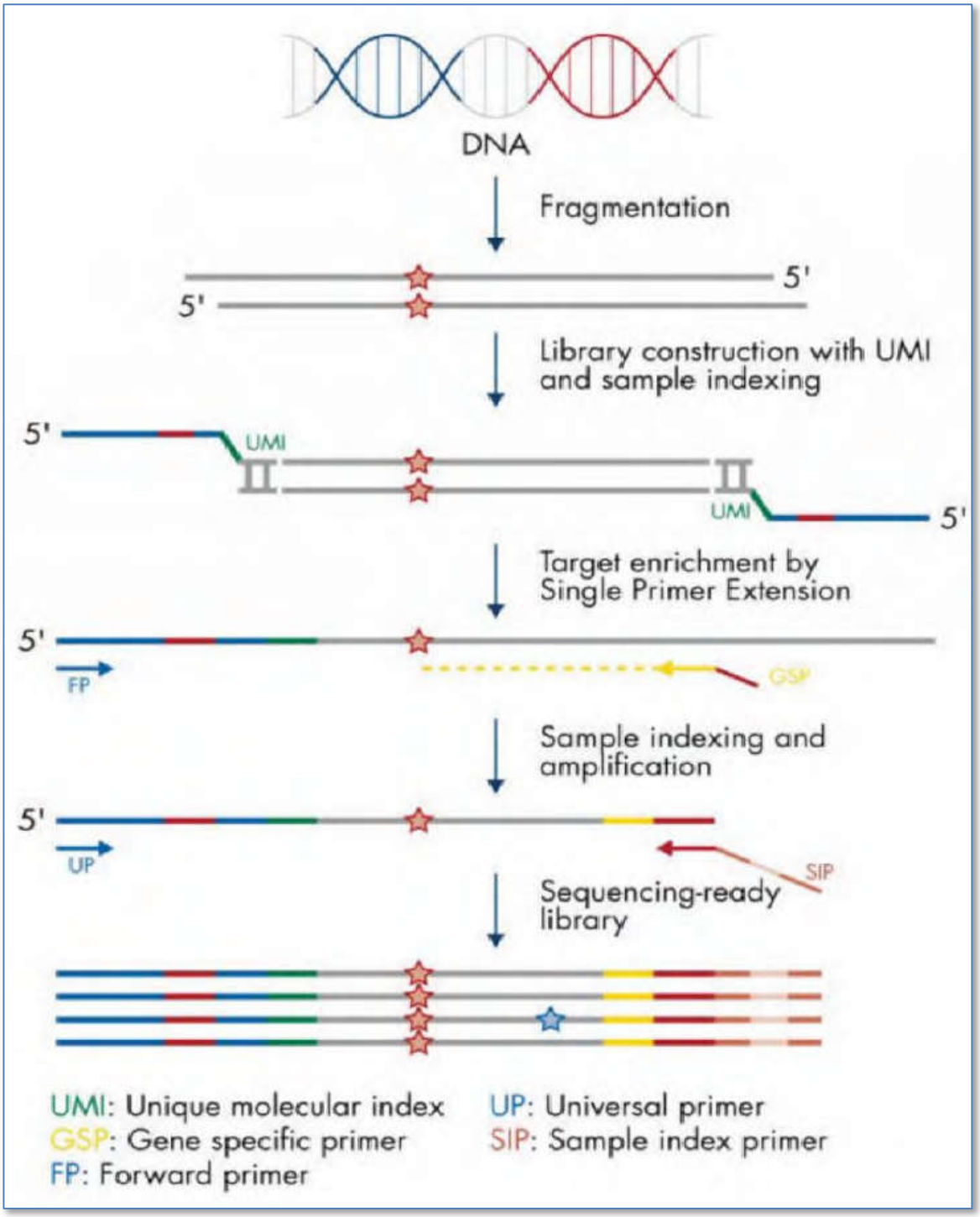
are identical except for the UP having one additional nucleotide at its 3′ end;

2. a gene specific primer ("GSP") that is designed to anneal to sequence in the gene one seeks to analyze, and
3. a sample index primer ("SIP") that anneals to a sequence the GSP adds on by way of a tail.

There is no dispute that the only unique primer that anneals to a sequence one is interested in analyzing, *i.e.*, the target, is the GSP, the gene specific primer. Appx595 at 518:9-15, 518:23-519:4. All of the other primers are the same across all of QIAGEN's kits. Appx599 at 532:16-25, 533:8-10. This includes the UP, FP and SIP. *Id*. QIAGEN carefully designed its process and constructed its GSPs so that its process can enrich for the sequence to be analyzed with a single pair of primers: the GSP and the FP.

JX 8.12 that both parties relied on depicts the QIAGEN workflow:

Tecan Ex. 2207, Page 23 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029



Appx1046. The third molecule from the top shows the red/blue/green artificial adaptor sequences ligated to all fragments in a sample. Appx611 at 581:24-582:6. In an initial round, the GSP is depicted with a yellow portion

that anneals to the gene and a red tail that does not anneal to anything. The FP anneals to the complement of the blue adaptor sequence. In the next round, the UP anneals to the same complement of the blue adaptor sequence, while the SIP (which Plaintiffs accuse of being a target-specific primer) anneals to the red tail added by the GSP. The fact that the UP, FP and SIP primers in the QIAGEN kits contain the same common, universal sequences, regardless of what gene mutation is being studied, is but one of many things showing the fallacy of Plaintiffs' infringement theories that classify these primers as being both "universal" ***and*** "target-specific" under their various theories, depending on the patent.

C. The Asserted '810 Patent

The '810 patent describes the claimed methods as relating to enriching target sequences prior to sequencing while limiting time consuming sequencing of near off target bases. Appx90 col. 1:53-55, 2:1-2. Inventor Dr. Iafrate testified, "our technology in the '810 patent is a method of preparing DNA so that it can be sequenced on one of these very sophisticated high throughput DNA sequencers." Appx508 at 171:16-20. One problem that has always been encountered when looking for mutations using the PCR enrichment method is that the sequence that one primer in the pair is designed against, like a portion of a gene, is known, while the sequence of its opposing

Tecan Ex. 2207, Page 25 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

paired primer is not. Appx509 at 176:2-16. To approach this problem, in the '810 patent, the DNA is broken into shorter fragments whose lengths can be read on current-day sequencing machines, like those Illumina makes. Then a known artificial sequence—an adaptor—is added to one end of each strand of the fragment. *Id.* The adaptor allows one to design an opposing primer against a known sequence common to all the fragments. *Id.* Through the use of target-specific primers designed against the known gene sequence, and "universal" primers designed against the common adaptor sequence, one can then create an enriched library for sequencing. Appx509-10 at 176:17-179:1, 180:11-19. The adaptor, however, also serves another important purpose. It contains sequences the makers of sequencing machines require for DNA to be anchored in a certain part of the sequencer so that various operations, such as the actual sequencing read outs, may be performed on the anchored DNA. Appx522 at 225:23-226:6.

Dr. Iafrate called "synthetic DNA" added to the end of each fragment of DNA "tail adaptors." Appx509-10 at 176:7-16, 177:7-8; Appx522 at 225:5-8. As the annotated claim sets forth below,[2] the method requires first ligating an adaptor to each fragmented piece of DNA, and then conducting

[2] The numbers in brackets [1], [2], [3] etc., are added to clarify the different molecules within the method.

Tecan Ex. 2207, Page 26 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

two rounds of amplification with ***two*** separate pairs of primers: a first adaptor primer and first target-specific primer, and a second adaptor primer and second target-specific primer, each of which have specific definitions in the patent:

> 16. A method for preparing a nucleic acid for sequencing, the method comprising:
>
> (i) ***ligating***
>
> [1] a universal oligonucleotide tail adaptor that comprises a first ligatable duplex end and a second unpaired end to
>
> [2] ***a nucleic acid comprising a known target nucleotide sequence*** to produce
>
> [3] a ligation product,
>
> […];
>
> (ii) ***amplifying*** the [3] ligation product [of step] (i) using
>
> [4] ***a first target-specific primer*** that specifically anneals to the known target nucleotide sequence and
>
> [5] ***a first adaptor primer*** having a nucleotide sequence identical to a first portion of the amplification strand; and
>
> (iii) ***amplifying*** an amplification product of [step] (ii) using
>
> [6] ***a second target-specific primer*** that specifically anneals to the amplification product of [step] (ii) and
>
> [7] ***a second adaptor primer*** having a nucleotide sequence identical to a second portion of the amplification strand,
>
> wherein ligating in step (i) comprises performing an overhang ligation reaction, and wherein the universal oligonucleotide tail adaptor further comprises a barcode portion.

The techniques described in the ’810 patent were adaptations of older techniques now applied to current-day sequencing machines. Because of this, the claims are very specific about the way the enrichment must be carried out

and the sequences that the primers must have and anneal to. The '810 patent requires enrichment to be accomplished using ***two rounds*** of PCR with different pairs of opposing primers. The first round uses a first target-specific primer and a first adaptor primer. The second round acts as a "proofreading" step of the first round by, for example, requiring the second target-specific primer to also anneal to a portion of the known target nucleotide sequence, rather than artificial sequence that may be added by the first target-specific primer. QIAGEN's accused process lacks a second target-specific primer for multiple reasons.

The inventors acted as lexicographers in describing the specific sequences that both the adaptors and claimed primers must contain. Appx97-98 col. 16:64-17:13. Of note, the specification defines the "second target-specific primer" as follows:

> a single-stranded oligonucleotide comprising a 3′ portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (b), and a 5′ portion comprising a nucleic acid sequence that is identical to a second sequencing primer. The second target-specific primer is nested with respect to the first target-specific primer.

Appx97 col. 16:33-41. The district court adopted this definition except for the requirement that second target-specific primer is "nested" with respect to

the first target-specific primer. Appx2.[3] Nonetheless, there are two components to the definition the court adopted: (1) the second target-specific primer must anneal to the known target nucleotide sequence that the first target-specific primer anneals to from the prior step of the method, and (2) it must contain sequence "identical to a second sequencing primer." QIAGEN's accused primer meets neither of these requirements.

Likewise, the term "second adaptor primer" is defined as:

> a nucleic acid molecule comprising a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first adaptor primer.

Appx98 col. 17:5-9. The district court adopted this definition verbatim, which requires the "second adaptor primer," like the first adaptor primer,[4] to include "sequence identical to a portion of the first sequencing primer." Appx2. By definition, the ligated universal oligonucleotide tail adaptor must also include sequence "identical" to sequencing primer sequence. Appx95 col. 11:37-40. This language ("identical to a portion") is different from the definition for

[3] QIGEN asserted that "nested" was part of the definition. Appx9-11. "Nested" means "annealed to a nucleic acid sequence 3′ downstream and in the same direction as another molecule." Appx2. Essentially, the first primer lands on the "left" and the second primer lands inside of the first primer, toward the "right." Appx518 at 211:25-212:7.

[4] The first adaptor primer also must have a sequence identical to a portion of the first sequencing primer. Appx97 col. 16:64-66.

Tecan Ex. 2207, Page 29 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

"second target-specific primer" which requires "identical" without any qualification.

One basis for non-infringement is that the second target-specific primer only includes nineteen out of the thirty-four nucleotides in common with what is accused of being the sequence of the second sequencing primer. It is therefore not "identical."

D. The Asserted '597 Patent

The '597 issued a year into suit, adding two additional inventors and reciting a specification that significantly overlaps with the '810 patent. Dr. Iafrate (inventor on both patents) described the '597 patent as a way to "improve" on the methods of the '810 patent by requiring fewer steps and not using ligated tail adaptors. Appx512-13 at 188:1-189:23, 191:16-17 ("So in some ways it looks like the '810 patent, but it's not using the adapter."). Instead of attaching adaptors to the ends of each fragment of DNA, it uses a plurality of random primers that take the place of the adaptors. The random primers must have tails containing the sequences the sequencing machine manufacturers require for affixing the DNA to the machine. In the asserted claims, there are no references to adaptors or tail adaptors. Appx175 col. 77:11-48 (claim 1), 77:57-60 (claim 5), 78:48-51 (claim 19); Appx513 at 190:8-191:6.

Tecan Ex. 2207, Page 30 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

The ’597 patent specification contrasted the inventions with prior traditional sequencing methods that generated sequence information randomly (e.g. “shotgun” sequencing) or between two known sequences which are used to design primers. Appx139 col. 6:27-29. Thus, instead of using PCR primers designed against known sequences at each end, like the ’810 method, the methodology described in the ’597 patent uses the target-specific primer at one end and a plurality of random primers at the other end.

1. A method of preparing nucleic acids for analysis, the method comprising:

(a) ***contacting***

> [1] a first nucleic acid template comprising **a sequence of a first strand** of ***a double-stranded target nucleic acid*** with
>
> [2] a complementary ***target-specific primer***
>
> > that comprises **a target-specific hybridization sequence,** under conditions to promote template-specific hybridization and extension of the target-specific primer;

(b) ***contacting*** [3] a second nucleic acid template comprising a sequence of ***a second strand*** that is complementary to the sequence of ***the first strand*** of [1] ***the double-stranded target nucleic acid*** with [4] ***a plurality of different primers*** that share a common sequence that is 5′ to different hybridization sequences,

> […]
>
> wherein, following (a) and (b),
>
> [5] an extension product is generated to contain both a sequence that is characteristic of the target-specific primer and a sequence that is characteristic of the at least one of the plurality of different primers; and

(c) ***subjecting*** the [5] extension product to an amplification reaction comprising successive rounds of polymerase extension of

> i) [6] a tail primer that comprises a 3′ sequence that specifically anneals to the complement of the [4] common sequence and that comprises a 5′ tail sequence, and
>
> ii) [7] **a primer that specifically anneals to the complement of the [2] target-specific hybridization sequence.**

As described above, QIAGEN's kits ***do*** include ligated tail adaptors with known sequences that the inventors distinguished from the '597 invention. Appx139 col. 6:26-34. And as a result of trying to shoehorn QIAGEN's kits into the '597 patent claim language, Plaintiffs accuse one of QIAGEN's "universal" primers (the Forward Primer—which anneals to complements of the universal and artificial known tail adaptor sequences in QIAGEN's method) of being the "target-specific primer" for purposes of the '597 patent while referring to the GSP's that are designed to anneal to known genetic sequences, as the "plurality of different primers." The common adaptor primers in the QIAGEN accused products cannot be flipped around to satisfy the "target-specific primer" requirement in the '597 patent. Those common adaptor primers do not contain the claimed sequences.

A second non-infringement issue is that the only molecule Plaintiffs ever identified as the double-stranded target nucleic acid molecule undisputedly leads to the conclusion that the accused process lacks a target-specific primer under Plaintiffs' theory.

E. The District Court Case

Plaintiffs filed their original complaint the day the ’810 patent issued on July 10, 2018. Appx78, Appx185. On September 30, 2019, Plaintiffs requested leave to file an amended complaint adding what became the ’597 patent after the USPTO allowed the pending claims but before the ’597 patent issued. Appx129, Appx192. The ’597 patent issued on October 22, 2019 and Plaintiffs filed its amended complaint on October 30, 2019. Appx129, Appx194. QIAGEN had no prior notice of the ’810 or ’597 patents before being sued.

Plaintiffs’ original complaint included allegations of trade secret misappropriation, false advertising, breach of fiduciary duty, deceptive trade practices, tortious conduct, and other claims. Appx250-68. QIAGEN filed a motion for summary judgement on all but the fiduciary duty claims on March 1, 2021 and Plaintiffs responded by withdrawing those claims. Appx207, Appx331. Plaintiffs then accepted a $100 offer of judgment under Federal Rule of Civil Procedure 68 for the breach of fiduciary duty claims. Appx348.

On December 12, 2019, the district court construed terms from the ’810 patent. Appx1-17. The district court subsequently construed terms from the ’597 patent. Appx18-30. In the fall of 2020, non-party Invitae Corp. acquired ArcherDX, Inc. Appx538 at 288:7-8.

Tecan Ex. 2207, Page 33 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

On August 21, 2021, just prior to trial, the district court further construed one of the '810 patent terms—"portion of the known target nucleotide sequence" that is part of the definition of "second target-specific primer"—to mean "a portion of the ***same*** known target nucleotide sequence that is referenced in step (i) of claim 16." Appx808-09. Based on that construction, the court granted summary judgment of no literal infringement of the '810 asserted claims. *Id.* Plaintiffs proceeded to trial under the doctrine of equivalents ("DOE") for the '810 patent. Appx921. QIAGEN requested that the court construe two additional terms, "sequencing primer" and "identical" and proffered definitions to the court, but it did not give the jury any instruction on the meaning of those two terms. Appx848.

Before trial, the parties contested whether Plaintiffs should be able to recover foreign damages for methods performed overseas. Over QIAGEN's objection, the district court included Plaintiffs' instruction on foreign damages. Appx974-77.

The district court held a jury trial from August 23-27, 2021. The jury returned a verdict that QIAGEN willfully infringed the asserted claims of the '810 and '597 patents, that none of the asserted claims were invalid, and awarded damages, including lost profits damages in connection with QIAGEN's RNA-related products used in the U.S. in the amount of $841,756,

Tecan Ex. 2207, Page 34 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

reasonable royalty damages for sales of products used in the U.S. in the amount of $1,593,762, and reasonable royalty damages for sales of DNA- and RNA-related products used outside the U.S. in the amount of $2,240,303. Appx966-73.

On September 30, 2022, the district court denied QIAGEN's renewed motion for judgment as a matter of law or new trial but granted its motion for remittitur, reducing the $1,5973,762 award for sales of products used in the U.S. to $752,006 as the highest amount supported by the evidence for a royalty on sales of QIAGEN's DNA-related products in the U.S. Appx58-59. The district court also granted-in-part Plaintiffs' motion for an ongoing royalty, supplemental damages, and pre- and post-judgment interest but denied Plaintiffs' request for enhanced damages and a permanent injunction with leave to renew the latter after an evidentiary hearing. Appx60-70.

After the district court's order issued, Plaintiffs notified the district court on November 21, 2022, that it intended to pursue a different scope of injunction based on harm to non-party, Invitae, which had announced plans to unload the Archer business on July 18, 2022. Appx978-81. The district court denied Plaintiffs' request to pursue this new scope of a permanent injunction as untimely. Appx982-86. The district court entered a Final Judgment on August 8, 2023, from which this appeal is taken. Appx73-77.

Tecan Ex. 2207, Page 35 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

SUMMARY OF THE ARGUMENT

1. QIAGEN's accused kits do not infringe asserted claim 16 of the '810 patent. The claims include very specific definitions of the primers, their sequences, and where they are supposed to anneal and for what purpose. By definition, the "second target-specific primer" must contain sequences "identical" to a "second sequencing primer." There is no infringement because Plaintiffs assert a primer called "Read2" is the "second sequencing primer" for purposes of the definition. The "Read2" primer comprises thirty-four nucleotides. QIAGEN's SIP—accused of being the "second target-specific primer"—does not include the thirty-four Read2 nucleotides. It only includes nineteen of them. Thus, it is not identical: a sequence of nineteen is not identical to a sequence of thirty-four.

In addition, the "universal" QIAGEN SIP cannot be the "second target-specific primer" because the SIP does not anneal to a portion of the known target nucleotide sequence. The district court determined there was no literal infringement for this reason. Appx804-09. The SIP cannot be a target-specific primer under the doctrine of equivalents because it contains no sequence specific to the target. The point of the invention of the '810 patent is to have successive rounds of first and then second target-specific primers to enrich the sample and prevent amplification of non-target sequences. The

Tecan Ex. 2207, Page 36 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

"universal" SIP does not perform this function. And if the doctrine is permitted to expand the claims to cover non-nested primers directed to non-target sequences, then the claims lack support under 35 U.S.C. § 112.

Similarly, two other claim terms (the adaptor primers and the adaptor) require the presence of sequence identical to a portion of a first sequencing primer. But the primer that Plaintiffs assert is a "first sequencing primer" is not a sequencing primer; it is P7 sequence, which is sequence that Illumina requires to immobilize or "glue" the DNA to be sequenced in the Illumina instrument, not sequence used by a primer to initiate read out of nucleotide sequence. It has no other function aside from "gluing" the sample to the surface of the flow cell used in the sequencer and does not participate in analyzing (i.e., reading) the target gene sequence.

2. QIAGEN does not infringe the '597 patent. First, Plaintiffs turned their infringement theory on its head by accusing QIAGEN's FP primer—a "universal" adaptor primer having common sequence—of being the "target-specific primer" for the '597 patent. Neither the law nor the patents countenance such a switch. There is fundamental difference in both patents between primers that anneal to target sequences to be analyzed and those that anneal to known common sequences that are never analyzed.

Tecan Ex. 2207, Page 37 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

In addition, Plaintiffs' theory suffers from a failure of proof. This is because step (a) of the method requires contacting a target-specific primer (the accused FP) with "a sequence of a first strand of a double-stranded target nucleic acid" and step (b) requires contacting a plurality of primers to "a sequence of a second strand of ***the*** double-stranded target nucleic acid." The terms of the claim thus require the "double-stranded target nucleic acid" to be a constant, the same in both steps. The flaw in Plaintiffs' approach is that its theory requires changing the "double-stranded target nucleic acid" between the steps. This was made clear by Plaintiffs' expert's admissions about the inability for the accused target-specific primer, the FP, to hybridize to the sequence of the first strand of the double-stranded target nucleic acid that had been identified as the starting point of the method.

3. If this Court does not reverse on liability and negate the damages award in its entirety, it should reverse on the district court's improper imposition of damages for foreign uses of the patented methods. The sale of a product is not an infringing sale of a method claim. A person or entity must practice the steps of the method in the United States to infringe under 35 U.S.C. § 271(a). Over QIAGEN's objection, however, the district court included a damages instruction that allowed recovery for foreign uses based upon the district court's reading of *Carnegie Mellon v. Marvell,* 807 F.3d 1283

Tecan Ex. 2207, Page 38 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

(Fed. Cir. 2015) ("*CMU*"), despite other decisions to the contrary. If *CMU* can be read to cover damages for method claims practiced entirely abroad, it should be confined to the very particular situation of a hypothetical negotiation for United States infringement encompassing foreign uses. No such evidence exists here. The district court's upholding of such damages was error.

STANDARDS OF REVIEW

Claim construction is reviewed de novo. *TomTom, Inc. v. Adolph*, 790 F.3d 1315, 1322 (Fed. Cir. 2015). Under Third Circuit law, JMOL decisions on questions of fact are reviewed de novo. *Biogen MA Inc. v. EMD Serono, Inc.*, 976 F.3d 1326, 1331 (Fed. Cir. 2020). JMOL should be granted where, "viewing the evidence in the light most favorable to the nonmovant and giving it the advantage of every fair and reasonable inference, there is insufficient evidence from which a jury reasonably could find liability." *Id.*

Tecan Ex. 2207, Page 39 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

ARGUMENT

I. THE ACCUSED QIAGEN PRODUCTS DO NOT INFRINGE THE '810 PATENT

A. The Accused Products Do Not Use A "Second Target-Specific Primer"

1. The SIP Does Not Have Sequence Identical To A Second Sequencing Primer

The inventors' lexicography requires the "second target-specific primer" to include a "sequence that is identical to a second sequencing primer." Appx2, Appx97 col 16:33-39. Plaintiffs accuse QIAGEN's SIP primer of being the "second target-specific primer." Plaintiffs' argued that Illumina's Read2 Primer was the "second sequencing primer" and the SIP primer includes a sequence "identical" to the Read2 sequence. Appx584 at 474:9-475:8. The evidence cannot support this conclusion.

The Read2 sequencing primer is 34 nucleotides in length. Appx619 at 611:24-612:18. That means that 34-nucleotide sequence must appear in QIAGEN's SIP to meet the definition. But QIAGEN's SIP shares only 19 of the 34 nucleotides that comprise the Read2 sequence. *Id*. QIAGEN's SIP is missing fifteen out of the thirty-four nucleotides of Read2—over 40%. That is not "identical."

The district court erred in denying JMOL by reasoning that "identical" could mean "identical to a portion." Appx42 (jury "was not precluded from

Tecan Ex. 2207, Page 40 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

finding that 'identical' may include identical to a portion, and therefore was allowed to find for Plaintiff."). This cannot be the case because the patent specification and claim 16 itself differentiate between "identical" and "identical to a portion." Indeed, the district court adopted the specification's lexicography requiring the adaptor primers to be "***identical to a portion*** of the first sequencing primer" contained in the universal oligonucleotide tail adaptor. Appx2; Appx97 col. 16:64-66 (first adaptor primer); Appx98 col. 17:5-8 (second adaptor primer), Appx95 col. 11:37-40 (adaptor containing first and second sequencing primer sequences). This requirement, referring to the adaptor sequences being identical to ***portions*** of the universal oligonucleotide tail adaptor is also stated in the claim language itself. Appx128 col. 78:6-8, 11-13.

In contrast, the second target-specific primer must contain sequence "***identical*** to a second sequencing primer." The patent is very particular in the language it uses: "identical to a portion" (adaptor primers) versus "identical" (second target-specific primer). And the difference is important: the specificity of the process, which depends on the conditions under which primers anneal to particular sequences (Appx576 at 442:12-21), is dependent on the number of complementary nucleotides. Appx96 col. 13:58-61.

Tecan Ex. 2207, Page 41 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

In addition, the inventors were specific with respect to the word "portion" modifying nucleotide sequences. The inventors instructed that when "portion" modifies a nucleotide sequences, it means "all or only a subset of the nucleotides comprised by the molecule." Appx94 col. 9:1-5. Thus, "portion" as a modifier nucleic acid includes all or some of a sequence. By choosing to require "identical" only for the second target-specific primer, without using the "portion" modifier they used throughout the specification and for the adaptor primers compared to sequencing primer sequence, the inventors specified that they were limiting the requirement to actual identity *i.e.* 34 nucleotides that were the same as those in the second sequencing primer.

When different terms are used, different meanings attach. *Chicago Bd. Options Exch., Inc. v. Int'l Sec. Exch., LLC*, 677 F.3d 1361, 1369 (Fed. Cir. 2012) ("different terms have different meanings"). And where there is a distinction in degree between claim language, the construction should reflect that distinction. *Ecolab, Inc. v. Envirochem, Inc.*, 264 F.3d 1358, 1367 (Fed. Cir. 2001) (qualifying term "substantially" must be given meaning and broadens the term "uniform" to include some non-uniformity). Here, what the Court allowed the jury to decide renders the word "portion" superfluous.

Tecan Ex. 2207, Page 42 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

Not only is the plain meaning of "identical" not met as a matter of law, but QIAGEN requested the court further construe the term. Appx848. The court refused to do so. Allowing the jury to apply "plain meaning" to conclude that one sequence being identical to another, had the exact same meaning as the one sequence being identical to a portion of the other sequence was error.

B. QIAGEN's Sample Index Primer Cannot Be A "Second Target-Specific Primer" Under The DOE

The district court ruled that QIAGEN's Sample Index Primer ("SIP") could not literally be the "second target-specific primer" because it does not specifically anneal to a portion of the same known target nucleotide sequence that is referenced in the prior step of claim 16 and subject to annealing with the "first target-specific primer." Appx804-09. Consequently, Plaintiffs had to show this element was present pursuant to the doctrine of equivalents. The district court erred in not granting JMOL because the evidence failed to support that the QIAGEN SIP was the equivalent of the "second target-specific primer."

(a) Finding The "Universal" SIP To Be Equivalent To The Second Target-specific Primer Would Vitiate The Claim Element

As this Court recently confirmed, application of the doctrine of equivalents is "limited" and "exceptional." *VLSI Tech. v. Intel Corp.*, No.

2022-1906, slip op. at 14-15 (Fed. Cir. Dec. 4, 2023) (reversing JMOL on doctrine of equivalents); *Warner-Jenkinson Co. v. Hilton Davis Chem. Co.*, 520 U.S. 17, 40 (1997) (DOE cannot eliminate claim elements); *Carnegie Mellon Univ. v. Hoffmann-La Roche Inc.*, 541 F.3d 1115, 1129 (Fed. Cir. 2008) (affirming judgment of noninfringement under DOE). Here, the "universal" SIP is not equivalent to the required "second target-specific primer"; it is not an insubstantial substitute. Rather, it is a wholly different type of primer. The accused SIP anneals only to the common sequence added by the tail portion of the "first target-specific primer" and not to any of the known target nucleotide sequences. Appx 599 532:18-19, 532:22-25, 533:4-10. Indeed, all of the QIAGEN primers—except for the Gene Specific Primer—are "universal" primers that anneal to common sequences.

The Court's reasoning in *Power Integrations, Inc. v. Fairchild Semiconductor Int'l, Inc.*, 843 F.3d 1315 (Fed. Cir. 2016), applies here. In that case, the claim limitation required "a controller to generate the switching signal and to control the switching signal in response to ***a first feedback signal*** associated with a voltage control loop and ***a second feedback signal*** associated with a current control loop," i.e., there were two distinct "feedback signals." *Id.* at 1342 (emphasis added). The accused infringer had a single "feedback signal," but the jury found infringement under the doctrine of

Tecan Ex. 2207, Page 44 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

equivalents. *Id.* at 1343. The Court reversed the denial of JMOL, holding that "a finding of infringement under the doctrine of equivalents would vitiate the requirement" of two feedback signals. *Id.* at 1344. Similarly, in the case before this Court, the '810 patent inventors specifically required two separate target-specific primers to anneal to the known target nucleotide sequence. These target-specific primers had to work in conjunction with other two other primers, the adaptor primers, that have their own requirements and functions, to achieve a final goal. By allowing the jury to find that another primer—the SIP, that does not anneal to the required sequence—nonetheless satisfies this requirement, because the adaptor primer it is paired with performs a similar function, eliminates the claim limitation and blurs the requirement of step iii requiring two distinct primers each having different functions.

In the context of this patent, there is a fundamental difference between the "target" and other sequences within the DNA. The '810 patent confirms that the "specific" in "second target-specific primer" requires enrichment for ***target sequence*** by annealing to ***target sequence***, not common sequence. Appx93 col. 8:49-55. The district court's denial of JMOL was thus error because it eliminated the requirement of a second target-specific primer.

Tecan Ex. 2207, Page 45 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

(b) QIAGEN's SIP Fails The Function/Way/Result Test For DOE

The function/way/result test also demonstrates that denial of JMOL was error. *Genentech, Inc. v. Wellcome Found. Ltd.*, 29 F.3d 1555, 1567 (Fed. Cir. 1994) (equivalency requiring "substantial identity of function, way, and result."). "If any one of the prongs is unsupported, the finding of equivalency cannot stand." *Id*.

(i) QIAGEN's SIP Has A Different Function From The "Second Target-Specific Primer"

The function of a target-specific primer in the '810 patent is to anneal to the known target nucleic sequence to provide a second specificity step to achieve the desired specificity; it is a second "panning for gold" step that does not amplify unwanted sequences. Appx93 col. 8:49-55; Appx107 col. 36:28-32; Appx618 at 607:22-608:3. The function of a "***second*** target-specific primer" is to increase enrichment for the target sequences over non-target sequences. Here, the target sequence for the second target-specific primer is explicitly defined as "the known target nucleotide sequence." Appx97 col. 16:35-37. The '810 patent defines what it means to be specific with respect to a primer: a target-specific primer will anneal to the target nucleotide sequence and mediate amplification of that sequence and it will ***not*** anneal to

Tecan Ex. 2207, Page 46 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

non-target sequences and mediate amplification of those non-target sequences. Appx93 col. 8:49-55.

In contrast, QIAGEN's SIP binds to common sequence. Appx581 at 461:23-25. The SIP's function is to multiply what is present; not to differentiate target from non-target sequences. The SIP adds sequence that will be compatible with the sequencer. Appx587 at 487:5-8. Indeed, the SIP does not anneal to the DNA target sequence from the patient sample. Appx612-13 at 586:19-22, 587:2-6, 587:19-23. The SIP cannot distinguish between what is target sequence and what is not. Its function within the method is different from the function of a "second target-specific primer." Appx521 at 224:1-8 (inventor testimony distinguishing between target sequence and common sequence); *Genentech, Inc. v. Wellcome Found. Ltd.*, 29 F.3d 1555, 1567 (Fed. Cir. 1994) (defining the function "for purposes of equivalency analysis" must be made "in the context of the patent, the prosecution history, and the prior art").

To the extent Plaintiffs argue that the SIP works in combination with the adaptor primer, called the Universal Primer (UP), to somehow perform the same function as a second target-specific primer, that argument ignores that the UP has its own specific function as an adaptor primer. Appx592 at 504:24-505:11 (defining function of adaptor primer). The '810 specification already

Tecan Ex. 2207, Page 47 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

recognized that the second adaptor primer would work in conjunction with the second target-specific primer (Appx90 col. 2:11-13), but nonetheless required that the second target-specific primer perform its own function of creating specificity (Appx97 col. 16:46-50; Appx98 col. 18:24-29; Appx107 col. 36:28-32), which the accused SIP does not do. No evidence here indicates that the accused UP is some form of "super adaptor primer" that not only performs its function but that of the target-specific primer as well.

(ii) The Way QIAGEN's Sample Index Primer Works Is Different From The Claimed "Second Target-Specific Primer"

Even assuming the district court properly defined the function of the SIP and the second target-specific primer to be the same, it erred in finding that there was substantial evidence supporting a finding that the way the SIP works is the same as the claimed "second target-specific primer." Appx43-44. The way that the "second target-specific primer" works to amplify target nucleic acid sequences is to anneal to and amplify the ***known target nucleotide sequence*** and not anneal to sequences that are not the known target nucleotide sequence. And, not only does the specification say that it anneals to known target nucleotide sequence, but every example in the patent shows that it anneals in a nested fashion with respect to the first target-specific

primer.[5] Appx518 at 211:14-17 ("Q. And it is fair to say that in every single example that you disclosed in the '810 patent, the target-specific primers are nested; right? A. I think that's true, yes."); Appx97 col. 16:33, 16:39-41. The nesting provides both a double check and further specificity. Appx521 at 222:17-223:1; Appx534 at 276:4-8.

Here the SIP bears no resemblance to the second target-specific primer. It does not anneal to known target nucleotide sequence. It does not do so in a nested fashion. It does—as opposed to the requirement that it must not—anneal to sequences other than the known target nucleotide sequence and it fails to provide specificity, a function that Plaintiffs pass off to the second adaptor primer, which has a function of its own to perform which is not creating specificity. Appx521-22 at 224:1-8, 225:9-15; Appx533 at 272:15-21.

This case is similar to *Amgen Inc. v. Sandoz Inc.*, 923 F.3d 1023, 1029 (Fed. Cir.), *reh'g granted, opinion modified*, 776 F. App'x 707 (Fed. Cir. 2019). In that case, the patentee attempted to argue that the claimed multi-step process was equivalent to a one-step process. The Court rejected

[5] QIAGEN asserted that "nesting" was also part of the definition of "second target-specific primer." Appx9-11; Appx97 col. 16:33-41 ("As used herein…. The second target-specific primer is nested with respect to the first target-specific primer.").

Tecan Ex. 2207, Page 49 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

patentee's arguments. *Id.*; *Virnetx, Inc. v. Cisco Sys., Inc.*, 767 F.3d 1308, 1323 (Fed. Cir. 2014) (rejecting DOE where claim required "encrypted channel" and specification distinguished that from a "secure channel" and accused infringer only used a "secure channel").

(iii) The Results of QIAGEN's Method Are Different From The Claimed "Second Target-Specific Primer"

In the QIAGEN workflow, the "universal" SIP primer used with the "universal" UP primer results in amplifying ***everything*** that has the requisite common sequence from the first round of PCR, rather than further enriching for the target nucleic acid of interest. Appx699 at 929:15-21; Appx581 at 461:23-25. Thus, when the "first target-specific primer" (QIAGEN's GSP) mis-primes by annealing to non-target sequence, neither the accused SIP nor the accused UP will be able to remove that mis-priming event from the pool of nucleic acids to be amplified. That is the antithesis of what the '810 patent describes in that the "second target-specific primer" is supposed to remove those events.

Thus, in every material respect, the accused "universal" SIP primer is not equivalent to the "second target-specific primer."

C. If Non-Nested SIP Can Be Expanded Under The DOE To Be A "Second Target-Specific Primer" Then The Claim Fails For Lack Of Written Description

The upshot of the application of DOE to a non-nested "universal" primer, is the claim is now expanded far beyond what is disclosed in the specification as being in the possession of the inventors. The "test for sufficiency is whether the disclosure of the application relied upon reasonably conveys to those skilled in the art that the inventor had possession of the claimed subject matter as of the filing date." *Ariad Pharms. v. Eli Lilly*, 598 F.3d 1336, 1351 (Fed. Cir. 2010) (reversing denial of JMOL on written description). The specification and inventor testimony demonstrates the '810 patent inventors were not in possession of a non-nested invention.

QIAGEN's SIP primer is not "nested" with respect to the first target-specific primer (the GSP), meaning it does not anneal to the known target nucleotide sequence 3′ downstream and in the same direction as the first target-specific primer. Appx2 (claim construction of "nested"); Appx603 at 549:6-10.

Not only does the specification state that the second target-specific primer "is" nested with respect to the first target-specific primer (Appx97 col. 16:39-41), but every example in the specification uses nested target-specific primers (Appx518 at 211:14-17) (Inventor Iafrate agreeing that "in every

single example you disclosed in the '810 patent, the target-specific primers are nested"); Appx601 at 540:24-541:4 (Plaintiffs' expert agreeing that there is no "example in the '810 patent where the second target-specific primer wasn't nested with respect to the first").

Using primers pairs where one is nested and one is not, results in a "hemi-nested" configuration. The specification explicitly instructs that such a configuration can result in "problems." Appx98 col. 17:35-45. Indeed, Inventor Iafrate agreed with the following question: "at the time you filed your patent application, you did not include non-nested gene specific primers because you had never gotten that to work in a satisfactory way; correct?" Appx519 at 214:12-18. In every instance where the inventors described their invention, including in their invention disclosure statement, they emphasized the need to use nested target-specific primers. Further, when the inventors saw how QIAGEN had designed its products not using nested primers, they admitted that QIAGEN had done something they could not do. Appx523 at 230:19-231:2; Appx535 at 278:15-279:24; Appx987; Appx1018-20. The inventors Le and Zheng tested a non-nested approach and were left with a specificity of only 6.8%. Appx535 at 278:15-279:24; Appx1018-20. It would be odd indeed to find that the inventors had described a methodology which they consistently admitted they could not get to work properly and therefore

Tecan Ex. 2207, Page 52 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

purposefully did not describe in their patent. And it is entirely appropriate for those admissions to support a finding of no written description for a non-nested method. *Nuvo v. Dr. Reddy's Lab*., 923 F.3d 1368, 1381 (Fed. Cir. 2019). Here the lack of written description is even weaker than in *Nuvo*, where the inventors at least expressed a wish or hope that a compound would work. *Id*. For the '810 patent, there is no such wish, and the specification actually teaches away from hemi-nested methodologies like QIAGEN uses.

D. The District Court Erred With Respect To The "Sequencing Primer" Requirement Of The Claims

The term of art "sequencing primer" arises in definitions of claim terms: the first and second adaptor primers in claim 16 must have "sequence that is identical to a portion of a ***first sequencing primer***." Appx97-98 col. 16:66-17:7; Appx2. This requirement is not met.

To meet the definition of "first adaptor primer," Plaintiffs identified a sequence called P7 as the "first sequencing primer" sequence. Appx583 at 468:15-18. P7, however, is not a sequencing primer. It is a different type of primer.

The issue boils down to the following: Plaintiffs claim that the sequence of any primer used in the overall sequencing process is a sequencing primer. QIAGEN, on the other hand, asserts that a sequencing primer is one

Tecan Ex. 2207, Page 53 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

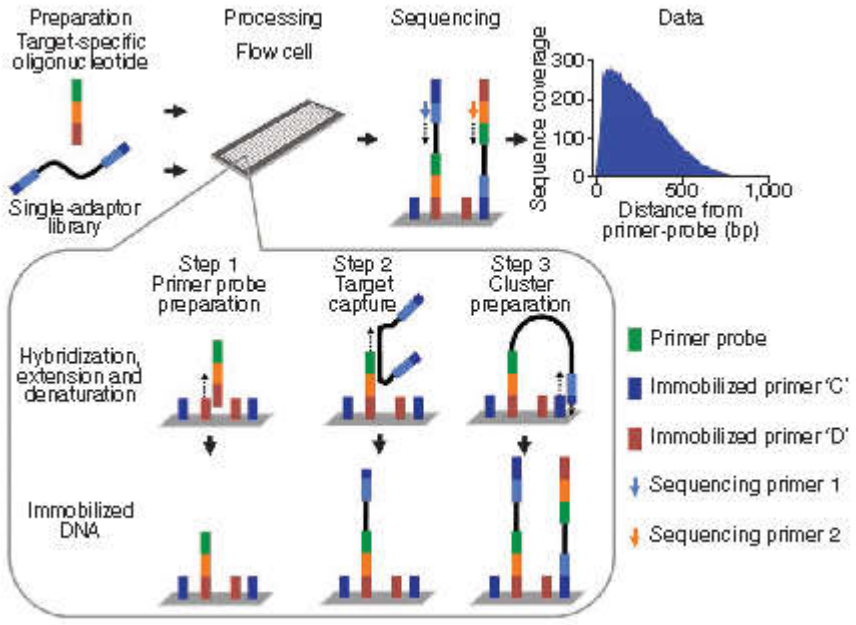
that reads out the sequence of nucleotides. QIAGEN's definition is consistent with the specification and the common understanding in the field.

The specification provides that "the sequencing step relies on the use of a first and second sequencing primers." Appx99 col. 19:35-39; Appx108 col. 38:17-21 ("After sequencing, the reads are mapped to the human genome…."). Figure 9 also draws a distinction between primers used anywhere in the sequencing process called "sequencing platform primers" and those used to read out the identity of nucleotides called "sequencing primer binding sites." Appx89. An example of sequencing primers are the Illumina primers named "Read1" and "Read2." These sequences will be recognized by the primers within the instrument used to sequence the nucleic bases as Illumina publications confirm. Appx616 at 601:6-19 (referring to "Read1" and "Read2" sequences as "binding sites that initiate sequencing").

The P7 primer, however, is not a sequencing primer like Read1 or Read2. P7 is a sequence that Illumina created to anchor or "glue" DNA fragments to the flow cell surface within their instruments where they will be sequenced. Appx522 at 225:19-226:6; Appx596 at 523:5-18. Illumina itself calls P7 "sequences that allow the library to bind and generate clusters on the flow cell" and refers to Read1 and Read2 sequences as "sequencing primer binding sites," i.e., sequence used to identify and read out nucleotides. Appx

615-16 at 598:1-601:19. Plaintiffs' expert agreed with this characterization, admitting that the P7 sequence is used to "glue" the fragments to be sequenced into the instrument. Appx596 at 523:7-23, Appx615 at 595:16-596:16; Appx614-15 at 594:22-595:15 (Read1 and Read2 primer sequences function to read out the identity of the nucleotides).

Those of skill in the art understand the P7, Read1, and Read2 sequences the same way. PX-1037, an article from Nature Biotechnology referencing Illumina primers distinguished between Illumina sequencing primers and Illumina primers that immobilize or "glue" DNA on the flow cell. Appx617 at 604:6-20; Appx1111-16. Figure 1 from the article provided:



Appx1112. As can be seen, primers like P7 are referred to as "immobilized primers" and the primers used for sequencing such as Read1 and Read2 are

"sequencing primers." There is clear difference based on their function. Appx618 at 607:7-21. Indeed, Dr. Lennon did not claim those of skill in the art would refer to the "P7" sequence as a sequencing primer. Appx615 at 596:21-597:2. Rather, he testified that those of skill in the art "typically refer to the primers in the process by what they do." Appx615 at 597:3-8.

Because it was error to allow the jury to conclude that any primer sequence used in the overall sequencing process is a sequencing primer, after QIAGEN requested construction of the term, which the court declined to construe, the finding of infringement should be reversed. Appx848; Appx738 at 1086:20-1087:5; Appx917-18; *O2 Micro Int'l v. Beyond Innovation Tech.*, 521 F.3d 1351, 1360 (Fed. Cir. 2008) (court must resolve disputes regarding claim scope).

II. THE ACCUSED QIAGEN PRODUCTS DO NOT INFRINGE THE '597 PATENT

A. QIAGEN's Accused "Forward Primer" Cannot Be A "Target-Specific Primer"

For the '597 patent, Plaintiffs reverse positions and no longer accuse QIAGEN's gene specific primer of being a "target-specific primer." Instead, Plaintiffs accuse QIAGEN's Forward Primer (FP)—which anneals to common sequence—of being the "target-specific primer." Appx597 at

Tecan Ex. 2207, Page 56 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

525:24-526:2. FP is the same primer that Plaintiffs accused of being a "universal" adaptor primer for the '810 patent. Appx585 at 478:20-23.

1. QIAGEN's FP Contains Common Sequences Not Target-specific Hybridization Sequences

For the '597 patent, the district court construed "target-specific primer" as follows:

> [A] primer that has ***a level of complementarity between the primer and the target*** such that there exists an annealing temperature at which the primer will anneal to and mediate amplification of the target nucleic acid and ***will not anneal to*** or mediate amplification of ***non-target sequences*** present in a sample.

Appx18-19 (emphasis added).

In contrast, QIAGEN's accused FP anneals to the complement of artificial sequence added onto the sample. The '597 specification defines such artificial sequence as "common sequence" rather than "target sequence." Appx144 col. 16:59-61; Appx132 fig. 1A, step 101 (adding common sequence); Appx133 fig. 1B (same). These common sequences are non-target sequences—exactly what the construction requires the "target-specific primer" to avoid. There is a clear distinction between the target, *e.g.*, the gene of interest, and "non-target sequences." The FP anneals to non-target sequences and therefore cannot be the target-specific primer.

Tecan Ex. 2207, Page 57 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

2. Plaintiffs' "Target" Molecule Theory Cannot Be Upheld

In denying JMOL, the district court relied on two pieces of evidence to reach the conclusion that the jury had substantial evidence supporting infringement: (1) testimony from Plaintiffs' expert that "'the FP primer in the QIAGEN workflow, targets the ligated adaptor' and thus satisfies the Court's construction of 'target-specific primer,'" (Appx591 at 501:2-502:25) and (2) testimony that "the FP binds only to 'a very small proportion of the molecules' in a sample" (Appx620 at 617:11-618:2); Appx37-38. Plaintiffs' theory is that because adaptor sequences are added to the DNA fragment that contains the target sequence, then the FP primer that anneals to those newly-added adaptor sequences is annealing to the same molecule as the actual target sequence. This argument fails.

First, it cannot be that what was an adaptor primer for the '810 patent is now a target-specific primer for the '597 patent. It is the same primer. Plaintiffs' argument (adopted by the district court) is that the definitions of target-specific primer are different between the patents. But the differences are not material and do not justify the flip-flop.

'810 Patent Definition	***'597 Patent Definition***
"second target-specific primer" means "a single-stranded oligonucleotide comprising a 3'	"target-specific primer" means "a primer that has a level of complementarity between the primer

Tecan Ex. 2207, Page 58 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (b), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer" (Appx93)	and the target such that there exists an annealing temperature at which the primer will anneal to and mediate amplification of the target nucleic acid and will not anneal to or mediate amplification of non-target sequences present in a sample" (Appx18-19)

Second, Plaintiffs' position adopted by the district court fails because QIAGEN's FP does exactly what the construction says it cannot do: "anneal to or mediate amplification of ***non-target*** sequences present in a sample." Appx19 (emphasis added). There is a fundamental distinction in the patent and claim construction between "target" sequences—what is of interest and derived from the patient sample—and "non-target sequences." Artificial sequences added to a molecule are never referred to as target sequence, and therefore a primer that anneals to those sequences cannot be a "target-specific primer." Appx144 col. 16:59-61; Appx155 col. 38:15-25 (defining common sequence to differentiate it from target sequence).

Furthermore, trial testimony from all witnesses agreed that adaptor sequence is never analyzed—it is not the "target" of the process. For instance, Illumina's P7 sequence is included in the QIAGEN adaptor, yet none of the inventors ever analyzed that sequence. Appx522 at 227:25-228:4. The P7 adaptor sequence is not indicative of any disease state and in fact, it is thrown

Tecan Ex. 2207, Page 59 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

away in the rare instances where the sequence is read out. Appx596 at 521:11-522:21.

Similarly, Inventor Le testified that the adaptor sequence is common sequence that is ligated to both target and non-target sequences in a sample. Appx532 at 268:4-19. Thus, inventor Le confirmed that adaptor primers like QIAGEN's FP are "common primers" incapable of distinguishing between target and non-target fragments. Appx533-34 at 270:6-19, 272:15-273:6. Inventor Iafrate testified that adaptors are designed to be on every fragment in a sample. Appx522 at 225:5-8. The adaptor primers in such a sample do not enrich for anything, they "make random copies of everything" containing the adaptor sequence and would not be sifting the unwanted dirt from the gold (the target nucleic acid). Appx522 at 225:9-18.

Plaintiffs' expert Dr. Lennon concurred. He testified that the QIAGEN FP primer, which is designed to anneal to common adaptor sequence that is ligated to all fragments, will bind to all sequences in a sample. Appx599 at 534:19-21. Dr. Lennon further testified that it is the QIAGEN GSP, not the FP, that is responsible for selecting and amplifying the molecules of interest. Appx621 at 619:2-10. Nothing in the record supports a finding that a "universal" primer like QIAGEN's FP is a "target-specific primer."

Tecan Ex. 2207, Page 60 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

B. Insufficient Evidence Established That Sequences In The First And Second Strands Of The "Double-Stranded Target Nucleic Acid" Interacted With The Accused Primers To Satisfy Steps (a) And (b) Of Claim 1

It is axiomatic that when a method claim requires contacting parts of a molecule with various components, the molecule must actually exist. Here, Plaintiffs' infringement theory on the '597 patent can succeed only by attempting to ignore this fundamental truth. In relevant part, the claim requires two contacting steps with specified molecules:

> "(a) contacting a first nucleic acid template comprising a sequence of a first strand of ***a* double-stranded target nucleic acid** with a complementary target-specific primer that comprises a target-specific hybridization sequence to promote template specific hybridization …"; and
>
> "(b) contacting a second nucleic acid template comprising a sequence of a second strand that is **complementary to the sequence of the first strand of *the* double-stranded target nucleic acid** with a plurality of different primers … to promote template specific hybridization…"

Appx175. As can be seen, under basic principles of antecedent basis, the double-stranded target nucleic acid must be the same in steps (a) and (b).

1. Plaintiffs Expert Admitted That Step (a) Is Not Satisfied

On direct examination, Plaintiffs' expert did not present any evidence of what molecule was the "double-stranded target nucleic acid" and therefore what could be the sequence of strand one and strand two of that molecule for

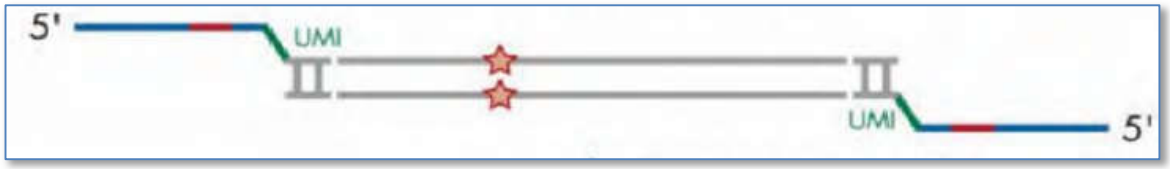
Tecan Ex. 2207, Page 61 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

any of the steps.[6] The following cross-examination exchange of Plaintiffs' expert confirms this lack of evidence:

> Q. So I looked through every single one of your slides, and I couldn't find any place where you actually put up a molecule that you call the double-stranded target nucleic acid for the '597. Am I wrong?
>
> A. I have to look through the slides, but I believe that might be correct.

Appx610 at 576:11-16. Plaintiffs' expert later confirmed, "I don't believe we presented it that way." Appx610 at 576:24. Thus, despite having the burden of proof, Plaintiffs did not want to identify for the jury the actual molecules that are the "double-stranded target nucleic acid."

Under cross-examination, Plaintiffs' expert admitted that the following image from JX-8, the third molecule from the top in the QIAGEN workflow, was the "double-stranded target nucleic acid" for step ***(b)*** of the claim. Appx613 at 590:13-19 (identifying the steps in JX 8.12 of the QIAGEN workflow); Appx614 at 591:22-25, 592:7-13 ("Step B starts on that molecule" referencing the image below).



[6] Plaintiffs infringement analysis for the '597 patent starts with (b) and then proceeds with (a). Appx588 at 491:1-21.

Given that both steps (a) and (b) must refer to the same double-stranded target nucleic acid molecule with the same first and second strands, regardless of which step one begins with, this is the double-stranded target nucleic acid molecule which defines the sequence of the first strand that the target-specific primer must hybridize to and the sequence of the second strand that the plurality of primers must hybridize to. Dr. Lennon then admitted that the FP, which he accused of being the "target-specific primer" that can satisfy element 1(a) cannot anneal to any sequence in any strand of that molecule. Appx614 at 592:9-22.

The district court erred by allowing the jury to decide, as a question of fact, that ***different*** double-stranded target nucleic acid molecules could satisfy steps 1(a) and 1(b) despite the antecedent basis ("the double-stranded target nucleic acid") referred to in step 1(b). Appx39-40. Dr. Lennon admitted that the molecule depicted above was what he relied on in analyzing infringement for step 1(b) and further admitted that this molecule could not satisfy step 1(a).

The district court appears to believe that Dr. Lennon's reference to some other molecules could be ***the*** double-stranded target nucleic acid that the jury could rely on to find infringement of elements 1(a) and 1(b). Appx39 (citing Appx624 at 634:1-635:5). This was error. Plaintiffs' expert simply saying other molecules exist without identifying what they are in any admitted

Tecan Ex. 2207, Page 63 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

evidence is not sufficient to support a finding that steps 1(a) and 1(b) were met. Appx612 at 585:7-11 ("I believe in my report I had it a couple of double-stranded target nucleic acids."); Appx612-13 at 586:23-587:1 ("I have said that there is other double-stranded target nucleic acids."); *CommScope Techs. LLC v. Dali Wireless Inc.*, 10 F.4th 1289, 1298 (Fed. Cir. 2021) (reversing denial of JMOL, noting that patentee's mapping of the claims to the accused products was "unclear," and holding, "[t]he burden is on a patent owner to show that 'the properly construed claim reads on the accused device exactly.'") (internal quotation omitted).

2. On JMOL, The District Court Erred In Allowing The "Double-Stranded Target Nucleic Acid" To Change Between Steps

Plaintiffs cannot solve the problem of a lack of evidence by changing the double-stranded target nucleic acid molecule and sequences of its first and second strands throughout the different steps of the process. In a pretrial summary judgment hearing, the district court was troubled by the possibility of that occurring:

> THE COURT: Double-stranded target nucleic acid, do you disagree that that is the thing in the red box?
>
> MR. RUTT: That's not the only double-stranded target nucleic acid.
>
> THE COURT: I got to tell you, it is very confusing to me that it seems like everything is something different in a different step of the claim in claim interpretation, and maybe in these particular

Tecan Ex. 2207, Page 64 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

> claims that's okay, but it seems odd to me. The double-stranded target nucleic acid and then after we go a few steps we have something else, and what do we indicate, we call it the double-stranded target nucleic acid and it's very different. So that gives me pause.

Appx421. When the issue arose again at the pre-trial conference, Plaintiffs represented it would not change the molecule being accused throughout the process. Appx474 at 36:2-19. But that is exactly what happened before the jury and in the denial of JMOL.

(a) Claim Language And Antecedent Basis Rules Require The Molecule To Be The Same

It was error for the district court to rule that whether there could be different "double-stranded target nucleic acids" in steps (a) and (b) was "a question of fact for the jury." Appx39. The district court's initial error was to treat this question as one of indefiniteness. Appx39-40 (citing *Energizer Holdings, Inc. v. Int'l Trade Comm'n*, 435 F.3d 1366, 1370 (Fed. Cir. 2006)). But *Energizer* is inapposite. That case concerned lack of antecedent basis as an indefiniteness challenge. *Energizer*, 435 F.3d at 1369-70. Here, the question is not whether a POSA would find an antecedent basis where there is none in the claim. The issue is whether the claims require the same structure/molecule for both steps when there is an antecedent basis identified. They do.

Tecan Ex. 2207, Page 65 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

The claim language controls. Step (a) requires contacting a sequence contained in **a** first strand of a certain molecule—"**a** double-stranded target nucleic acid." Appx175 (emphasis added). And then step (b) requires contacting a sequence of a second strand that is complementary to the sequence of **the** first strand of the same molecule—"**the** double-stranded target nucleic acid"—with a plurality of primers. *Id.* (emphasis added). Step (b) then references "the double-stranded target nucleic acid" another time in requiring that the plurality of primers not anneal to the same sequence in "**the** double-stranded target nucleic acid" as any other primer. Step (b) refers back to the molecule in step (a), as do many dependent claims.[7]

The law of antecedent basis provides that where the definite article "the" is used the claim refers back to the prior recitation. Plaintiffs and the district court's position to the contrary is not consistent with the law. For example, in *Bicon, Inc. v. Straumann Co.*, 441 F.3d 945, 954 (Fed. Cir. 2006), the Court affirmed summary judgment of noninfringement where a limitation required "the bore having a taper generally matching that of the conical

[7] In the file history, Plaintiffs distinguished the Anderson prior art by stating that in the '597 invention, the extension reactions in steps (a) and (b) created sequence contained in a single molecule whereas the extension reactions in Anderson created sequence contained in two different molecules. Appx1126.

Tecan Ex. 2207, Page 66 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

surface portion of ***the abutment***." *Id.* at 948, 954. The Court rejected the patentee's argument that this limitation could be met by pointing to ***any*** abutment. *Id.* at 954 ("the" abutment refers to a particular abutment). Here the "double-stranded target nucleic acid" of step (b) is necessarily referring back to the molecule of step (a). In fact, it is even more critical to adhere to the law of antecedent basis in this claim than in *Bicon*, because the first strand referred to in step (a) and the second strand referred to in step (b) are defined by reference to each being complementary to the other and being part of the same double-stranded target nucleic acid. There is simply no justification for departing from the well-established rule on antecedent basis. *Haemonetics Corp. v. Baxter Healthcare Corp*., 607 F.3d 776, 782 (Fed. Cir. 2010) (finding error because "district court… ignore[d] the antecedent basis… and fail[ed] to give effect to the claim language"); *In re Fought*, 941 F.3d 1175, 1178 (Fed. Cir. 2019) (collecting cases holding that even preamble language is "limiting when it serves as antecedent basis"); *see also Rexnord Corp. v. Laitram Corp.*, 274 F.3d 1336, 1342 (Fed. Cir. 2001) ("a claim term should be construed consistently with its appearance in other places in the same claim or in other claims of the same patent"); *Phonometrics, Inc. v. Northern Telecom Inc*., 133 F.3d 1459, 1465, 45 USPQ2d 1421, 1426 (Fed.Cir.1998) ("A word or phrase used consistently throughout a claim should be interpreted consistently.");

Tecan Ex. 2207, Page 67 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

CVI/Beta Ventures, Inc. v. Tura LP, 112 F.3d 1146, 1159, 42 USPQ2d 1577, 1586 (Fed.Cir.1997).

3. It Is Impossible For Plaintiffs To Satisfy The Elements Of Claim 1 Without Impermissibly Referring To Different Molecules

That Plaintiffs' theory improperly relies upon using different molecules for the "double-stranded target nucleic acid" for the different steps of the claim, can be seen by comparing steps (a) and (c). Under the principles of DNA complementarity, the same primer sequence cannot be used for steps (a) and (c). But Plaintiffs do accuse the same primer sequence for (a) and (c); therefore, its theory must be improperly relying on different molecules for "double-stranded target nucleic acid."

- Step (a) requires the target-specific primer to anneal to the target-specific hybridization sequence. The accused target-specific primer in the QIAGEN process is the "FP."
- Step (c) requires a primer that anneals to the complement of **the** target-specific hybridization sequence.

Thus the sequence of the primers for steps (a) and (c) must be the complement of each other. Where one has an A nucleotide, the other must have a T nucleotide.[8]

The fatal flaw in Plaintiffs' theory is that it accuses QIAGEN's UP primer of being the primer for step (c). ***But UP is the same sequence as FP***. Appx603 at 548:11-13. UP has only a single base added to the end. FP is not the complement of UP.

This demonstrates the error because Plaintiffs could only satisfy the claim elements by impermissibly changing the double-stranded target nucleic acid molecule, and the constituent sequences in the first and second strands, as they progressed through comparing the elements of the claim to the accused products.

III. THE DISTRICT COURT ERRED IN PERMITTING DAMAGES FOR FOREIGN INFRINGEMENT

Because Plaintiffs assert only method claims, it is the customers' ***use*** of QIAGEN's kits that infringes—not the kits themselves. The district court, however, allowed the jury to impose a 7% reasonable royalty on overseas uses of the kits—when such use is non-infringing under 35 U.S.C. § 271(a). *NTP,*

[8] The principle of complementarity follows the laws of Watson Crick base pairing. *Enzo Biochem, Inc. v. Applera Corp.*, 599 F.3d 1325, 1328 (Fed. Cir. 2010) (background to base pairing).

Tecan Ex. 2207, Page 69 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

Inc. v. Rsch. In Motion, Ltd., 418 F.3d 1282, 1318 (Fed. Cir. 2005) (holding "a process cannot be used 'within' the United States as required by section 271(a) unless each of the steps is performed within this country."). The district court's imposition of damages for overseas use violates the bedrock presumption against extraterritoriality. *Halo Elec. v. Pulse Elec*., 769 F.3d 1371, 1380 (Fed. Cir. 2014) ("'[o]ur patent system makes no claim to extraterritorial effect; these acts of Congress do not, and were not intended to, operate beyond the limits of the United States, and we correspondingly reject the claims of others to such control over our markets.'") (quoting *Microsoft Corp. v. AT&T Corp.,* 550 U.S. 437, 444 (2007)).

A. There Is No Legal Basis For Plaintiffs To Recover Damages For The Practice of Method Claims In Foreign Countries

1. Fundamental U.S. Patent Law Precludes Damages For Foreign Infringement Under Section 271(a)

Plaintiffs asserted infringement of method claims under Section 271(a), which provides that one who "makes, uses, offers to sell, or sells any patented invention, ***within the United States*** or imports into the United States any patented invention during the term of the patent therefor, infringes the patent." 35 U.S.C. § 271(a) (emphasis added); Appx974-77.

The QIAGEN accused products are kits that may be used to practice the methods—but the kits themselves do not infringe. Appx975. Because a

"method claim is directly infringed only by one practicing the method," *Cardiac Pacemakers, Inc. v. St. Jude Medical, Inc*., 576 F.3d 1348, 1359 (Fed. Cir. 2009) (quoting *Joy Tech. v. Flakt, Inc*., 6 F.3d 770, 775 (Fed. Cir. 1993)), the direct infringement only occurs when the kits are ***used***, not when the kits are sold. Indeed: "The law is unequivocal that the sale of equipment to perform a process is not a sale of the process." *Id*. (quoting *Joy Tech*., 6 F.3d at 773).

Because sale of a kit—whether domestic or foreign—is not sale of a process, a use of the patented method outside the United States is not infringement under 271(a). *NTP,* 418 F.3d at 1318. Applying this and other fundamental principles, it was error for the district court to deny JMOL on foreign damages.

2. The *CMU* Exception Does Not Apply Here

In allowing a damages jury instruction on foreign sales, and denying JMOL after the jury awarded such damages, the district court relied on the panel decision in *Carnegie Mellon Univ. v. Marvell Technology Group, Ltd.*, 807 F.3d 1283 (2015) ("*CMU*"). Appx974-77 (allowing jury instruction); Appx54-58 (denying JMOL on foreign damages). The district court recognized that the issue "is not whether the foreign uses of the patented method are infringing – they are not and Plaintiffs concede that." Appx974-

Tecan Ex. 2207, Page 71 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

77. The district court, however, concluded that *CMU* compelled an instruction that if "QIAGEN's infringement in the United States was a substantial cause of the sale of that [overseas] product" then the foreign sale could be subject to damages. Appx976. The district court confirmed the conclusion post-trial. Appx54-58.

The district court's interpretation of *CMU* was incorrect. Within the context of precedent and the strong presumption against extraterritoriality, *CMU* is directed to a particular type of hypothetical negotiation in which foreign revenue is accounted for as part of a domestic license. In *CMU*, defendant sold computer chips that practiced the patented method both in the United States and overseas. The *CMU* panel observed that a "physical product is being employed to measure damages for the infringing use of the patented methods," 807 F.3d at 1306; "territoriality is satisfied when and only when any one of those domestic actions for that unit (e.g., sale) is proved to be present," *id*.; and a "new trial is needed to determine whether the sales [of the foreign chips] are properly said to have been in the United States." 807 F.3d at 1307. Significantly, however, *CMU* focused on the "***hypothetical negotiation*** [that] would have employed the number of units sold to measure the value of the method's domestic use." 807 F.3d at 1307 (emphasis added).

Tecan Ex. 2207, Page 72 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

In other words, the *CMU* panel posited a hypothetical negotiation that would specifically account for foreign non-infringing uses.

The *CMU* panel's hypothetical negotiation is not present in this case because Plaintiffs' theory was a straightforward license of a single rate assuming every sold product infringed. Plaintiffs' damages expert presented to the jury a single hypothetical negotiation, starting with the exclusive license agreement between the patent owner MGH and Archer (Appx649-50 at 733:10-735:21), and then adjusting the 4.5% rate upwards. Appx650-51 at 735:22-739:4. The expert then relied on the *Georgia Pacific* factors to increase the rate to 7%. Appx651 at 739:5-742:19. He then applied a single 7% rate to the entire royalty base: "***that's everything*** except the RNA products in the U.S."[9]—including foreign sales. Appx651-52 at 742:20-743:4. Nowhere in the hypothetical negotiation is there acknowledgement or accounting for the foreign sales. They are treated exactly like domestic sales.

CMU cannot be a principle of general application that allows United States plaintiffs to reach beyond the territoriality of the patent laws to include foreign sales within a royalty base ***as if*** the infringement occurred in the

[9] The RNA products in the U.S. were subject to lost profits calculations. The DNA products in the U.S. and all overseas sales were subject to reasonable royalty.

Tecan Ex. 2207, Page 73 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

United States. If *CMU* is a valid methodology, it should be treated as a narrow exception, subject to presentation of substantial evidence of a specific type of hypothetical negotiation accounting for non-infringing foreign revenue as part of a value proposition derived from domestic infringement. There was no evidence of such a hypothetical negotiation in this case and thus *CMU* should not apply.

3. If Applicable To This Case, The *CMU* Exception Is In Conflict With Prior Precedent And Recent SCOTUS Guidance And Should Be Clarified Or Overruled

Even applied narrowly, *CMU* appears in conflict with prior precedent regarding method claims. Despite addressing liability for an overseas practice of a patented method for products that did not enter the United States, the *CMU* panel did not cite or address *Joy Tech* and its progeny that hold "sale of equipment to perform a process is not a sale of the process." 6 F.3d at 773; *Cardiac*, 576 F.3d at 1359 (same). This Court's rulings in *Cardiac* and earlier *Joy Tech* should control. *Newell Companies, Inc. v. Kenney Mfg. Co.,* 864 F.2d 757, 765 (Fed. Cir. 1988).

In admitting "the Court struggled with this issue" at trial (Appx57), the district court recognized the potential reach of *CMU* to swallow up the territorial limit of patents: "by allowing the patentee to obtain the same damages it would have obtained if the use of the patented method overseas

were an infringement, even though that use is not an infringement." Appx977. The district court's concern was correct. Indeed, the court noted that "the Court takes Defendants' point that to some extent *CMU* is difficult to reconcile with other governing law." Appx57.

QIAGEN respectfully requests that this Court clarify the *CMU* holding with respect to foreign sales as inconsistent with *Joy Tech* and *Cardiac*, and more recent Supreme Court precedent regarding territoriality. In *WesternGeco* (decided after *CMU*), the Supreme Court noted that "[c]ourts presume that federal statutes apply only within the territorial jurisdiction of the United States" and that this principle has "deep roots." *WesternGeco LLC v. ION Geophysical*, 138 S.Ct. 2129, 2136 (2018). To rebut the presumption, within the statute there must be a "clear indication of an extraterritorial application." *Id*. The statutes at issue in *WesternGeco* were Sections 271(f)(2) and 284 and "taken together" they "permit the patent owner to recover lost foreign profits." 138 S.Ct. at 2139. Section 271(f)(2) provides liability for supplying components from the United States to be "combined outside of the United States in a manner that would infringe the patent if combined within the United States." By contrast, Section 271(a) is directed by its own terms to actions ***within*** the United States.

Tecan Ex. 2207, Page 75 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

The limits on the territorial reach of United States intellectual property laws was also recently confirmed in *Abitron Austria GmbH v. Hetronic Int'l, Inc.*, 600 U.S. 412, 417 (2023). In reversing application of trademark laws overseas, the Court noted it "is a longstanding principle of American law that legislation of Congress, unless a contrary intent appears, is meant to apply only within the territorial jurisdiction of the United States." *Id.* (quoting *Morrison v. Nat'l Australia Bank Ltd.*, 561 U.S. 247, 255 (2010)). Drawing on patent precedent, the Court concluded: "'[f]oreign conduct is generally the domain of foreign law.'" *Id*. (quoting *Microsoft Corp. v. AT&T Corp*., 550 U.S. 437, 455 (2007)).

Importantly, Congress had the opportunity to address how United States patent law would treat movement of goods from the United States abroad with respect to patent infringement and chose only to impose liability for such movement under 271(f) related to apparatus claims. Congress chose not to extend the reach of 271(f) to method claims. *Cardiac*, 576 F.3d at 1359 (*en banc*). Extending the reach of *CMU* to capture method claims practiced abroad, as was permitted by the district court, is in conflict with Congressional intent. Indeed, permitting damages for overseas uses, especially when Congress has not authorized such liability, may result in conflict with foreign

Tecan Ex. 2207, Page 76 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

jurisdictions that have their own patent laws—exactly the result that the policy of international comity is designed to avoid.

By allowing a damages theory to be a springboard to capture overseas infringement—when Section 271(a) contemplates only domestic liability—*CMU* is in conflict with both recent and longstanding precedent regarding the territorial reach of United States patents.

B. Even If *CMU* Remains Good Law, It Does Not Support The Foreign Damages Jury Instruction

Over QIAGEN's objection, the district court provided the following erroneous instruction to the jury: "[d]amages may also be awarded on sales of products that practice the patented methods in their normal intended use outside of the United States if, for those products, you find that (1) QIAGEN's infringement in the United States was a substantial cause of the sale of that product, and (2) QIAGEN made or sold the product within the United States." Appx768 at 1207:9-15.

There is no support in *CMU* for "substantial cause" being the standard to swallow up overseas sales into a royalty base given the strong presumption against extraterritoriality. *CMU* provided no specific guidance but there must be significantly more connection between domestic and foreign sales than what exists here. *Halo*, 769 F.3d at 1378-1380, provides guidance for what constitutes a domestic "sale." Under *Halo*, courts should look at the sales

Tecan Ex. 2207, Page 77 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

transaction itself and related conduct. *Id*. "For such a sale, one must examine whether the activities in the United States are sufficient to constitute a 'sale' under Section 271(a) recognizing that a strong policy against extraterritoriality liability exists in patent law." *Id.* at 1378. "Substantial cause" does not meet that standard.

C. Even If The Jury Instruction Was Correct, Insufficient Evidence Supported The Foreign Damages Verdict

Even if the jury instruction was correct, QIAGEN's domestic activities were not a "substantial" cause of the foreign sales and therefore the district erred in not granting JMOL. Appx54-58.

To uphold the verdict, the district court relied upon QIAGEN's quality control "testing" occurring in the United States. Appx57-58. There are no bills of lading or other evidence of United States connections.

Indeed, the accused kits are one-time use products. Appx1039-42. Thus, quality control testing is on a representative unit from a large batch. Appx1134. That is unlike the situation were a single machine can be tested to carry out a patented method and then that same machine is shipped overseas. Thus, Plaintiffs proved no actual causation. *Halo*, 769 F.3d at 1380 (doubts about United States activity under § 271(a) "resolved by the presumption against extraterritorial application of United States laws."). Thus, even if *CMU* remains applicable law, and even if the jury instruction

Tecan Ex. 2207, Page 78 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

was correct, there was insufficient evidence connecting the overseas uses to any activities in the United States.

CONCLUSION

For the foregoing reasons, QIAGEN requests the Court reverse the decision below.

Respectfully submitted,

Dated: December 15, 2023

s/ *Brian C. Cannon*

Brian C. Cannon
QUINN EMANUEL URQUHART & SULLIVAN, LLP
555 Twin Dolphin Drive, 5th Floor
Redwood Shores, California 94065
Telephone: (650) 801-5000
briancannon@quinnemanuel.com

Attorneys for Defendants-Appellants

Tecan Ex. 2207, Page 79 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

CERTIFICATE OF COMPLIANCE WITH FED. R. APP. P. 32(a)(7) AND FEDERAL CIRCUIT RULE 32

Counsel for Defendants-Appellants certifies that the brief contained herein has a proportionally spaced 14-point typeface, and contains 13,849 words, based on the "Word Count" feature of Word for Microsoft 365, including footnotes and endnotes, excluding the parts of the brief exempted by Fed. R. App. 32(a)(7) and Fed. Cir. R. 32(b).

Dated: December 15, 2023

s/ *Brian C. Cannon*
Attorney for Defendants-Appellants

ADDENDUM

TABLE OF CONTENTS
ADDENDUM

Docket No.	Document Description	Filed	Appendix Pages
146	Memorandum Order	12/12/2019	Appx1
254	Memorandum Order	06/18/2020	Appx18
521	Memorandum Opinion	09/30/2022	Appx31
522	Order	09/30/2022	Appx71
551	Final Judgement	08/08/2023	Appx73
JX-001	United States Patent No. 10,017,810		Appx78
JX-003	United States Patent No. 10,450,597		Appx129

Tecan Ex. 2207, Page 82 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

ARCHERDX, INC. and THE GENERAL HOSPITAL CORPORATION d/b/a MASSACHUSETTS GENERAL HOSPITAL,	)	
Plaintiffs,	)	
v.	)	C.A. No. 18-1019 (MN)
QIAGEN SCIENCES, LLC, QIAGEN LLC f/k/a QIAGEN INC., QIAGEN BEVERLY, LLC f/k/a QIAGEN BEVERLY, INC., QIAGEN GAITHERBURG, LLC f/k/a QIAGEN GAITHERSBURG, INC., QIAGEN GMBH, QIAGEN N.V. and JONATHAN ARNOLD,	)	
Defendants.	)	

MEMORANDUM ORDER

At Wilmington this 12th day of December 2019:

As announced at the hearing on November 12, 2019[1], IT IS HEREBY ORDERED that the disputed claim terms of U.S. Patent No. 10,017,810 ("the '810 Patent") are construed as follows:

1. "universal oligonucleotide tail adaptor" means "a nucleic acid molecule comprised of two strands (a blocking strand and an amplification strand) and comprising a first ligatable duplex end and a second unpaired end" ('810 Patent, claims 1, 15, & 16);

2. "blocking strand" means "a strand of the universal oligonucleotide tail adaptor that comprises a 5' duplex portion" ('810 Patent, claims 1-4 & 16);

3. "amplification strand" means "a strand of the universal oligonucleotide tail adaptor that comprises an unpaired 5' portion, a 3' duplex portion, and a 3' T overhang, and nucleic acid sequences identical to a first and second sequencing primers" ('810 Patent, claims 1-4 & 16);

[1] The construction of all the listed terms except the term "nested" were announced at the November 12, 2019 hearing.

Appx1

4. "second target-specific primer" means "a single-stranded oligonucleotide comprising a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (b), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer" ('810 Patent, claims 1 & 16);

5. "known target nucleotide sequence" means "a portion of a target nucleic acid for which the sequence (e.g. the identity and order of the nucleotide bases comprising the nucleic acid) is known" ('810 Patent, claims 1, 13, 14, & 16); and

6. "second adaptor primer" means "a nucleic acid molecule comprising a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first adaptor primer" ('810 Patent, claims 1 & 16).

In addition, for the reasons set forth below,

7. "nested" means "annealed to a nucleic acid sequence 3' downstream and in the same direction as another molecule . . ." ('810 Patent, claim 1).

The parties briefed the issues (*see* D.I. 123) and submitted a Joint Claim Construction Chart containing intrinsic evidence (*see* D.I. 134, Ex. A[2]). They also submitted an extensive Appendix that included three expert declarations. (D.I. 124). Defendants submitted a tutorial describing the relevant technology (*see* D.I. 140). The Court carefully reviewed all submissions in connection with the parties' contentions regarding the disputed claim terms, heard oral argument and expert testimony (*see* D.I. 143), and applied the following legal standards in reaching its decision.

I. LEGAL STANDARDS

A. Claim Construction

"[T]he ultimate question of the proper construction of the patent [is] a question of law," although subsidiary fact-finding is sometimes necessary. *Teva Pharms. USA, Inc. v. Sandoz, Inc.*,

[2] The parties amended their original claim construction chart (D.I. 76) three times. (D.I. 90); (D.I. 93); (D.I. 134, Ex. A). The Court references the third, and final, amended claim construction chart.

Appx2

Tecan Ex. 2207, Page 84 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

135 S. Ct. 831, 837-38 (2015). "[T]he words of a claim are generally given their ordinary and customary meaning [which is] the meaning that the term would have to a person of ordinary skill in the art in question at the time of the invention, i.e., as of the effective filing date of the patent application." *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312-13 (Fed. Cir. 2005) (en banc) (internal citations and quotation marks omitted). Although "the claims themselves provide substantial guidance as to the meaning of particular claim terms," the context of the surrounding words of the claim also must be considered. *Id.* at 1314. "[T]he ordinary meaning of a claim term is its meaning to the ordinary artisan after reading the entire patent." *Id.* at 1321 (internal quotation marks omitted).

The patent specification "is always highly relevant to the claim construction analysis . . . [as] it is the single best guide to the meaning of a disputed term." *Vitronics Corp. v. Conceptronic, Inc.*, 90 F.3d 1576, 1582 (Fed. Cir. 1996). It is also possible that "the specification may reveal a special definition given to a claim term by the patentee that differs from the meaning it would otherwise possess. In such cases, the inventor's lexicography governs." *Phillips*, 415 F.3d at 1316. "Even when the specification describes only a single embodiment, [however,] the claims of the patent will not be read restrictively unless the patentee has demonstrated a clear intention to limit the claim scope using words or expressions of manifest exclusion or restriction." *Hill-Rom Servs., Inc. v. Stryker Corp.*, 755 F.3d 1367, 1372 (Fed. Cir. 2014) (internal quotation marks omitted) (quoting *Liebel-Flarsheim Co. v. Medrad, Inc.*, 358 F.3d 898, 906 (Fed. Cir. 2004)).

In addition to the specification, a court "should also consider the patent's prosecution history, if it is in evidence." *Markman v. Westview Instruments, Inc.*, 52 F.3d 967, 980 (Fed. Cir. 1995) (en banc), *aff'd*, 517 U.S. 370 (1996). The prosecution history, which is "intrinsic evidence, . . . consists of the complete record of the proceedings before the [Patent and Trademark Office]

Tecan Ex. 2207, Page 85 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

and includes the prior art cited during the examination of the patent." *Phillips*, 415 F.3d at 1317. "[T]he prosecution history can often inform the meaning of the claim language by demonstrating how the inventor understood the invention and whether the inventor limited the invention in the course of prosecution, making the claim scope narrower than it would otherwise be." *Id.*

In some cases, courts "will need to look beyond the patent's intrinsic evidence and to consult extrinsic evidence in order to understand, for example, the background science or the meaning of a term in the relevant art during the relevant time period." *Teva*, 135 S. Ct. at 841. Extrinsic evidence "consists of all evidence external to the patent and prosecution history, including expert and inventor testimony, dictionaries, and learned treatises." *Markman*, 52 F.3d at 980. Expert testimony can be useful "to ensure that the court's understanding of the technical aspects of the patent is consistent with that of a person of skill in the art, or to establish that a particular term in the patent or the prior art has a particular meaning in the pertinent field." *Phillips*, 415 F.3d at 1318. Nonetheless, courts must not lose sight of the fact that "expert reports and testimony [are] generated at the time of and for the purpose of litigation and thus can suffer from bias that is not present in intrinsic evidence." *Id.* Overall, although extrinsic evidence "may be useful to the court," it is "less reliable" than intrinsic evidence, and its consideration "is unlikely to result in a reliable interpretation of patent claim scope unless considered in the context of the intrinsic evidence." *Id.* at 1318-19. Where the intrinsic record unambiguously describes the scope of the patented invention, reliance on any extrinsic evidence is improper. *See Pitney Bowes, Inc. v. Hewlett-Packard Co.*, 182 F.3d 1298, 1308 (Fed. Cir. 1999) (citing *Vitronics*, 90 F.3d at 1583).

B. Indefiniteness

"The primary purpose of the definiteness requirement is to ensure that the claims are written in such a way that they give notice to the public of the extent of the legal protection afforded

Tecan Ex. 2207, Page 86 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

by the patent, so that interested members of the public, *e.g.* competitors of the patent owner, can determine whether or not they infringe." *All Dental Prodx, LLC v. Advantage Dental Prods., Inc.*, 309 F.3d 774, 779-80 (Fed. Cir. 2002) (citing *Warner-Jenkinson Co. v. Hilton-Davis Chem. Co.*, 520 U.S. 17, 28-29 (1997)). Put another way, "[a] patent holder should know what he owns, and the public should know what he does not." *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., Ltd.*, 535 U.S. 722, 731 (2002).

A patent claim is indefinite if, "viewed in light of the specification and prosecution history, [it fails to] inform those skilled in the art about the scope of the invention with reasonable certainty." *Nautilus, Inc. v. Biosig Instruments, Inc.*, 134 S. Ct. 2120, 2129 (2014). A claim may be indefinite if the patent does not convey with reasonable certainty how to measure a claimed feature. *See Teva Pharm. USA, Inc. v. Sandoz, Inc.*, 789 F.3d 1335, 1341 (Fed. Cir. 2015). But "[i]f such an understanding of how to measure the claimed [feature] was within the scope of knowledge possessed by one of ordinary skill in the art, there is no requirement for the specification to identify a particular measurement technique." *Ethicon Endo–Surgery, Inc. v. Covidien, Inc.*, 796 F.3d 1312, 1319 (Fed. Cir. 2015).

Like claim construction, definiteness is a question of law, but the Court must sometimes render factual findings based on extrinsic evidence to resolve the ultimate issue of definiteness. *See, e.g.*, *Sonix Tech. Co. v. Publications Int'l, Ltd.*, 844 F.3d 1370, 1376 (Fed. Cir. 2017); *see also Teva*, 135 S. Ct. at 842-43. "Any fact critical to a holding on indefiniteness . . . must be proven by the challenger by clear and convincing evidence." *Intel Corp. v. VIA Techs., Inc.*, 319 F.3d 1357, 1366 (Fed. Cir. 2003); *see also Tech. Licensing Corp. v. Videotek, Inc.*, 545 F.3d 1316, 1338 (Fed. Cir. 2008).

I. THE COURT'S RULING

The Court's rulings regarding all of the disputed claim terms of the '810 Patent, except "nested," were announced from the bench at the conclusion of the hearing as follows:

> At issue is U.S. Patent No. 10,017,810, titled "Methods for Determining a Nucleotide Sequence Contiguous to a Known Target Nucleotide Sequence." There are seven terms in dispute. All of the terms but one – "nested" – appear in claims 1 and 16 of the '810 Patent. "Nested" appears only in claim 1.
>
> I am prepared to rule on six of those disputes. I will not be issuing a written opinion on those six, but I will issue an order stating my rulings. I will issue an opinion – either with that order or otherwise – regarding the term "nested."
>
> As to my rulings today, I want to emphasize before I announce my decisions that while I am not issuing a written opinion, we have followed a full and thorough process before making the decisions I am about to state. I have reviewed the '810 Patent and the portions of the prosecution history and parent prosecution history submitted. There was full briefing on each of the disputed terms. There was an extensive appendix that included two expert declarations from Niall Lennon on behalf of Plaintiffs and one from Michael Metzker on behalf of Defendants. Defendants submitted a tutorial. There has been argument here today, as well as testimony from Dr. Metzker and from Dr. Lennon. All of that has been carefully considered.
>
> Now as to my rulings. As an initial matter, I am not going to read into the record my understanding of claim construction law generally. I have a legal standard section that I have included in earlier opinions, including somewhat recently in *Omega Flex v. Ward Manufacturing*, C.A. No. 18-1004. I incorporate that law and adopt it into my ruling today, and will also set it out in the order that I issue.
>
> Additionally, I note that through Dr. Lennon, Plaintiffs define the person of ordinary skill in the art as someone having "a Ph.D. in Biology, Molecular Biology, Genetics, Biochemistry, Pharmacology, Biomedical Engineering or related fields, with 2 to 5 years of experience in next generation sequencing OR a BSc in Biology, Molecular Biology, Genetics, Biochemistry, Pharmacology, Biomedical Engineering, or related fields and 5 to 6 years of experience in next generation sequencing."

Defendants, through Dr. Metzker, define the person of ordinary skill as one having "either an M.D./Ph.D. or Ph.D. in Molecular Biology, Molecular Genetics, Chemistry, Engineering, or equivalent disciplines with two years of experience, or a Bachelor of Science in such fields with five years of experience, with such experience including library preparation methods, PCR, and next generation sequencing." The competing definitions differ as to whether the experience should be limited to next gen sequencing, but the parties appear to agree that the differences in the proposals are not relevant to claim construction.

First, we have three terms: "Universal oligonucleotide tail adaptor," "blocking strand," and "amplification strand," all of which appear in subsection little roman numeral "i" of claims 1 and 16. Plaintiffs propose that no construction is necessary for any of the three. In their briefing, in the alternative, for "universal oligonucleotide tail adaptor," Plaintiffs offer "[a] nucleic acid molecule comprised of two strands (a blocking strand and an amplification strand) and comprising a first ligatable duplex end and a second unpaired end." Today they offered a construction of amplification strand that "amplification strand comprises an unpaired 5' portion, a 3' duplex portion, and a 3' T overhang and nucleic acid sequences identical to a first and second sequencing primer."

For universal oligonucleotide tail adaptor, Defendants propose "[a] nucleic acid molecule comprised of two strands (a blocking strand and an amplification strand) and comprising a first ligatable duplex end and a second unpaired end. The blocking strand of the universal oligonucleotide tail adaptor comprises a 5' duplex portion. The amplification strand comprises an unpaired 5' portion, a 3' duplex portion, and a 3' T overhang and nucleic acid sequences identical to a first and second sequencing primers. The duplex portions of the blocking strand and the amplification strand are substantially complementary and form the first ligatable duplex end comprising a 3' T overhang and the duplex portion is of sufficient length to remain in duplex form at the ligation temperature." For the other two terms – "blocking strand" and "amplification strand" – they assert that these terms are "defined as part of the definition of 'universal oligonucleotide tail adaptor.'"

In column 11, lines 32 to 36 of the '810 Patent, the specification states "[a]s used herein, the term 'universal oligonucleotide tail adaptor' refers to a nucleic acid molecule comprised of two strands (a blocking strand and an amplification strand) and comprising a first ligatable duplex end and a second

unpaired end." It then continues [in lines 36 to 40] that the blocking strand "comprises a 5' duplex portion" and the "amplification strand comprises an unpaired 5' portion, a 3' duplex portion, and a 3' T overhang and nucleic acid sequences identical to a first and second sequencing primer."

Here, the first sentence of the paragraph provides a clear definition and lexicography for the term. The parties do not dispute that. Moreover, here today the parties agreed that two sentences discussing the "blocking strand" and "amplification strand" are acceptable definitions for those terms, and also agreed that whether those terms are included as part of the "universal oligonucleotide tail adaptor" or separately does not matter.

The dispute that remains is whether it is appropriate to add the last sentence of Defendants' proposal – that "[t]he duplex portions of the blocking strand and the amplification strand are substantially complementary and form the first ligatable duplex end comprising a 3' T overhang and the duplex portion is of sufficient length to remain in duplex form at the ligation temperature."

Here, the prosecution history is instructive. In the Notice of Allowability of the application that became the ['810's parent patent, U.S. Patent Number 9,487,828], the Examiner referred to the universal oligonucleotide tail as being one of the terms having an "explicit and limiting definition[] present in the specification." She then went on to say that "[t]he universal oligonucleotide tail adaptor must contain a blocking strand and an amplification strand as well as a ligatable duplex portion and an unpaired end [(Specification at page 14, paragraph 65)]. The blocking strand must contain a 5' duplex portion, and the amplification strand must contain an unpaired 5' portion, a 3' duplex portion, and a 3' T overhang, and sequences identical to the first and second sequencing primers [(Specification at page 14, paragraph 65)]."

Similarly, in the June 7, 2017 Office Action during the prosecution of the '810 Patent, the Examiner rejected several dependent claims as not further limiting because they recited structural features "required by the explicit definition of the 'universal tail adaptor' set forth in the specification." [June 7, 2017 Office Action at page 7]. Among the claims that were rejected on that basis were then claim 23, which claimed that the amplification strand must contain an unpaired 5' portion, and also then claim 24, which specified a 3' T overhang.

The Examiner, however, did not include the additional requirements that Defendants seek to add, i.e., that the strands be substantially complimentary and of sufficient length. In looking at the intrinsic evidence as a whole, I find that those requirements are not clearly part of the patentees' definition.

Thus, I will construe the term "universal [oligonucleotide] tail adaptor" to mean: "a nucleic acid molecule comprised of two strands (a blocking strand and an amplification strand) and comprising a first ligatable duplex end and a second unpaired end."

I will construe "blocking strand of the universal oligonucleotide tail adaptor" to comprise a 5' duplex portion.

I will construe the amplification strand to comprise an unpaired 5' portion, a 3' duplex portion, and a 3' T overhang, and nucleic acid sequences identical to a first and second sequencing primers.

Second, we have the term "second target-specific primer." Plaintiffs state that no construction is necessary, but "[t]o the extent the Court seeks to construe the term," proposes the construction: "A single-stranded oligonucleotide comprising a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the amplification product resulting from step (ii), and a 5' portion comprising a nucleic acid sequence that is identical to the second sequencing primer."

Defendants counter with "[a] single-stranded oligonucleotide comprising a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon (amplification product) resulting from step (ii), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer. The second target-specific primer is nested with respect to the first target-specific primer."

Here, again, the '810 Patent defines the term at issue, and the parties do not dispute that. In column 16, lines 33 to 39, the patent states: "As used herein, the term 'second target-specific primer' refers to a single-stranded oligonucleotide comprising a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (b), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer."

The parties do not dispute that this is a definition. Instead, Plaintiffs remove part of the definition – the words "the known target nucleotide sequence comprised by" before the word amplicon [–] and change other words, e.g., they change "a second sequencing primer." And Defendants add the sentence in the specification following the definition that I just read, i.e., that "[t]he second target-specific primer is nested with respect to the first target-specific primer."

It is settled that the specification may define a claim term and that, when it does, "the inventor's lexicography governs." *Phillips v. AWH Corp.*, 415 F.3d 1303, 1316 (Fed. Cir. 2005). "To act as its own lexicographer, a patentee must 'clearly set forth a definition of the disputed claim term' other than its plain and ordinary meaning." *Cont'l Circuits LLC v. Intel Corp.*, 915 F.3d 788, 796 (Fed. Cir. 2019)[; *Pacing Techs., LLC v. Garmin Int'l, Inc.*, 778 F.3d 1021, 1024 (Fed. Cir. 2015); *Thorner v. Sony Comput. Entm't Am. LLC*, 669 F.3d 1362, 1365 (Fed. Cir. 2012); *CCS Fitness, Inc. v. Brunswick Corp.*, 288 F.3d 1359, 1366 (Fed. Cir. 2002)].

Here, I reject all of the changes proposed by the parties to the definition set forth in the specification.

Plaintiffs clearly set forth a definition in the specification. And they have not offered any compelling reason to diverge from that. Instead, they change the words asserting that it is not a difference in substance but rather will make it more jury friendly.

I cannot say whether that is true or not at this point, but the intrinsic evidence does not support the changes Plaintiffs suggest. For example, during the prosecution of the parent application, applicants submitted a new claim (claim 56), which claimed *inter alia* "[t]he method of claim 21, wherein the second target-specific primer comprises a 3' portion comprising a nucleic acid that can specifically anneal to a portion of the nucleic acid comprising the known target nucleotide sequence comprised by the amplification product of step (ii)." [June 16, 2015 Amendment at page 8.] The Examiner rejected that dependent claim because it was not further limiting because the added features were already required in claim 21 by the explicit definition in the specification. [September 22, 2015 Office Action at page 8.]

Similarly, in the prosecution of the '810 Patent itself, virtually the same proposed amendment was rejected for the same

reason. [May 2, 2017 Preliminary Amendment at page 3, claim 32; June 7, 2017 Office Action at page 7.]

Additionally, in connection with a double patenting rejection, the Examiner noted that "the second amplification . . . necessarily requires the second target-specific primer to be capable of annealing to a portion of the known target sequence present in the first amplification product." [June 7, 2017 Office Action at pages 10–11].

As to Defendants' proposal, I do not find that it is clear that the language they propose to add is part of the lexicography for the term. That this is not part of the definition is consistent with the prosecution history where the Patent Examiner referenced the "explicit" definition of "second target-specific primer," [in the June 7, 2017 Office Action at page 3,] and she did not include the "nesting requirement" that Defendants now propose.

Similarly, in the Notice of Allowability of the parent application, the Examiner referred to the definition of "second target-specific primer" as requiring "a 3' target-specific portion and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer." [Notice of Allowance at page 8]. Again, there is no requirement for nesting recognized by the Examiner as part of the construction.

Moreover, "nesting" is specifically included in claim 1, but is not included in claim 16. In the absence of a clear reason to do so, which I have not seen, I will not read into the claim a term that the patentee chose not to include.

To the extent that Defendants assert that the reference to a second sequencing primer in the construction renders the claim indefinite, for a claim to be held invalid for indefiniteness, there must be clear and convincing evidence. *See Nautilus, Inc. v. Biosig Instruments, Inc.*, 572 U.S. 898, 912 n.10 [(citing *Microsoft Corp. v. i4i Ltd. P'ship*, 564 U.S. 91, 95 (2011))]. At this time, the Court finds that Defendants have not met their burden to show indefiniteness. Should there still be a disagreement regarding this claim in the future, Defendants may raise the issue later, if appropriate, after full fact and expert discovery.

The third disputed term is "nested." We have heard testimony and argument on this issue today. I need to review that testimony and argument, and as I stated at the beginning of this ruling, I will address that issue separately.

The fourth disputed term is "known target nucleotide sequence." Plaintiffs state that no construction is necessary, but then offer that to the extent the Court seeks to construe the term … "[a] portion of a target nucleic acid for which the sequence (e.g. the identity and order comprising the nucleotide bases comprising the nucleic acid) is known."

Defendants offer the same sentence at the start of their proposal and then tack on the sentence: "The known target nucleotide sequence can be of any length of 10 or more nucleotides."

Here, again the patent specification provides a definition of the term. In column 10, lines 55 to 59, it states: "As used herein, the term 'known target nucleotide sequence' refers to a portion of a target nucleic acid for which the sequence (e.g. the identity and order of the nucleotide bases comprising the nucleic acid) is known."

It is not clear that [Defendants'] additional sentence is part of the lexicography. And in the absence of that being clear, I will not read the limitation in.

The final term is "second adaptor primer." Here, again, Plaintiffs offer that no construction is necessary, and in the alternative propose the construction "[a] nucleic acid molecule comprising a nucleic acid sequence identical to a portion of the first sequencing primer."

Defendants propose the construction "[a] nucleic acid molecule comprising a nucleic acid sequence identical to a portion of the first sequencing primer and that is nested with respect to the first adaptor primer."

The specification again defines [the] term. In column 17, lines 5 to 8, it states: "As used herein, the term 'second adaptor primer' refers to a nucleic acid molecule comprising a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first adaptor primer."

Plaintiffs agree that the first portion of that sentence is definitional but omit the second portion. They support their argument [by] noting that during prosecution of the parent the Examiner in an Interview Summary noted that "new claim 79 does not require nesting of the adaptor primers." As Defendants point out, however, the term at issue in claim 79 was not "second adaptor primer" it was "an adaptor primer."

> Here, the patentees included a nesting requirement in the definition of "second adaptor primer." They did not refer to it as an embodiment. They included it in the definition. I will follow suit and define the term as the patentees did in the specification.

As noted above, the Court did not construe the term "nested" at the hearing. After further review of the papers submitted and the transcript of the argument and expert testimony, it will do so now. Claim 1 of the '810 Patent requires that "the second target-specific primer is relative to the first target-specific primer." Plaintiffs assert that "[a] second primer is nested relative to a first primer when the second primer is not identically positioned with respect to the first primer." (D.I. 134, Ex. A at 11). Defendants assert "[a] second primer is nested with respect to a first primer when the second primer anneals to a nucleic acid sequence 3' downstream and in the same direction as the first primer (inside)." (*Id.*). Thus, the primary dispute is whether the second primer must be downstream of and in the same direction as the first primer or may anneal anywhere that is different from the first primer's binding location.

The Court construes "nested" to mean "annealed to a nucleic acid sequence 3' downstream and in the same direction as another molecule." This construction is consistent with the ordinary and customary meaning of "nested" to a person of ordinary skill in the relevant art and is supported by the intrinsic evidence.

As already stated, "the words of a claim are generally given their ordinary and customary meaning [which is] the meaning that the term would have to a person of ordinary skill in the art in question at the time of the invention, i.e., as of the effective filing date of the patent application." *Phillips*, 415 F.3d at 1312-13. The Federal Circuit has further instructed that "[b]roadening of ordinary meaning of a term" to "encompass relatively obscure definitions" "in the absence of support in the intrinsic record indicating that such a broad meaning was intended violates the

Tecan Ex. 2207, Page 95 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

principles articulated in *Phillips*." *Nystrom v. Trex Co., Inc.*, 424 F.3d 1136, 1145-46 (Fed. Cir. 2005).

Both experts agree that, by far, the "most common" way in which "nested" is used in the relevant art is to describe a second primer 3' downstream from and in the same direction as a first primer. (*See, e.g.*, D.I. 143 at 85:6-20, 95:15-22; D.I. 124 Ex. 2 ¶ 42 (Metzker Declaration)). Plaintiffs, however, maintain that the "ordinary" meaning is broader. Setting aside any semantic argument regarding the difference between "most common" and "ordinary," in light of *Nystrom* the Court is hesitant to adopt Plaintiffs' construction without intrinsic support for such an expansion of the vastly more common definition. Yet Plaintiffs rely entirely on extrinsic evidence – the use of "nested" in three publications[3] and the personal experiences of their expert, Dr. Lennon. (*E.g.*, D.I. 123 at 47-49, 55-57).[4] Moreover, the intrinsic record here does not support Plaintiffs' construction; rather, it supports an "inside" construction of nested.

---

[3] Defendants assert, and Plaintiffs do not dispute, that the three sources Plaintiffs cite are extrinsic. (*E.g.*, D.I. 123 at 55 & n.16 (Defendants' first articulation of "extrinsic" argument in Joint Claim Construction Brief); *id.* at 55-57 (Archer's Response)); D.I. 143 at 109:11-17, 121:6-7 (Defendants repeatedly describing Plaintiffs' references as "extrinsic")). Two of them – "Schuelke" and "Kanizay," (*e.g.,* D.I. 124 Ex. 1, Exs. 3-4) – indisputably are, as they are entirely external to the patent and prosecution history. The third – the "Illumina" or "Solexa" reference, (*e.g.*, D.I. 124 Ex. 1, Ex. 2), however, is cited on the face of the '810 Patent. ('810 Patent, at [56]). Although "prior art cited in a patent or cited in the prosecution history of the patent constitutes intrinsic evidence," *V-Formation v. Benetton Grp. S.P.A.*, 401 F.3d 1307, 1311 (Fed. Cir. 2005) (citations omitted), the Federal Circuit has also held that references included by a patentee in its disclosures but not otherwise cited by the applicant or the examiner during prosecution have little probative value in claim construction. *Osram GmbH v. Int'l Trade Comm'n*, 505 F.3d 1351, 1357-58 (Fed. Cir. 2007) (citations omitted). The dissenting judge in *Osram* even described such uncited references as "extrinsic evidence." *Id.* at 1361. Given this, along with Plaintiffs' failure to dispute Defendants' "extrinsic" argument or to articulate how the "Illumina/Solexa" patent application was utilized in the prosecution of the '810 Patent, the Court considers that reference to be extrinsic evidence.

[4] The parties dispute the import and disclosures of Plaintiffs' cited publications; however, the Court finds no need to resolve those issues in light of its other findings.

First, the articulated goals and advantages of the claimed invention require "internally" nested primers. The '810 Patent is directed towards increasing the specificity with which target oligonucleotide sequences are identified, ('810 Patent, at [57] (Abstract) ("The technology described herein is directed to methods of determining oligonucleotide sequences, e.g. by enriching target sequences prior to sequencing the sequences."); *id.* at col. 35 l. 60 – col. 36 l. 12 ("Advantages of the methods described herein include: . . . 3. High specificity . . . .")), and a second round of PCR with primers "internal" to those used in the first round is key to achieving that high specificity, (*see* '810 Patent col. 35 ll. 6-9 ("High specificity in this system is achieved with the unidirectional tandem, nested primers . . . ."); *id.* at col. 36 ll. 28-32 ("Additional specificity is achieved by using two unidirectional primers, with [the second] 3' downstream of [the first] . . . .")). Moreover, the Court finds compelling Dr. Metzker's testimony that the use of "externally" nested primers (one option under Plaintiffs' construction) would not increase specificity if utilized in a PCR reaction because the second, external primers would simply "amplify everything that's present" rather than target the desired sequence. (D.I. 143 at 117:18 – 119:23).

Second, the definitions and descriptions of "nested" that are incorporated into the '810 Patent define or describe that term to mean "3' downstream and in the same direction as." *See, e.g.*, Molecular Biology and Biotechnology: A Comprehensive Desk Reference 644 (Meyers et al eds., 1995) ("Another approach that can improve PCR specificity is to follow the initial amplification reaction with an additional PCR with internal, single, or double-nested primers. Like the use of oligonucleotide hybridization probes, this approach utilizes sequence information internal to the two 'outer' primers to identify the subset of amplification products that corresponds to the target fragment."); Sambrook & Russell, Molecular Cloning: A Laboratory Manual 8.61

(3d ed. 2001) ("If necessary, the products of the first PCR can be used as templates for a second, nested PCR, which is primed by a gene-specific sense oligonucleotide internal to the first, and a second antisense oligonucleotide . . . ."); Davis, Kuehl, & Battey, Basic Methods in Molecular Biology 117 (2d ed. 1995) ("'[N]ested' PCR can be used to increase the specificity of the amplification reaction. The nested PCR procedure involves two sequential PCR reactions. After 10 to 20 cycles of PCR with the first pair of gene-specific primers . . . a second 25- to 30-cycle PCR reaction [is conducted] with a new set of gene-specific primers that hybridize to sequences internal to, or nested between, the first pair."); (*see also* '810 Patent col. 9 l. 65 – col. 10 l. 19 (incorporating definitions of common cell and molecular biology terms or descriptions of "standard procedures" from, *inter alia*, Meyers, Sambrook, and Davis)). Additionally, Dr. Metzker testified that at least two of these incorporated resources would be used by someone of ordinary skill in the relevant art at the time of the invention. (D.I. 143 at 129:12 – 132:22).

Finally, although not definitive, Plaintiffs have failed to point to a single instance in the patent or elsewhere in the intrinsic evidence where "nested" is used in the manner they advocate. Rather, as they admit, the term is repeatedly used in the specification to mean "annealed to a nucleic acid sequence 3' downstream and in the same direction as another molecule." (*See* D.I. 123 at 49 (disputing the relevance of, but noting that "the examples of the '810 patent show the second primers nested downstream of the first")); *see also, e.g.*, '810 Patent fig. 1 & col. 5 ll. 56-59 (instructing that, in Figure 1, after being subjected to a first round of PCR with a first target-specific primer, "the sample is subjected to an additional 14 cycles of PCR amplification using multiplex nested gene specific primers (3' downstream of [the first target-specific primer] and in the same direction)"); *id.* at fig. 3A & col. 6 ll. 6-7 (describing Figure 3A, which shows a second target-specific primer 3' downstream and in the same direction as a first target-specific primer, as

Tecan Ex. 2207, Page 98 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

"depict[ing] a schematic presentation of nested primer targeting strategy"); *id.* at fig. 8 & col. 6 ll. 22-23 (stating Fig. 8, which shows "nested adaptor primers and nested gene specific primers" 3' downstream and in the same direction as adaptor and gene specific primers used in an earlier round, as "a schematic presentation of the targeting sequencing approach described herein")).

In light of the foregoing, the Court does not adopt the broad construction Plaintiffs request. Instead, it construes "nested" to mean "annealed to a nucleic acid sequence 3' downstream and in the same direction as another molecule."

Maryellen Noreika

The Honorable Maryellen Noreika
United States District Judge

Appx17

Tecan Ex. 2207, Page 99 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

ARCHERDX, INC. and THE GENERAL HOSPITAL CORPORATION d/b/a MASSACHUSETTS GENERAL HOSPITAL,	)	
Plaintiffs,	)	
v.	)	C.A. No. 18-1019 (MN)
QIAGEN SCIENCES, LLC, QIAGEN LLC f/k/a QIAGEN INC., QIAGEN BEVERLY, LLC f/k/a QIAGEN BEVERLY, INC., QIAGEN GAITHERBURG, LLC f/k/a QIAGEN GAITHERSBURG, INC., QIAGEN GMBH, QIAGEN N.V. and JONATHAN ARNOLD,	)	
Defendants.	)	

MEMORANDUM ORDER

At Wilmington this 18th day of June 2020:

IT IS HEREBY ORDERED that the agreed-upon construction of the claim term "the same sequence" in U.S. Patent No. 10,450,597 ("the '597 Patent") is: "the identical sequence," ('597 Patent, cl. 1).

Further, as announced at the hearing on May 28, 2020, IT IS HEREBY ORDERED that the disputed claim terms of the '597 Patent are construed as follows:

1. "double-stranded target nucleic acid" means "a target nucleic acid comprising a first strand and a second strand that are complementary to each other and hybridized to each other but that may have 3' or 5' overhangs or tails," ('597 Patent, cl. 1, 3, 11, 25-27, & 29);

2. "target nucleic acid" means "a nucleic acid molecule of interest (e.g., a nucleic acid to be analyzed)," ('597 Patent, cl. 1, 3, 11, 25-27, & 29);

3. "target-specific primer" means "a primer that has a level of complementarity between the primer and the target such that there exists an annealing temperature at which the primer will anneal to and mediate

amplification of the target nucleic acid and will not anneal to or mediate amplification of non-target sequences present in a sample," ('597 Patent, cl. 1, 14, 19, & 21);

4. "target-specific hybridization sequence" means "a sequence of the target-specific primer that has sufficient complementarity with a sequence of the double-stranded target nucleic acid to enable hybridization between the target-specific primer and a sequence in/of the double-stranded target nucleic acid," ('597 Patent, cl. 1, 13, 14, 19, & 20);

5. "plurality of different primers" means "two or more different primers," ('597 Patent, cl. 1, 5, 24, 25, & 26);

6. "a sequence that is characteristic of the target-specific primer" means "a sequence from one of the strands of the single target nucleic acid recited in step (a)," ('597 Patent, cl. 1);

7. "[contacting a] first nucleic acid template comprising a sequence of a first strand of a double-stranded target nucleic acid [with a complementary target-specific primer that comprises a target-specific hybridization sequence]" needs no construction, ('597 Patent, cl. 1 & 2);

8. "[contacting a] second nucleic acid template comprising a sequence of a second strand that is complementary to the sequence of the first strand of the double-stranded target nucleic acid" needs no construction, ('597 Patent, cl. 1).[1]

The parties briefed the issues (*see* D.I. 226) and submitted a Joint Claim Construction Chart containing intrinsic evidence (*see* D.I. 249, Ex. A[2]). They also submitted an extensive Appendix that included four expert declarations. (D.I. 227.) The Court carefully reviewed all submissions in connection with the parties' contentions regarding the disputed claim terms, heard oral argument (*see* D.I. 251), and applied the following legal standards in reaching its decision.

---

[1] Pursuant to the parties' request, the Court considered the disputed terms in the order they were argued rather than the order in which they were presented in the final amended claim construction chart. (*See* D.I. 249 at 5; D.I. 251.)

[2] The parties amended their original claim construction chart (D.I. 194) three times. (D.I. 225; D.I. 246, Ex. A; D.I. 249, Ex. A). The Court references the third, and final, amended claim construction chart.

Tecan Ex. 2207, Page 101 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

I. LEGAL STANDARDS

"[T]he ultimate question of the proper construction of the patent [is] a question of law," although subsidiary fact-finding is sometimes necessary. *Teva Pharms. USA, Inc. v. Sandoz, Inc.*, 135 S. Ct. 831, 837-38 (2015). "[T]he words of a claim are generally given their ordinary and customary meaning [which is] the meaning that the term would have to a person of ordinary skill in the art in question at the time of the invention, i.e., as of the effective filing date of the patent application." *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312-13 (Fed. Cir. 2005) (en banc) (internal citations and quotation marks omitted). Although "the claims themselves provide substantial guidance as to the meaning of particular claim terms," the context of the surrounding words of the claim also must be considered. *Id.* at 1314. "[T]he ordinary meaning of a claim term is its meaning to the ordinary artisan after reading the entire patent." *Id.* at 1321 (internal quotation marks omitted).

The patent specification "is always highly relevant to the claim construction analysis . . . [as] it is the single best guide to the meaning of a disputed term." *Vitronics Corp. v. Conceptronic, Inc.*, 90 F.3d 1576, 1582 (Fed. Cir. 1996). It is also possible that "the specification may reveal a special definition given to a claim term by the patentee that differs from the meaning it would otherwise possess. In such cases, the inventor's lexicography governs." *Phillips*, 415 F.3d at 1316. "Even when the specification describes only a single embodiment, [however,] the claims of the patent will not be read restrictively unless the patentee has demonstrated a clear intention to limit the claim scope using words or expressions of manifest exclusion or restriction." *Hill-Rom Servs., Inc. v. Stryker Corp.*, 755 F.3d 1367, 1372 (Fed. Cir. 2014) (internal quotation marks omitted) (quoting *Liebel-Flarsheim Co. v. Medrad, Inc.*, 358 F.3d 898, 906 (Fed. Cir. 2004)).

In addition to the specification, a court "should also consider the patent's prosecution history, if it is in evidence." *Markman v. Westview Instruments, Inc.*, 52 F.3d 967, 980 (Fed. Cir.

Tecan Ex. 2207, Page 102 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

1995) (en banc), *aff'd*, 517 U.S. 370 (1996). The prosecution history, which is "intrinsic evidence, . . . consists of the complete record of the proceedings before the [Patent and Trademark Office] and includes the prior art cited during the examination of the patent." *Phillips*, 415 F.3d at 1317. "[T]he prosecution history can often inform the meaning of the claim language by demonstrating how the inventor understood the invention and whether the inventor limited the invention in the course of prosecution, making the claim scope narrower than it would otherwise be." *Id.*

In some cases, courts "will need to look beyond the patent's intrinsic evidence and to consult extrinsic evidence in order to understand, for example, the background science or the meaning of a term in the relevant art during the relevant time period." *Teva*, 135 S. Ct. at 841. Extrinsic evidence "consists of all evidence external to the patent and prosecution history, including expert and inventor testimony, dictionaries, and learned treatises." *Markman*, 52 F.3d at 980. Expert testimony can be useful "to ensure that the court's understanding of the technical aspects of the patent is consistent with that of a person of skill in the art, or to establish that a particular term in the patent or the prior art has a particular meaning in the pertinent field." *Phillips*, 415 F.3d at 1318. Nonetheless, courts must not lose sight of the fact that "expert reports and testimony [are] generated at the time of and for the purpose of litigation and thus can suffer from bias that is not present in intrinsic evidence." *Id.* Overall, although extrinsic evidence "may be useful to the court," it is "less reliable" than intrinsic evidence, and its consideration "is unlikely to result in a reliable interpretation of patent claim scope unless considered in the context of the intrinsic evidence." *Id.* at 1318-19. Where the intrinsic record unambiguously describes the scope of the patented invention, reliance on any extrinsic evidence is improper. *See Pitney Bowes, Inc. v. Hewlett-Packard Co.*, 182 F.3d 1298, 1308 (Fed. Cir. 1999) (citing *Vitronics*, 90 F.3d at 1583).

Tecan Ex. 2207, Page 103 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

II. THE COURT'S RULING

The Court's rulings regarding all of the disputed claim terms of the '597 Patent were announced from the bench at the conclusion of the hearing as follows:

> Good afternoon again, counsel. Thank you for the arguments today. The Court previously construed seven terms in Plaintiffs' U.S. Patent No. 10,017,810. Today we are addressing their U.S. Patent No. 10,450,597, which has been added to the case. There are another eight terms in dispute in that patent.
>
> I am prepared to rule on all eight of those disputes. As with my earlier constructions, I will not be issuing a written opinion on the constructions of these eight terms, but I will issue an order stating my rulings.
>
> Also, I want to emphasize that, as with my prior ruling, although I am not issuing a written opinion, we have followed a full and thorough process before making the decisions I am about to state. I have reviewed the '597 Patent and the portions of the prosecution history submitted. There was full briefing on each of the disputed terms. There was an extensive appendix that included two more expert declarations from Niall Lennon on behalf of Plaintiffs and two more from Michael Metzker on behalf of Defendants. There has been argument here today. All of that has been carefully considered.
>
> I also incorporate the law regarding claim construction set out in my earlier order and will also set it out in the order that I issue on the '597 Patent.
>
> Additionally, I note that as to a person of ordinary skill in the art, the parties appear to agree that any differences in their respective proposals are not relevant to claim construction.[3]

---

[3] Through Dr. Lennon, Plaintiffs define the person of ordinary skill in the art as someone having "a Ph.D. in Biology, Molecular Biology, Genetics, Biochemistry, Pharmacology, Biomedical Engineering or related fields, with 2 to 5 years of experience in next generation sequencing OR a BSc in Biology, Molecular Biology, Genetics, Biochemistry, Pharmacology, Biomedical Engineering, or related fields and 5 to 6 years of experience in next generation sequencing." (*See* D.I. 124-1 ¶ 16 (JA007).) Defendants, through Dr. Metzker, define the person of ordinary skill as one having "either an M.D./Ph.D. or Ph.D. in Molecular Biology, Molecular Genetics, Chemistry, Engineering, or equivalent disciplines with two years of experience, or a Bachelor of Science in such fields with five years of experience, with such experience including library preparation methods, PCR, and next generation sequencing." (*See* D.I. 124-2 ¶ 26 (JA211).)

Now as to my rulings. First, we have the term "double-stranded target nucleic acid" in claims 1, 3, 11, 25 through 27, and 29.

Plaintiffs propose the construction "a target nucleic acid comprising a first strand and a second strand that are complementary to each other and that are sufficiently hybridized to maintain a double-stranded nature." Defendants propose the construction "a target nucleic acid consisting of a first strand and a second strand that are complementary to each other."

The dispute here appears to be over whether some portion of the molecule may be single-stranded, such as a tail, or whether it must be 100 percent double-stranded.

Here, I mostly agree with Plaintiffs and conclude that the term "double-stranded" does not require that a molecule be entirely or 100 percent double-stranded. And I will construe the term to mean "a target nucleic acid comprising a first strand and a second strand that are complementary to each other and hybridized to each other but that may have 3' or 5' overhangs or tails."

This construction is consistent with the ordinary meaning of the term "double-stranded" as used in references cited by the parties and in a number of Qiagen's patents.[4]

Defendants argue that Plaintiffs, in essence, define "double-stranded" in the intrinsic evidence to mean wholly double-stranded. They point to references cited in the patent, which are incorporated by reference into the patent. Although I agree that those references are part of the intrinsic evidence, I disagree that those references provide a clear definition for the term that alters the ordinary meaning. For example, Defendants cite U.S. Patent No. 7,282,337 [("the '337 Patent")] and the Encyclopedia of Molecular Biology. The '337 Patent describes nucleic acid molecules as "single-stranded, double-stranded, or double-stranded with single-stranded regions." [*See* '337 Patent, col. 6 ll. 36-38.] I do not read this to clearly redefine the term "double-stranded," however. Instead, it appears to allow for the term "double-stranded" to be used to encompass molecules that are fully double-stranded as well as those that have some portions that are single-stranded. And nothing in

---

4 *See, e.g.*, The Encyclopedia of Molecular Biology 956, 1036 (John Kendrew et al. eds. 1994) (D.I. 227, Ex. E, Ex. 5); U.S. Patent No. 8,628,914 (D.I. 227, Ex. A, Ex. 2); N.H. Dekker et al., Joining of Long Double-Stranded RNA Molecules Through Controlled Overhangs, 32 Nucleic Acids Research e140, 1, Fig. 1 (2004) (D.I. 227, Ex. A, Ex. 4).

that sentence suggests that a nucleic acid molecule must be 100 percent single-stranded or 100 percent double-stranded, but not both.

Similarly, although the Encyclopedia, in discussing viruses and fill-in reactions, refers to molecules having portions that are single-stranded and portions that are double-stranded, [*see, e.g.*, The Encyclopedia of Molecular Biology 377, 484 (John Kendrew et al. eds. 1994),] it also refers to double-stranded molecules having single-stranded ends. For example, in discussing "sticky ends" it refers to "single-stranded DNA termini protruding from fragments of double-stranded DNA." [*Id.* at 1036.] And in discussing restriction enzymes, it refers to double-stranded DNA fragments with single-stranded extensions at the ends. [*Id.* at 956.] This, too, allows for use of the term "double-stranded" in connection with molecules having single-stranded termini.

The parties also dispute whether the definition of "portions" and Figure 2A in the patent support Plaintiffs' or Defendants' construction. I do not find that either of these parts of the patent clearly define or alter the ordinary meaning of "double-stranded." For example, with Figure 2A, there is some incongruity between the illustration and the description. The illustration suggests that a double-stranded molecule having a single-stranded end is the "double-stranded" molecule purified in Step 205. Yet the written description suggests that the double-stranded illustration after Step 206 in the Figure is representative of the product of Step 204 because it includes the common sequences. [*See* '597 Patent, col. 10 l. 63 – 11 l. 15.]

And finally, Defendants argue that the "double-stranded target nucleic acid" of the claims has a first strand and a second strand that are complementary to each other. And then argue that the word "complementary" in the patent means 100 percent complementary and thus the "double-stranded target nucleic acid" must be 100 percent complementary. I do not, however, read the patent to use the word "complementary" on its own to require 100 percent complementarity. Indeed, the patent, at column 38, line 2, refers to varying levels of complementarity[,] including 100 percent. If the word "complementary" on its own meant 100 percent, then the reference to 100 percent [in that sentence] would be redundant.

The second disputed term is "target nucleic acid," also in claims 1, 3, 11, 25 through 27, and 29. Plaintiffs propose the construction "a nucleic acid molecule of interest (e.g., a nucleic acid to be analyzed)." Defendants' proposal includes all of Plaintiffs'

proposal and then adds to the end: "which is a genomic nucleic acid molecule or cDNA molecule."

Here, I agree with Plaintiffs and will construe the term to mean "a nucleic acid molecule of interest (e.g., a nucleic acid to be analyzed)."

It is settled law that when the specification defines a claim term, "the inventor's lexicography governs." [*Phillips*, 415 F.3d at 1316.] "To act as its own lexicographer, a patentee must 'clearly set forth a definition of the disputed claim term' other than its plain and ordinary meaning." [*Cont'l Circuits LLC v. Intel Corp.*, 915 F.3d 788, 796 (Fed. Cir. 2019); *accord Pacing Techs., LLC v. Garmin Int'l, Inc.*, 778 F.3d 1021, 1024 (Fed. Cir. 2015); *Thorner v. Sony Comput. Entm't Am. LLC*, 669 F.3d 1362, 1365 (Fed. Cir. 2012); *CCS Fitness, Inc. v. Brunswick Corp.*, 288 F.3d 1359, 1366 (Fed. Cir. 2002).]

The patentee has done so here, stating "[a]s used herein, the term 'target nucleic acid' refers to a nucleic acid molecule of interest (e.g., a nucleic acid to be analyzed)." ['597 Patent, col. 15 ll. 30-32.] Neither party disputes that that is the definition.

Instead, Defendants contend that additional limitations must be added to resolve a dispute regarding what it means to be "a nucleic acid molecule of interest" and/or "a nucleic acid to be analyzed" – that is, to resolve their disputes regarding the meaning of certain aspects of the definition provided in the specification.

It is not, however, the Court's job in claim construction to resolve every derivative dispute between the parties regarding the meaning of the terms.[5]

---

5 *See Acumed LLC v. Stryker Corp.*, 483 F.3d 800, 806 (Fed. Cir. 2007) ("[A] sound claim construction need not always purge every shred of ambiguity. The resolution of some line-drawing problems—especially easy ones like this one—is properly left to the trier of fact." (citing *PPG Indus. V. Guardian Indus. Corp.*, 156 F.3d 1351, 1355 (Fed. Cir. 1998) ("[A]fter the court has defined the claim with whatever specificity and precision is warranted by the language of the claim and the evidence bearing on the proper construction, the task of determining whether the construed claim reads on the accused product is for the finder of fact."))); *see also GPNE Corp. v. Apple Inc.*, 830 F.3d 1365, 1372 (Fed. Cir. 2016) ("Where a district court has resolved the questions about claim scope that were raised by the parties, it is under no obligation to address other potential ambiguities that have no bearing on the operative scope of the claim. . . . This is because '[s]uch an endeavor could proceed ad infinitum, as every word—whether a claim term itself, or the words a court uses

> And moreover, here, I am not convinced that the disputes raised by Defendants are those of claim construction rather than factual questions relating to infringement – e.g., whether the particular molecule asserted to be a target nucleic acid is "of interest" or "to be analyzed." Thus, I reject the changes proposed by Defendants to the definition set forth in the specification.
>
> The third disputed term is "target-specific primer" in claims 1, 14, 19, and 21. During the argument today, Plaintiffs agreed that the "target-specific primer" must target the "target nucleic acid" and that Defendants' construction is correct if "target nucleic acid" is not limited to genomic DNA. In my construction of "target nucleic acid," I declined to read in the limitation that the target nucleic acid must be "a genomic nucleic acid molecule or cDNA molecule." Thus, I understand that Plaintiffs agree to the construction advocated by Defendants and I will adopt that.
>
> And to be clear, that construction is "a primer that has a level of complementarity between the primer and the target such that there exists an annealing temperature at which the primer will anneal to and mediate amplification of the target nucleic acid and will not anneal to or mediate amplification of non-target sequences present in a sample."[6]

---

to construe a claim term—is susceptible to further definition, elucidation, and explanation.'" (internal citations omitted)).

[6] The agreed-upon construction is supported by the intrinsic evidence. The relevant claim language says: "A method of preparing nucleic acids for analysis, the method comprising: (a) contacting a first nucleic acid template comprising a sequence of a first strand of a double-stranded target nucleic acid with a complementary target-specific primer that comprises a target-specific hybridization sequence, under conditions to promote template-specific hybridization and extension of the target specific primer. *See* '597 Patent, cl. 1. The only thing referred to as a "target" in the claim language is the "double stranded target nucleic acid." The fact that the nucleic acid template is "contacted" by the target-specific primer does not alter that, as the "double stranded target nucleic acid" is included within the nucleic acid template. The term at issue is "target-specific primer," not "template-specific primer" or "contact primer." Thus, based on the language of the claim itself, the Court is inclined to find that the "target" of the "target-specific primer" is the "target nucleic acid." Additionally, the prosecution history indicates that the applicants intended for the "target-specific primer" to "target" the "target nucleic acid" recited in the claims. During prosecution, the applicants stated that "[t]he present claims . . . relate to two extension reactions to generate an extension product containing 'a sequence that is characteristic of the target-specific primer' – that is, an extension product encoding a sequence from the target nucleic acid recited in step (a)." (*See* D.I. 227, Ex. J, Ex. 5 at ADX-000490591.)

Tecan Ex. 2207, Page 108 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

The fourth disputed term is "target-specific hybridization sequence" in claims 1, 13, 14, 19, and 20. Plaintiffs offer the construction "a sequence of the target-specific primer, that has sufficient complementarity with a sequence of another nucleic acid (e.g., a template molecule, a target sequence) to enable hybridization between the nucleic acids." Defendants counter with the construction "a sequence of the target-specific primer that that has sufficient complementarity with a sequence of the double-stranded target nucleic acid to enable hybridization between the target-specific primer and a sequence in/of the double-stranded target nucleic acid."

The dispute here is essentially the same as the dispute over "target-specific primer" – whether the hybridization sequence must be specific to the target nucleic acid or may enable hybridization to any part of the nucleic acid template.

In the papers, Plaintiffs propose that I adopt the specification's definition of "hybridization sequence." Yet that definition is general and does not account for the "target-specific" modifier in the term at issue. Defendants' proposal, on the other hand, modifies the specification's "hybridization sequence" definition to match its argument that the "target-specific primer" must target the "double-stranded target nucleic acid."

As I already stated and the parties agreed, the target of the "target-specific primer" is the "double-stranded target nucleic acid." As the "target-specific hybridization sequence" is a part of the "target-specific primer," its target must be the same.

Thus, I will construe "target-specific hybridization sequence" to mean "a sequence of the target-specific primer that has sufficient complementarity with a sequence of the double-stranded target nucleic acid to enable hybridization between the target-specific primer and a sequence in/of the double-stranded target nucleic acid."[7]

The fifth disputed term is "plurality of different primers" in claims 1, 5, 24, 25, and 26. Plaintiffs propose the construction "two or more different primers." Defendants propose: "two or more different primers that are not target-specific primers." The dispute

---

7 The Court intentionally uses "that has" rather than "that that has" in its construction. This, however, is solely a grammatical modification. Although "that that has" appears in the specification's definition of "hybridization sequence," the repetition of "that" appears to be a typographical error.

is over Defendants' addition of the language "that are not target-specific primers."

Here, I agree with Plaintiffs and will construe the term to mean "two or more different primers."

This construction is supported by the intrinsic evidence, which allows for the "different primers" to be target-specific primers. For example, at column 8, lines 60 through 64, the patent states: "Subsequent rounds of PCR or extension may be performed using different primers (e.g., second tail primers or nested primers or a second target-specific primer) to incorporate different sequences into the resulting products." Similarly, at column 30, lines 23 through 26, the patent states: "In some embodiments, target nucleic acids in a sample, or separate portions of a sample, can be amplified with a plurality of primers (e.g., a plurality of first and second target-specific primers)."

Defendants assert that the specification distinguishes the claimed method from prior art sequencing in which primers were designed based on known sequences at each end of the target molecule. [*See* '597 Patent, col. 6 ll. 26-30.] And thus, they argue, I should except target-specific primers because if the "plurality of different primers" are designed based on known sequences, as the claimed "target-specific primer" is, then the invention is covered by the prior art. Although it is possible that Plaintiffs' construction will allow for more arguments based on the prior art, there is no argument that construing the term as I am will necessarily leave the claims invalid. And I reject Defendants' arguments.

The sixth disputed term is "a sequence that is characteristic of the target-specific primer" in claim 1. During the argument, the parties agreed to the construction "a sequence from one of the strands of the single target nucleic acid recited in step (a)." I will adopt that construction.

Lastly, we have two terms – (1) "contacting a first nucleic acid template comprising a sequence of a first strand of a double-stranded target nucleic acid with a complementary target-specific primer that comprises a target-specific hybridization sequence" as used in claims 1 and 5; and (2) "contacting a second nucleic acid template comprising a sequence of a second strand that is complementary to the sequence of the first strand of the double-stranded target nucleic acid," also as used in claim 1. These terms comprise the bulk of steps (a) and (b), respectively, described in claim 1.

Plaintiffs propose that no construction is necessary for these terms and the Court should leave them as-is. Defendants, however, want me to construe the first of these terms as "contacting a set of nucleotides in/of a first strand of a double-stranded target nucleic acid that make up a first nucleic acid template with a complementary target-specific primer that comprises a target-specific hybridization sequence," and ask that I construe the second to mean "contacting a set of nucleotides in/of a second strand of the double-stranded target nucleic acid that is complementary to the first strand that make up a second nucleic acid template."

I will note first that the parties agree that steps (a) and (b) may be interchangeable and sequential. That is not the dispute. [*See also* '597 Patent, col. 6 l. 64 – col. 7 l. 8 ("In some embodiments, steps (a) and (b) above are performed sequentially.").]

Instead, the dispute seems to be Defendants' request that I include language that requires the "nucleic acid template" claimed to be made up only of the "target nucleic acid." Defendants' proposed limitations, however, are contradicted by the plain language of the claims, which say that the nucleic acid template comprises a sequence of a first strand of the double-stranded target nucleic acid and not that it is limited to only that sequence.

"Comprising" is a term of art that denotes inclusion but not limitation.[8] And that is also the way "comprising" is defined in the patent, starting at column 38, line 66, where it says: "As used herein the term 'comprising' or 'comprises' is used in reference to compositions, methods, and respected component(s) thereof, that are essential to method or composition, yet open to the inclusion of unspecified elements, whether essential or not."

Defendants' proposed additions are also inconsistent with claim 5, which specifies that "the first nucleic acid template comprises an extension product resulting from the hybridization and extension of the at least one of the plurality of different primers in step (b)."

Thus, I see no reason to limit the first and second strands as proposed by Defendants, and I agree with Plaintiffs that no further construction of the terms are necessary.

---

8 *Waters v. Agilent*, 410 F. Supp. 3d 702, 709-20 (D. Del. Sep. 20, 2019) (citing *Mars, Inc. v. H.J. Heinz Co.*, 377 F.3d 1369, 1376 (Fed. Cir. 2004)).

The Honorable Maryellen Noreika
United States District Judge

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

ARCHERDX, LLC and THE GENERAL HOSPITAL CORPORATION d/b/a MASSACHUSETTS GENERAL HOSPITAL,	)	
Plaintiffs,	)	
v.	)	C.A. No. 18-1019 (MN)
QIAGEN SCIENCES, LLC, QIAGEN LLC f/k/a QIAGEN, INC., QIAGEN BEVERLY, LLC F/K/A QIAGEN BEVERLY, INC., QIAGEN GAITHERSBURG, LLC f/k/a QIAGEN GAITHERSBURG, INC., QIAGEN GMBH, QIAGEN N.V. and JONATHAN ARNOLD,	)	
Defendants.	)	

MEMORANDUM OPINION

Daniel M. Silver, Alexandra M. Joyce, MCCARTER & ENGLISH, LLP, Wilmington, DE; Leigh J. Martinson, Keith Toms, Jill Mello, Ph.D., Wyley S. Proctor, MCCARTER & ENGLISH, LLP, Boston, MA; Michael A. Albert, Eric J. Rutt, WOLF, GREENFIELD & SACKS, P.C., Boston, MA – Attorneys for Plaintiffs

Edward R. Reines, WEIL, GOTSHAL & MANGES LLP, Redwood Shores, CA – Attorney for Plaintiff ArcherDX, LLC

David E. Moore, Bindu A. Palapura, POTTER ANDERSON & CORROON LLP, Wilmington, DE; David Bilsker, Andrew Naravage, QUINN EMANUEL URQUHART & SULLIVAN LLP, San Francisco, CA; Anne Toker, James E. Baker, Anastasia M. Fernands, QUINN EMANUEL URQUHART & SULLIVAN LLP, New York, NY; Jeffrey C. Wu, QUINN EMANUEL URQUHART & SULLIVAN LLP, Salt Lake City, UT – Attorneys for Defendants

September 30, 2022
Wilmington, Delaware

NOREIKA, U.S. DISTRICT JUDGE

The Court presided over a five-day jury trial from August 23, 2021 to August 27, 2021. (*See* D.I. 485, 486, 487, 488, 489). At the end, the jury found Defendants Qiagen Sciences, LLC, Qiagen LLC f/k/a Qiagen, Inc., Qiagen Beverly, LLC f/k/a Qiagen Beverly, Inc., Qiagen Gaithersburg, LLC, f/k/a Qiagen Gaithersburg, Inc., Qiagen GmbH, Qiagen N.V. and Jonathan Arnold (collectively "Defendants" or "Qiagen") to have willfully infringed claims of two patents of Plaintiffs ArcherDX, LLC and the General Hospital Corporation (collectively, "Plaintiffs"). (*See* D.I. 465). Presently before the Court are Defendants' renewed motion for judgment as a matter of law or, alternatively, for a new trial on infringement, invalidity and damages, motion for a new trial on willfulness and motion for remittitur (*See* D.I. 495). In addition, Plaintiffs move for injunctive relief, ongoing royalty, enhanced damages, supplemental damages and pre- and post-judgment interest. (*See* D.I. 494). For the reasons set forth below, the Court will deny Defendants' motions for judgment as a matter of law and new trial, grant Defendants' motion for remittitur, deny Plaintiffs' motion for enhanced damages, deny Plaintiffs' motion for injunction with leave to renew after an evidentiary hearing, grant-in-part Plaintiffs' motion for ongoing royalty and grant Plaintiffs' motions for supplemental damages and pre- and post-judgment interest.

I. BACKGROUND

Plaintiffs and Defendants are in the business of biomedical technology. Plaintiffs and Defendants provide products referred to as assay "kits" that allow users to prepare and analyze nucleic acids using next generation sequencing technology. (*See* D.I. 1 ¶¶ 19-20, D.I. 330 at 1). The kits enable users to detect gene mutations associated with various cancers and have applications in both the clinical and research space. (*See* D.I. 1 ¶ 1, D.I. 130 ¶ 19, D.I. 330 at 1).

At issue in this case are two of Plaintiffs' patents: U.S. Patent No. 10,017,810 ("the '810 patent") and U.S. Patent No. 10,450,597 ("the '597 patent"). The '810 patent is directed to

methods of determining oligonucleotide sequences. (*See* D.I. 130 ¶ 18). The '597 patent is directed to methods of preparing and analyzing nucleic acids. (*See id.*). Plaintiffs have developed and now sell various assay kits that use technology covered by both patents. (*Id.* ¶¶ 19-20). Defendants also developed and now sell kits used to detect gene mutations. (*See* D.I. 330 at 1).

On July 10, 2018, Plaintiffs filed suit alleging that Defendants infringed the '810 patent.[1] (*See* D.I. 1 ¶¶ 41-55). On October 30, 2019, after the '597 patent had issued, Plaintiffs filed an amended complaint, adding a claim for infringement of the '597 patent. (*See* D.I. 130 ¶¶ 41-55). On August 21, 2021 the Court reversed its previous denial of summary judgment of no literal infringement of the '810 patent after further argument. (*See* D.I. 447). Thus, the issues left for trial were literal infringement of the '597 patent and infringement under the doctrine of equivalents for the '810 patent.

From August 23, 2021 to August 27, 2021, the Court presided over a jury trial. (*See* D.I. 485, 486, 487, 488, 489). The jury found that Defendants willfully infringed claims 16, 17 and 19 of the '810 patent and claims 1, 5 and 19 of the '597 patent. (*See* D.I. 465 at 2-4). In addition, the jury found none of these claims invalid. (*See id.* at 5-7). The jury awarded Plaintiff $841,756 in lost profits damages for sales of RNA-related products in the United States, $1,593,762 in royalty damages for sales of DNA-related products in the United States and $2,240,303 in royalty damages for sales outside the United States. (*See id.* at 8).

[1] Plaintiff also initially asserted claims of misappropriation of trade secrets, false advertising, breach of fiduciary duty, deceptive trade practices and tortious conduct. (*See* D.I. 1 at ¶ 1). Prior to trial, Plaintiffs elected to withdraw their claims of trade secret misappropriation, false advertising, deceptive trade practices and tortious interference. (*See* D.I. 413). On August 3, 2021, Plaintiffs accepted an offer of judgment from Defendants pursuant to Rule 68 of the Federal Rules of Civil Procedure for their breach of fiduciary duty claims. (*See* D.I. 423).

Tecan Ex. 2207, Page 115 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

On September 20, 2021, the Court entered judgment on the jury verdict under Rule 58(b) of the Federal Rules of Civil Procedure. (*See* D.I. 482). On October 18, 2021, Defendants renewed their motion for judgment as a matter of law on the issues of infringement, invalidity and damages, or, in the alternative, moved for a new trial. (*See* D.I. 495). In addition, Defendants moved for a new trial on willfulness and remittitur of damages. (*See id.*). On October 18, 2021, Plaintiffs moved for enhanced damages, injunction, ongoing royalty, supplemental damages and pre- and post-judgment interest. (*See* D.I. 494). The parties briefing on post-trial motions was completed on December 15, 2021. (*See* D.I. 513, 514).

II. LEGAL STANDARDS

A. Judgment as a Matter of Law

Judgment as a matter of law may be entered against a non-moving party if the Court "finds that a reasonable jury would not have a legally sufficient evidentiary basis to find for the party on [an] issue." FED. R. CIV. P. 50(a)(1). Judgment as a matter of law is appropriate "only if, viewing the evidence in the light most favorable to the nonmovant and giving it the advantage of every fair and reasonable inference, there is insufficient evidence from which a jury reasonably could find liability." *Lightning Lube, Inc. v. Witco Corp.*, 4 F.3d 1153, 1166 (3d Cir. 1993). Entry of judgment as a matter of law is a remedy to be invoked only "sparingly." *CGB Occupational Therapy, Inc. v. RHA Health Servs. Inc.*, 357 F.3d 375, 383 (3d Cir. 2004).

Following a jury trial, a renewed motion for judgment as a matter of law under Rule 50(b) may be granted only if the movant demonstrates "that the jury's findings, presumed or express, are not supported by substantial evidence or, if they were, that the legal conclusion(s) implied [by] the jury's verdict cannot in law be supported by those findings." *Pannu v. Iolab Corp.*, 155 F.3d 1344, 1348 (Fed. Cir. 1998) (alteration in original) (internal quotation marks omitted). Substantial evidence is such relevant evidence that a reasonable mind might accept as adequate to support the

finding under review. *See Enplas Display Device Corp. v. Seoul Semiconductor Co.*, 909 F.3d 398, 407 (Fed. Cir. 2018). In determining whether substantial evidence supports the jury verdict, the Court may not make credibility determinations, weigh the evidence or substitute its own conclusions for that of the jury where the record evidence supports multiple inferences. *See Lightning Lube*, 4 F.3d at 1166.

B. Motion for a New Trial

"A new trial may be granted to all or any of the parties and on all or part of the issues in an action in which there has been a trial by jury, for any of the reasons for which new trials have heretofore been granted in actions at law in the courts of the United States." FED. R. CIV. P. 59(a). Common reasons for granting a new trial are: (1) the jury's verdict is against the clear weight of the evidence and a new trial is necessary to prevent a miscarriage of justice; (2) there exists newly discovered evidence that would likely alter the outcome of the trial; (3) improper conduct by an attorney or the Court unfairly influenced the verdict; or (4) the jury's verdict was facially inconsistent. *See Ateliers de la Haute-Garonne v. Broe.je Automation-USA Inc.*, 85 F. Supp. 3d 768, 775 (D. Del. 2015).

The decision of whether to grant a new trial is a question committed to the Court's discretion. *See Allied Chem. Corp. v. Da.flon, Inc.*, 449 U.S. 33, 36 (1980). Unlike the standard for judgment as a matter of law, the Court need not view the evidence in the light most favorable to the verdict winner when ruling on a motion for a new trial. *See Ateliers*, 85 F. Supp. 3d at 775. "[N]ew trials because the verdict is against the weight of the evidence are proper only when the record shows that the jury's verdict resulted in a miscarriage of justice or where the verdict, on the record, cries out to be overturned or shocks [the] conscience." *Williamson v. Consol. Rail Corp.*, 926 F.2d 1344, 1353 (3d Cir. 1991).

Tecan Ex. 2207, Page 117 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

III. DISCUSSION

A. Defendants' Motions

1. Patent Infringement

"To prove infringement, the patentee must show that an accused product embodies all limitations of the claim either literally or by the doctrine of equivalents." *Cephalon, Inc. v. Watson Pharms., Inc.*, 707 F.3d 1330, 1340 (Fed. Cir. 2013). "A two-step analysis is employed in making an infringement determination." *Intell. Ventures I, LLC v. Canon Inc.*, 104 F. Supp. 3d 629, 637-38 (D. Del. 2015) (citing *Markman v. Westview Instruments, Inc.*, 52 F.3d 967, 976 (Fed. Cir. 1995)). "First, the court must construe the asserted claims to ascertain their meaning and scope." *Id.* Second, the trier of fact must "compare the properly construed claims with the accused infringing product" to determine whether the product embodies the claims as construed. *Id.* "This second step is a question of fact." *Id.* (citing *Bai v. L & L Wings, Inc.*, 160 F.3d 1350, 1353 (Fed. Cir. 1998)).

The jury unanimously found that Defendants infringed the asserted claims of the '597 patent and the '810 patent. Defendants argue that no reasonable jury could have found infringement of either patent. For the reasons set forth below, the Court disagrees.

a. Infringement of the '597 Patent

The jury found that using QIAseq Targeted DNA Panels, QIAseq Targeted RNAscan Panels, QIAseq Immune Repertoire RNA Library Kits, QIAseq Index Kits for the Illumina and Ion Torrent platforms and GeneRead QIAact Kits (collectively "the '597 Accused Products") infringes claims 1, 5 and 19 of the '597 patent. (*See* D.I. 465 at 3-4).[2] Qiagen challenges the jury's

[2] The jury found that Defendants directly infringed by using the '597 Accused Products and also induced and contributed to infringement by their customers' use of the '597 Accused Products. The arguments Defendants raise challenging the findings of induced and

finding that the '597 Accused Products include two elements in the asserted claims: the "target-specific primer" and the "target-specific hybridization sequence." (*See* D.I. 497 at 1-8).

The asserted claims of the '597 patent require a "target-specific primer." (*See* JX-003 ('597 patent), 77:10-20, 77:57-60, 78:48-51). The Court construed this term to mean:

> a primer that has a level of complementarity between the primer and the target such that there exists an annealing temperature at which the primer will anneal to and mediate amplification of the target nucleic acid and will not anneal to or mediate amplification of non-target sequences present in a sample

(D.I. 254 at 1-2). Defendants contend that "[n]o reasonable jury could have concluded" that the accused forward primer ("FP") is a target-specific primer because it "anneals to the complement of the artificial adaptor sequence ligated to all fragments in a sample." (D.I. 497 at 2). The crux of Defendants' argument is that the FP cannot be target-specific because it effectively binds "to all sequences in a sample." (*See id.* at 3). In support of this contention, Defendants point to testimony that an adaptor primer "[b]y itself" will not "distinguish target from off-target" sequences. (Tr. at 268:4-24). Defendants argue that this evidence shows that the adaptor sequence "is [a] sequence shared across all or most libraries." (D.I. 497 at 3).

The jury, however, heard substantial evidence to the contrary. First, at trial, Plaintiffs' expert Dr. Lennon testified that "the FP primer in the QIAGEN workflow, targets the ligated adaptor" and thus satisfies the Court's construction of "target-specific primer." (Tr. at 501:2-502:25). Second, contrary to Defendants' contention, the jury heard testimony that the FP binds only to "a very small proportion of the molecules" in a sample. (Tr. at 617:11-618:2). Dr. Lennon explained that although the adaptors are present in all molecules in a sample, the FP binds to the

---

contributory infringement are directed only to the underlying act of direct infringement. Therefore, the Court does not separately address indirect infringement.

Tecan Ex. 2207, Page 119 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

adaptor *complement*, which is only present in a fraction of the molecules. (Tr. at 618:21-620:19 & 635:8-636:20). The jury was entitled to credit Dr. Lennon's testimony.

In addition, Defendants argue that the jury's finding is "inconsistent with the Court's claim construction" because the FP works "in conjunction with the GSP" to target and amplify sequences. (*See* D.I. 497 at 4). As Plaintiffs point out, however, "nothing in the Court's construction precludes the target-specific primer from 'mediat[ing] amplification of the target nucleic acid' (D.I. 254 at 1-2) in concert with another primer." (D.I. 500 at 5). The jury was thus allowed to believe Dr. Lennon's testimony that "the FP in concert with the GSP selects for and amplifies the molecules that we're interested in outlining." (Tr. at 497:10-18). In the face of conflicting evidence about whether the claim limitation "target-specific primer," as construed, was met by the '597 Accused Products, the jury was entitled to believe the evidence that the FP meets the Court's construction.

Finally, Defendants assert that no jury could find the FP to be the target-specific primer of the '597 patent because the FP was identified as a universal primer for the '810 patent. (D.I. 497 at 2). The Court sees no reason why the FP could not serve different functions with respect to the two different patents, and Defendants cite no case law that forecloses such a finding as a matter of law. Furthermore, as Plaintiffs argue, the '597 patent adopts a different definition of "target nucleic acid" than the '810 patent.[3] (D.I. 500 at 2-3 (citing JX-001 ('810 patent), 10:50-53, D.I. 254 at 1 and JX-003 ('597 patent), 15:30-32)). Therefore, it was proper for the jury to find that

[3] The '810 patent defines "target nucleic acid" as follows: "a nucleic acid molecule comprising both the nucleic acid sequence which is to be determined and the known target nucleotide sequence." (JX-001 ('810 patent), 10:50-53). In contrast, the Court used the '597 patent's definition of "target nucleic acid" to construe the term with respect to that patent: "a nucleic acid molecule of interest (e.g., a nucleic acid to be analyzed)." (D.I. 254 at 1; JX-003 ('597 patent), 15:30-32).

Tecan Ex. 2207, Page 120 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

the FP is both the '810 patent's "first adaptor primer" and the '597 patent's "target-specific primer."

The asserted claims of the '597 patent also require the target-specific primer to have a "target-specific hybridization sequence." (*See* JX-003 ('597 patent), 77:10-20, 77:57-60, 78:48-51). Per the Court's construction, a target-specific hybridization sequence is one that:

> has sufficient complementarity with a sequence of the double-stranded target nucleic acid to enable hybridization between the target-specific primer and a sequence in/of the double-stranded target nucleic acid.

(D.I. 254 at 2). Defendants contend that "Plaintiffs put on no evidence identifying the accused Double Stranded Target Nucleic Acid ("DSTNA") at trial," and thus no jury could find that the '597 Accused Products met this claim limitation. (D.I. 497 at 6). Defendants proceed to undercut their own argument, however, by citing to testimony Plaintiffs' expert offered at trial, identifying the accused DSTNA. (*See* D.I. 497 at 6 (citing Tr. at 591:17-592:13)). Plaintiffs' expert, Dr. Lennon, identified the DSTNA, explaining that "after the GSP extends, we have a double-stranded DNA molecule" to which the FP then binds (Tr. at 634:1-635:5). Defendants' own expert conceded that Dr. Lennon had identified this molecule as the required DSTNA. (Tr. at 988:10-20). Furthermore, as Plaintiffs note, "Qiagen does not dispute that [the identified DSTNA] includes a sequence to which a hybridization sequence on the FP anneals." (D.I. 500 at 7).

Defendants argue, however, that the identified molecule cannot be the claimed DSTNA because it does not yet exist at the outset of the process and thus there would be a lack of antecedent basis for "the double stranded target nucleic acid" in step b of claim 1. (*See* D.I. 497 at 7). Defendants contend that "a POSA [*i.e.,* a person of ordinary skill in the art] would not know which [DSTNA] is being referred to." (*See id.*). This, however, is a question of fact for the jury, and the jury heard testimony that "a person of ordinary skill in the art would" understand what is being

referred to because "[i]t's like reading a recipe" in which the order of the ingredients is interchangeable. (Tr. at 1135:2-15). Therefore, the jury had a legally sufficient evidentiary basis to conclude that the identified DSTNA met the '597 patent's claim limitation. *See Energizer Holdings, Inc. v. Int'l Trade Comm'n*, 435 F.3d 1366, 1370 (Fed. Cir. 2006) (quoting *Bose Corp. v. JBL, Inc.*, 274 F.3d 1354, 1359 (Fed. Cir. 2001) ("[D]espite the absence of explicit antecedent basis, '[i]f the scope of a claim would be reasonably ascertainable by those skilled in the art, then the claim is not indefinite.'")).

In sum, the jury heard substantial evidence to support its finding that the '597 Accused Products contained both a target-specific primer and a target-specific hybridization sequence. Thus, the Court must deny the Defendants' motion for judgment as a matter of law on infringement of the '597 patent.

b. Infringement of the '810 Patent

The jury found that the use of QIAseq Targeted DNA Panels, QIAseq Targeted RNAscan Panels, QIAseq Immune Repertoire RNA Library Kits and QIAseq Index Kits for the Illumina platform (collectively "the '810 Accused Products") infringe claims 16, 17 and 19 of the '810 patent under the doctrine of equivalents.[4] (*See* D.I. 465 at 2-3). Defendants assert that no reasonable jury could have found infringement because the '810 Accused Products lack three elements required by the asserted claims: a "sequence identical to a first and second sequencing primers," "adaptor primers" and the "sequence identical to first and second sequencing." (*See* D.I. 497 at 8-17).

[4] As with the '597 patent, the jury found that Defendants directly infringed and also induced and contributed to infringement from their customers' use of the '810 Accused Products, and the arguments Defendants raise as to indirect infringement address only the underlying act of direct infringement.

Tecan Ex. 2207, Page 122 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

First, the asserted claims of the '810 patent require the universal oligo tail adaptor to have a sequence identical to a first and second sequencing primers. (*See* JX-001 ('810 patent), 75:44-67, 77:19-78:18, 78:22-23; D.I. 146 at 1). At trial, Plaintiffs identified P7 and Read 2 as the first and second sequencing primers. (Tr. at 468:15-19, 474:9-475:11 & 594:2-14). Defendants' sole contention with respect to this limitation is that Plaintiffs failed to prove that P7 was a "sequencing primer." (D.I. 497 at 8-11). "Sequencing primer" was not a term the parties sought to have construed. (*See* D.I. 123). Therefore, the jury was free to rely on the plain and ordinary meaning of the term in deciding infringement. *See ePlus, Inc. v. Lawson Software, Inc.,* 700 F.3d 509, 520 (Fed. Cir. 2012). Both parties agree that P7 is a primer that "anneals fragments to a flow cell," a process called cluster generation. (*See* D.I. 497 at 8, D.I. 500 at 9 (citing Tr. at 595:19-596:16)). Dr. Lennon testified that "people of ordinary skill, understand" cluster generation to be "part of the sequencing process." (Tr. at 473:18-474:8). In addition, Dr. Lennon stated that this understanding was evidenced by one of Qiagen's own documents. (*See id.*). Although Defendants' expert testified that he did not think P7 was a "sequencing primer" (Tr. at 911:1-912:19), the jury was entitled to assess credibility and believe Plaintiffs' expert over that of Defendants. Therefore, the Court sees no reason to disturb the jury's verdict.

Second, the claims require a second adaptor primer ("SAP") with a "portion identical to a first sequencing primer." (*See* JX-001 ('810 patent), 75:44-67, 77:19-78:18, 78:22-23; D.I. 146 at 2). Qiagen argues that the jury had a legally insufficient basis to find that (a) P7 is the first sequencing primer and (b) the SAP has a sequence identical to a "portion" of P7. (*See* D.I. 497 at 11-12). Defendants' first argument fails for the reasons noted above. As to the second argument, the jury heard testimony that all 24 nucleotides of the SAP are identical to all 24 nucleotides of P7. (Tr. at 915:15-19). Defendants do not dispute this testimony, but rather argue that it was

Tecan Ex. 2207, Page 123 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

improper for the jury to find that identical "to a portion" may include identical to the whole. (*See* D.I. 497 at 11-12). As Defendants note in the next section of their brief, however, the '810 specification states that "[a] portion can comprise all or only a subset of the nucleotides comprised by the molecule." (JX-001 ('810 patent), 9:1-5). Thus, the jury was entitled to credit Dr. Lennon's testimony that the SAP "has a sequence identical to a portion of the first sequencing primer." (Tr. at 481:18-21).

Third, Qiagen contends that "no reasonable jury could have concluded that the [second target-specific primer ("STSP")] has [a] sequence identical to a second sequencing primer," as required by the asserted claims. (D.I. 497 at 13). As Plaintiffs point out, however, the jury heard evidence that the SIP (the accused STSP) has a 19-base sequence that is identical to 19 of the 34 bases in Read 2 (the second sequencing primer). (D.I. 500 at 11 (citing Tr. at 610:14-613:1 & 916:12-24)). The jury was not precluded from finding that "identical" may include identical to a portion, and therefore was allowed to find for Plaintiff. *See ePlus, Inc. v. Lawson Software, Inc.*, 700 F.3d 509, 520 (Fed. Cir. 2012) ("In the absence of . . . a construction, however, the jury was free to rely on the plain and ordinary meaning of the term . . . ."). !

Finally, Defendants challenge the jury's finding that the SIP infringes the second target-specific primer limitation under the doctrine of equivalents. (*See* D.I. 497 at 13-17). An accused device that does not literally infringe may still infringe under the doctrine of equivalents "if there is 'equivalence' between the elements of the accused product or process and the claimed elements of the patented invention." *Warner-Jenkinson Co. v. Hilton Davis Chem. Co.*, 520 U.S. 17, 21 (1997). Analysis under the doctrine of equivalents follows one of two tests endorsed by the Supreme Court – the insubstantial differences test or function-way-result test – both of which are performed on an element-by-element basis. Both parties use the function-way-result framework

Tecan Ex. 2207, Page 124 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

in their post-trial briefs. (*See* D.I. 497 at 13, D.I. 500 at 12). The function-way-result test evaluates whether the element in the accused product performs substantially the same function, in substantially the same way, to achieve substantially the same result as the claimed element. *Warner-Jenkinson,* 520 U.S. at 39-40.

Defendants argue that Plaintiffs put forth insufficient evidence to meet this test. Regarding the first element, Defendants contend that "the function of a target-specific primer in the '810 patent is to anneal to and mediate amplification of [the] target nucleic acid" and thus "enrich[] the target sequence." (D.I. 497 at 13-14). They argue that the SIP does not perform substantially the same function because it (a) "bind[s] to [the] common tail sequence added by the GSP" and (b) enriches the target sequence only "when paired with the adaptor primer." (*Id.*). Defendants' arguments rely on the incorrect assumption that the jury was required to credit Defendants' expert's testimony of the STSP's "function" over that of Dr. Lennon, Plaintiffs' expert. Dr. Lennon identified three functions of the STSP that the SIP performs: "[1] amplifying the amplification product of Step 2, [2] producing an amplification product that is capable of being sequenced, and also [3] reducing undesirable amplification products by being paired with the second adaptor primer." (Tr. at 487:21-488:15). The jury was entitled to credit this testimony.

With respect to the second element, Defendants argue that the SIP does not perform in substantially the same way because (a) it does not "bind to [the] known target nucleotide sequence" and (b) it is not "nested with respect to the first-target-specific primer." (D.I. 497 at 14). Again, Defendants fail to show why the jury was required to credit Defendants' narrow interpretation of "substantially the same way" over that of Plaintiffs. Dr. Lennon testified that the SIP performs the functions he identified in "the same way," by, for example, amplifying "the product of the first

Tecan Ex. 2207, Page 125 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

amplification reaction by pairing with a nested opposing primer." (Tr. at 488:2-15). The jury was free to make credibility determinations or to accept Plaintiffs' expert testimony.

Defendants also contend that the '810 patent "specifically teaches not to use hemi-nested primers." (D.I. 497 at 14). Therefore, they argue that it was improper for the jury to find that the accused SIP works in substantially the same way as the claimed STSP because the SIP is not nested with respect to the first target-specific primer, and thus their products use a hemi-nested method. (*See id.*). Per the Court's construction, a "nested" primer is one that "anneal[s] to a nucleic acid sequence 3' downstream and in the same direction as" the other primer. (D.I. 146 at 2). At trial, experts explained that primers are said to be "hemi-nested" when "the primers on one side are nested, and the primers on the other side are not." (*E.g.*, Tr. at 219:20-22). As Plaintiffs explain, however, "there are two types of hemi-nested methods – those in which only the adaptor primers are nested, and those in which only the target specific primers are nested." (D.I. 500 at 13). The jury heard evidence from at least three witnesses that the '810 patent discourages only one of the hemi-nested methods: methods in which the target-specific primers are nested and adaptor primers are not nested. (Tr. at 220:4-24, 538:4-7 & 779:10-14). Qiagen does not dispute that its products use the opposite approach: non-nested target-specific primers and nested adaptor primers. (*See* D.I 499 at 14; *see also* Tr. at 779:10-18; D.I. 500 at 14). As Plaintiffs point out, the '810 patent in fact "expressly discloses" the use of nested adaptor primers. (D.I. 500 at 14 (citing JX-001 ('810 patent), 17:43-45) ("The use of two adaptor primers, as described herein [*i.e.*, nested adaptor primers] can reduce, and in some embodiments eliminate, these problems.")). The jury, therefore, had an ample basis for rejecting Defendants' argument and determining that the SIP worked in substantially the same way as the claimed invention.

With respect to the third element, Defendants assert that no reasonable jury could find that the SIP achieves substantially the same result as the STSP because the result of Qiagen's process is "simply to copy everything that was amplified in the first round of PCR rather than further enriching for the target nucleic acid of interest." (D.I. 497 at 15). Again, Defendants' argument asks the Court to disregard Plaintiffs' proffered evidence. At trial, Dr. Lennon explained that the SIP and STSP achieve substantially the same result with respect to the three functions he identified. (Tr. at 488:16-489:7). This evidence was unrebutted at trial and Defendants do not attempt to rebut it in their post-trial briefs. (*See* D.I 500 at 16, D.I. 497 at 15-17 & D.I. 514 at 7). Rather, Defendants merely repeat their own expert's opinion on the matter. In the face of conflicting testimony about whether Qiagen's products perform substantially the same function, in substantially the same way, to achieve substantially the same result, the jury was entitled to credit Plaintiffs' evidence.

In sum, substantial evidence supports the jury's finding that Defendants infringed the '810 patent. Therefore, the Court will deny Defendants' motion for judgment as a matter of law on infringement of the '810 patent.

2. Patent Validity

An issued patent is presumed valid. *See* 35 U.S.C. § 282. To overcome this presumption, a party must show by clear and convincing evidence that the patent is invalid. *See Hewlett–Packard Co. v. Bausch & Lomb, Inc.*, 909 F.2d 1464, 1467 (Fed. Cir. 1990). Defendants challenged the validity of the '810 patent and '597 patent at trial on multiple grounds. (*See* D.I. 465 at 5-7). Defendants now argue that no reasonable jury could have found that Defendants failed to meet their burden with respect to two of those grounds: written description and definiteness. For the reasons set forth below, the Court disagrees.

Tecan Ex. 2207, Page 127 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

a. Written Description

For a patent to be valid, its specification must contain a written description that "clearly allow[s] persons of ordinary skill in the art to recognize that [the inventor] invented what is claimed." *Ariad Pharms., Inc. v. Eli Lilly & Co.*, 598 F.3d 1336, 1351 (Fed. Cir. 2010) (*en banc*). "In other words, the test for sufficiency is whether the disclosure of the application relied upon reasonably conveys to those skilled in the art that the inventor had possession of the claimed subject matter as of the filing date." *Id.* !

Starting with the '810 patent, Defendants contend that although the patent "covers a non-nested STSP," a "POSA reading the '810 specification would not understand the '810 inventors were in possession of such an invention." (D.I. 497 at 17). Defendants point to testimony that the patent only explicitly includes examples in which the STSP was nested. (D.I. 497 at 17 (citing Tr. at 953:16-23 & 211:14-17)). Plaintiffs, however, point to testimony that identified support for non-nested STSP's in the specification. (Tr. at 1126:17-1127:15 ("[T]here is a written description . . . in the body of the patents, description text, that describes the workflow in a way that's not nested.")). The jury was entitled to credit Plaintiffs' evidence.

In addition, Defendants assert that inventor Dr. Iafrate claimed that "he and the other inventors had tried" the process with non-nested target-specific primers "but it did not work." (D.I. 514 at 9). Defendants overstate that evidence. At trial, Dr. Iafrate testified that he had worked on the non-nested approach and "[i]t worked," only "the nested approach worked better." (Tr. at 215:12-18). As Plaintiffs state, however, "the claims do *not* require the STSP to achieve an on-target rate or level of specificity." (D.I. 500 at 21). Thus, Dr. Iafrate's testimony does not support Defendants' contention that '810 fails the written description requirement. The jury had sufficient basis for finding that Defendants failed to meet their burden.

Turning to the '597 patent, Defendants argue that "a POSA reading the [patent] would not understand the inventors possessed an invention that ligated a universal adaptor onto the sequence of interest and called a primer that anneals to the universal sequence, a target-specific primer." (D.I. 497 at 18). As Plaintiffs note, Defendants fail to point to any relevant testimony in support of this contention. (*See* D.I. 500 at 21-22). Rather, Defendants cite Dr. Metzker's testimony regarding his noninfringement theories which do not "even purport[] to apply the relevant legal standard." (D.I. 500 at 21-22 & D.I. 497 at 18-19). This is not clear and convincing evidence. Defendants have thus given no basis on which the Court may grant judgment as a matter of law with respect to their written description claims.

b. Indefiniteness

Patent claims must be "sufficiently definite." *Allen Eng'g Corp. v. Bartell Indus., Inc.*, 299 F.3d 1336, 1348 (Fed. Cir. 2002). "[A] patent is invalid for indefiniteness if its claims, read in light of the specification delineating the patent, and the prosecution history, fail to inform, with reasonable certainty, those skilled in the art about the scope of the invention." *Nautilus, Inc. v. Biosig Instruments, Inc.*, 572 U.S. 898, 901 (2014). Definiteness "is amenable to resolution by the jury where the issues are factual in nature." *BJ Servs. Co. v. Halliburton Energy Servs., Inc.*, 338 F.3d 1368, 1372 (Fed. Cir. 2003); *see also Bombardier Recreational Prods. Inc. v. Arctic Cat Inc.*, 785 F. App'x 858, 867 (Fed. Cir. 2019) ("Because the jury found the claims indefinite, we presume that it resolved the underlying factual issues relating to the construction of [the disputed term] in [the accused infringer's] favor.").

Starting with the '810 patent, Defendants argue that the '810 claims are indefinite "because there is no description in the specification of what constitutes a sequence in the adaptor primers being identical to a portion of the sequencing primers." (D.I. 497 at 19). That is, "the patent does

Tecan Ex. 2207, Page 129 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

not inform a POSA if a 'portion' means one, two, three, four, etc. nucleotides." (*Id.*). Defendants' entire argument rests on the following eight lines cited from the trial transcript:

> Q: And what is your opinion here?
>
> A: This is a different argument. This means it's indefinite. It means – the Court has construed two terms, universal oligonucleotide tail adaptor that has sequences identical to a first and second sequencing primer. It's also construed a second target-specific primer that also has sequences identical to a second sequencing primer. The specification doesn't help you whether these two second sequencing primers are the same of if they are different.

(Tr. at 954: 15-23). This excerpt fails even to mention the word "portion"[5] and is not clear and convincing evidence that the patent fails to inform a POSA, with reasonable certainty, about the scope of the invention as to what a portion of the sequencing primers means.[6] Therefore, the jury was entitled to find for Plaintiffs on the issue of indefiniteness.

Turning to the '597 patent, Defendants contend that "a POSA would have no idea which ' . . . double stranded target nucleic acid' is being referred to throughout claim 1 element b." (D.I.

---

[5] As Plaintiffs note, this testimony appears to address Defendants' other argument regarding indefiniteness (*i.e.*, regarding whether the second sequencing primer in the STSP and in the universal oligonucleotide tail adaptor must be the same primer). (*See* D.I. 500 at 22).

[6] Plaintiffs argue that Qiagen has waived this issue as this theory of indefiniteness was "never presented to the jury." (D.I. 500 at 22). Although Defendants dedicated a brief sentence to this theory (*i.e.*, that the word "portion" is indefinite) in their Rule 50(a) motion, Plaintiffs are correct in stating, "the only indefiniteness argument that Dr. Metzker presented to the jury involved a different theory – namely, whether the 'second sequencing primer' in the STSP and in the universal oligonucleotide tail adaptor must be the same primer." (Tr. at 1161:5-9; D.I. 500 at 22 (citing Tr. at 954:14-23)). Defendants counter that the question of indefiniteness is "ultimately a question of law for the Court to decide." (D.I. 514 at 10). If the Court should decide this question as a matter of law on the record before it, the Court finds in favor of Plaintiffs because (as discussed above) Defendants have failed to meet their burden of showing definiteness by any – let alone clear and convincing – evidence. Thus, the Court declines to reach the question of whether Defendants waived this theory.

497 at 20). Defendants repeat their arguments regarding non-infringement of the '597 patent (*i.e.*, that there is a lack of antecedent basis for "the" DSTNA in claim 1 element b). The Court dealt with these arguments above, and Defendants cite to no new testimony in support of their indefiniteness claim. (*See supra* § III.A.1.a). As noted above, Plaintiffs offered testimony that "a person of ordinary skill in the art would" understand what is being referred to because "[i]t's like reading a recipe" in which the order of the ingredients is interchangeable. (Tr. at 1135:2-15). The jury, therefore, had legally sufficient basis to find that Defendants failed to meet their burden in proving either patent invalid for indefiniteness. *See Energizer Holdings, Inc. v. Int'l Trade Comm'n*, 435 F.3d 1366, 1370 (Fed. Cir. 2006) ("[D]espite the absence of explicit antecedent basis, '[i]f the scope of a claim would be reasonably ascertainable by those skilled in the art, then the claim is not indefinite.'").

In sum, the Court sees no reason to disturb the jury's findings that the '810 and '597 patents are not invalid.

3. Willfulness

A determination of willfulness requires a finding of "deliberate or intentional" infringement. *SRI Int'l, Inc. v. Cisco Sys., Inc.*, 14 F.4th 1323, 1330 (Fed. Cir. 2021). Plaintiffs must show that "the accused infringer knew of the patent-in-suit, and knowingly or intentionally infringed the patent after acquiring that knowledge." *Jackson v. Seaspine Holdings Corp.*, No. CV 20-1784-RGA, 2022 WL 610703, at *7 (D. Del. Feb. 14, 2022) (citing *Eko Brands, LLC v. Adrian Rivera Maynez Enters., Inc.*, 946 F.3d 1367, 1378-79 (Fed. Cir. 2020)). The jury found that Defendants willfully infringed both the '597 and '810 patents. (*See* D.I. 465 at 8). Defendants ask the Court to grant a new trial on the jury's willfulness finding. (D.I. 497 at 20). For the reasons set forth below, the Court denies Defendants' request.

Tecan Ex. 2207, Page 131 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

Defendants first argue that a new trial on willfulness is warranted because there was "insufficient evidence to establish that infringement was willful." (*Id.* at 21.). The Court disagrees. First, Qiagen itself admits that Dr. Lader, a vice president of research and development at Qiagen, had been aware of Plaintiffs' patents since 2015 and reviewed the '810 patent in 2018 regarding its relevance to Qiagen products and the '597 patent later. (D.I. 497 at 20-21 & D.I. 500 at 19). Both he and Dr. Wang, senior director for research and development at Qiagen, "carefully studied" Plaintiffs' patents. (D.I. 497 at 20). In addition, Plaintiffs point to Dr. Lader's admission that Qiagen maintained a "patent watch list" for other companies but chose not to include Archer on the list. (D.I. 500 at 19 (citing Tr. at 681:15-683:21)). This is sufficient evidence to support the jury's finding that Defendants (a) had knowledge of the patents and (b) knowledge of infringement. Dr. Lader's assertion that he did not "believe" Qiagen infringed either patent does not preclude a finding of willfulness. (*See* D.I. 497 at 20-21 (citing Tr. at 691:7-692:20 & 701:4-10) ("I believe Qiagen's products do not infringe the '597 patent . . . I found the claims to be rather confusing, but, yes, I reached that conclusion.")). The jury was entitled to assess credibility and infer willfulness from the evidence presented, particularly given the fact that Dr. Lader admitted he was "not a patent expert." (Tr. at 701:19-20). This is thus not a case "where the verdict, on the record, cries out to be overturned." *Williamson*, 926 F.2d at 1353.

In addition, Defendants argue a new trial on willfulness is warranted because "Plaintiffs' counsel poisoned the well by making legally impermissible arguments" before the jury when counsel referenced Defendants' alleged failure to receive a freedom to operate opinion from in-house counsel.[7] (D.I. 497 at 21). First, as noted above, the jury heard substantial other evidence

[7] Defendants cite three instances in which Plaintiffs referenced Qiagen's lack of a freedom to operate ("FTO") at trial. (Tr. at 864:15-865:8, 1244:1-4 & 1245:14-1246:5). Plaintiffs counter that, in all three instances, Plaintiffs were attempting to mitigate Qiagen's

Tecan Ex. 2207, Page 132 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

on which it could properly base its finding. Defendants point to nothing that suggests the verdict was the result of improper prejudice. Second, as Plaintiffs note, "the Court gave Qiagen the choice between a curative instruction and an opportunity to address the purportedly prejudicial statements in closing arguments." (D.I. 500 at 18 (citing Tr. at 1266:1-4)). Defendants opted to address the issue in closing and then chose not to do so. (Tr. at 1266:5-8). Having refused to take either opportunity to repair any damage, Defendants cannot now argue that they believe the statements so prejudicial to have improperly influenced the jury verdict.

Given the evidence in support of the jury's finding of willfulness, the Court denies Defendants' motion for new trial on willfulness.

"improper insinuation that Qiagen had received an exculpatory opinion." (D.I. 500 at 18). The "improper insinuation" to which Plaintiffs refer is excerpted below:

> Q: [What does Qiagen do] when it thinks it's using somebody else's technology?
>
> A: What QIAGEN does when it thinks it's – first of all, we have – I mean, I would say, a pretty good IP department, intellectual, property department, and these people are interacting a lot with the R&D folks and they are discussing about patents or potential infringement, if there is any. We also very often use external counsel, so we try to do our best to make sure that we have the rights to sell the products we're selling.

(Tr. at 809:8-17). Defendants argue that Plaintiffs' references to the FTO improperly asks the jury to make an adverse inference based on the failure to provide an opinion of counsel in violation of *Knorr-Bremse*. (D.I. 497 at 21 (citing *Knorr-Bremse Systeme Fuer Nutzfahrzeuge GmbH v. Dana Corp.*, 383 F.3d 1337, 1344 (Fed. Cir. 2004))). Plaintiffs counter that such references were permissible because Defendants "opened the door" by implying they had received an FTO in the exchange cited above (and thus, *Knorr Bremse* is inapplicable). (D.I. 500 at 18) (citing *Hologic, Inc. v. Minerva Surgical, Inc.*, 2018 WL 3348998, at *2 (D. Del. July 9, 2018)). The Court finds that, regardless of whether Plaintiffs' statements violated *Knorr-Bremse*, Defendants are not entitled to a new trial. Therefore, the Court declines to revisit whether such statements were improper in light of *Knorr-Bremse* and Defendants' own testimony cited above.

Tecan Ex. 2207, Page 133 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

4. Damages

The jury awarded Plaintiffs $841,756 in lost profits in connection with Defendants' RNA-related products sold in the Unites States, $1,593,762 as a reasonable royalty for sales of DNA-related products in the Unites States and $2,240,303 as a reasonable royalty on products sold to customers outside the United States. (*See* D.I. 465 at 8). Defendants challenge the jury award of lost profits and foreign damages and seek remittitur of the royalty for U.S. sales.

a. Lost Profits

To be entitled to lost profits, Plaintiffs must "show a reasonable probability that, but for the infringement, [they] would have made the sales made by" Qiagen. *Presidio Components, Inc. v. Am. Tech. Ceramics*, 875 F.3d 1369, 1380 (Fed. Cir. 2017). At trial, Plaintiffs sought to establish lost profits using the *Panduit* factors. (Tr. at 721:24-731:22). "The Panduit test requires the patentee to show: (1) 'demand for the patented product'; (2) 'absence of acceptable noninfringing substitutes'; (3) 'manufacturing and marketing capability to exploit the demand'; and (4) 'the amount of profit that ... would have [been] made.'" *Georgetown Rail Equip. Co. v. Holland L.P.*, 867 F.3d 1229, 1240-41 (Fed. Cir. 2017) (quoting *Panduit Corp. v. Stahlin Brothers Fibre Works, Inc.*, 575 F.2d 1152, 1156 (6th Cir. 1978)).

Defendants challenge the lost profits award, arguing that (a) there was insufficient evidence for the jury to find an absence of non-infringing alternatives, (b) there was no evidence that "any customer purchased [Defendants'] RNA panels because of the" infringing technology and (c) Plaintiffs presented no market share analysis. For the reasons below, the Court denies Defendants' motion.

First, Defendants contend that Plaintiffs failed to show the absence of non-infringing alternatives because some evidence showed that Archer had "multiple market competitors." (*See* D.I. 497 at 22). As Plaintiffs note, however, "[m]ere existence of a competing device does not

make that device an acceptable substitute." *Presidio Components, Inc. v. Am. Tech. Ceramics Corp.*, 702 F.3d 1351, 1361 (Fed. Cir. 2012) (quoting *TWM Mfg. Co. v. Dura Corp.*, 789 F.2d 895, 901 (Fed. Cir. 1986)). Rather, whether a competing device should be considered an acceptable substitute is a question of fact for the jury. *See Minnesota Min. & Mfg. Co. v. Johnson & Johnson Orthopaedics*, Inc., 976 F.2d 1559, 1577 (Fed. Cir. 1992). Here, the jury heard testimony from numerous witnesses that the supposed competing products were "not viable options," used technology that was "a little bit ancient," could lead to patients being "wrongfully misdiagnosed," did not work with RNA at all, or were not "competitive." (Tr. at 510:25-512:4, 356:13-357:13, 353:18-22, 360:17-362:3, 709:7-21 & 712:20-713:11). Much of the evidence Defendants claim undermines the jury's finding pertains to Illumina's RNA products. (*See* D.I. 497 at 22-23 (citing Tr. 1-81:7-18, 383:11-384:7 & 388:12-389:2)). The jury, however, heard testimony that Illumina's products were inferior and thus not an acceptable substitute. (*See, e.g.*, Tr. at 356:22-357:13 (explaining that Illumina requires larger sample sizes from patients and thus leads to a more "traumatic" experience and that the technology used was "a little bit ancient")). In addition, much of the evidence Defendants cite in support of their argument refers to Plaintiffs' competition more broadly (rather than in the market for which lost profits was awarded, *i.e.*, RNA products in the United States). (*See* D.I. 497 at 22 (citing DX-327.145 (referring to Archer's competition in "life sciences more broadly") and Tr. at 331:1-1, 396:12-397:1, 705:20-25, 712:3-11 & 1020:12-1030:17 (discussing library preparation in general))). Based on this evidence, the jury could reasonably have found that Plaintiffs met their burden in proving the absence of non-infringing alternatives.

Second, Defendants argue that Plaintiffs are not entitled to lost profits because they "presented no evidence that any customer" purchased the infringing products because of the

patented technology. (D.I. 497 at 24). This argument fails as a matter of law. A patentee is not required to provide direct evidence that customers bought the infringing products because of the patented technology to be entitled to an award of lost profits. *See Panduit*, 575 F.2d at 1156 (establishing elements patentee is required to show for an award of lost profits). Regardless, Qiagen's witness testified that the patented "single primer extension" technology is "the underlying reason why [customers] work with us." (Tr. at 713:18-714:15).

Third, Defendants assert that Plaintiffs were required to put forth evidence of market share in order for the jury to properly award lost profits. The jury, however, awarded Plaintiffs lost profits for 100% of Defendants' RNA sales in the United States, thus implying a finding that the relevant market was a two-supplier market. In such a case, evidence of market share is irrelevant. *Mentor Graphics Corp. v. EVE-USA, Inc.*, 851 F.3d 1275, 1286 (Fed. Cir. 2017) ("The market share theory is irrelevant in this case because the jury made a factual finding . . . that the relevant emulator market for sales to Intel was a two-supplier market."). ‼

Defendants have failed to show that the jury's award of lost profits must be overturned.

b. Foreign Damages

Defendants argue that Plaintiffs were not entitled to royalties for products sold to purchasers outside of the United States because (a) allowing foreign damages when method claims are asserted and infringement is found under 35 U.S.C. 271(a) is contrary to the Supreme Court's decision in *WesternGeco LLC v. ION Geophysical Corp.*, 138 S. Ct. 2129 (2018) and (b) Plaintiffs presented insubstantial evidence that infringement in the United States was a substantial cause of the foreign sales.

First, at issue is the applicability of the Federal Circuit's holding in *Carnegie Mellon Univ. v. Marvell Technology Group, Ltd.*, 807 F.3d 1283 (2015) ("*CMU*"). This matter was hotly contested at trial. Plaintiffs argued that *CMU* allowed the jury to use sales made outside of the

United States to calculate damages for infringement in the United States if Archer could prove that (1) Defendants' infringement in the United States was a cause of the sale of that product and (2) Defendants made or sold the product within the United States. (*See* D.I 452 & D.I. 457). Defendants countered that doing so would "run[] contrary to longstanding precedent" that held that performing steps of a patented method abroad is not an act of infringement and thus a patentee has no right to damages for such uses. (D.I. 456 at 1-3). After briefing on the matter was completed, the Court issued an order agreeing to a modified version of Plaintiffs' proposed jury instructions on damages calculated based on sales to foreign customers. *ArcherDX, LLC v. QIAGEN Scis., LLC*, No. CV 18-1019 (MN), 2021 WL 3857460 (D. Del. Aug. 30, 2021). The Court reasoned,

> The issue here . . . is not whether the foreign uses of the patented methods are infringing – they are not and Plaintiffs concede that. The question is whether the sales of products that use the methods to foreign users can be used to measure damages for acts of infringement in the United States. United States patent law allows "damages adequate to compensate for the infringement." Here, the infringement is asserted under § 271(a), which, for the method claims at issue, makes it an act of infringement to use the claimed methods in the United States. The claims for induced infringement under § 271(b) and contributory infringement under § 271(c) also require direct infringement under § 271(a). And, as I said, that infringement must be performance of the method claim in the United States.
>
> In *Carnegie Mellon University v. Marvell Technology Group, Ltd.*, 807 F.3d 1283 (Fed. Cir. 2015) [("CMU")], the Federal Circuit addressed calculation of damages for infringement of a method claim that relied on the sales of products that perform that method. Recognizing the presumption against extraterritoriality, the *CMU* court nevertheless concluded that:
>
> > 'Where a physical product is being employed to measure damages for the infringing use of patented methods, . . . territoriality is satisfied when and only when any one of those domestic actions for that unit (e.g., sale) is proved to be present, even if others of

> the listed activities for that unit (e.g., making, using) take place abroad. Significantly, once one extends the extraterritoriality principle to confining how damages are calculated, it makes no sense to insist that the action respecting the product being used for measurement itself be an infringing action. Thus, here the claim is a method claim, but the damages-measuring product practices the method in its normal intended use.'
>
> . . . [Therefore,] *CMU* makes clear that it is addressing damages for infringement – that is, damages for actions in the United States – and is not expanding the statutory requirement for infringement.

Id. at *1-2 (internal citations removed). The jury was thus instructed that it could award use sales of products that practice the patented method outside the United States to measure damages if "(1) QIAGEN's infringement in the Unites States was a substantial[8] cause of the sale of that product, and (2) QIAGEN made or sold the product within the United States." (Tr. at 1207:9-15).

Now, Defendants argue that the Supreme Court's decision in *WesternGeco* implicitly overruled *CMU*. (*See* D.I. 497 at 26). The Court disagrees. In *WesternGeco*, the Court held that a patentee was entitled to damages for sales made abroad when infringement was found under 35 U.S.C. § 271(f)(2). *WesternGeco*, 138 S. Ct. at 2137-39. The Court reasoned that "the conduct relevant to the statutory focus in [*WesternGeco*] is domestic." *Id.* at 2137. That is, infringement under § 271(f)(2) regulated "the domestic act of 'suppl[ying] in or from the United States" components of patented inventions for combination abroad. *Id.* at 2138 (quoting 35 U.S.C.

---

[8] The Court amended Plaintiffs' proffered instructions by requiring a "substantial" causal connection. As noted in the Court's prior order, "[a]lthough *CMU* does not refer to a 'substantial' causal connection, the Court understands that there must be more than a tangential relationship between the infringement asserted and the volume of sales." *ArcherDX,* 2021 WL 3857460 at *2 n.8 (citing *WesternGeco LLC*, 837 F.3d at 1368 (Wallach, J., dissenting) (noting that "where the volume of non-infringing sales is independent of the extent of United States infringement, those sales should not be used as a measure of damages flowing from the domestic infringement")).

Tecan Ex. 2207, Page 138 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

271(f)(2)). Therefore, lost-profits damages for sales made abroad were properly awarded as "a domestic application of § 284." *Id.* That holding is not inconsistent with *CMU* and this Court's prior application of *CMU* to the case before it. Although in *CMU* and the present case, infringement was found under § 271(a), "the conduct relevant to the statutory focus . . . is domestic," *i.e.,* use in the United States. *See CMU*, 807 F.3d at 1306-07; *WesternGeco*, 138 S. Ct. at 2137. Therefore, by allowing the jury to calculate damages for domestic infringement by using sales made abroad related to that infringement, the Court is not implicitly finding that infringement occurred abroad but rather is allowing the patentee to recover fully for harm committed in the United States.[9]

Second, Defendants argue that "[e]ven under Plaintiffs' interpretation of [*CMU*], Plaintiffs failed to present substantial evidence to support damages for sales outside the U.S." (D.I. 497 at 28). The Court disagrees. As noted in the Court's prior order, "[i]n *CMU*, the Federal Circuit noted that the products 'practice[] the method in its normal intended use' and concluded that causation to domestic infringing uses was established given the design, simulation, and testing of the chips in California involved infringing uses and caused the worldwide sales." *ArcherDX,* 2021 WL 3857460, at *1 n.6 (citing *CMU*, 807 F.3d at 1306-07 ("all of Marvell's sales are strongly enough tied to its domestic infringement as a causation matter")). Here, Plaintiffs put forth evidence that "customers in this industry require data providing that products work;" "customers buy Qiagen's products for the infringing technology;" "generating data to prove that infringing technology works required Qiagen to infringe the asserted patents;" and "Qiagen's testing (*i.e.*,

[9] That being said, the Court takes Defendants' point that to some extent *CMU* is difficult to reconcile with other governing law. *See ArcherDX*, 2021 WL 3857460, at *2. Indeed, at trial, the Court struggled with this issue. Absent further direction from the Federal Circuit, however, the Court declines to amend its previous holding.

Tecan Ex. 2207, Page 139 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

infringement) of the infringing products occurred in [the] U.S." (D.I. 500 at 29 (citing Tr. at 358:2-359:20, 713:18-714:15, 453:1-508:1 & 679:8-24)). The jury could thus reasonably infer from this evidence that Qiagen's domestic infringement (use of the accused products) was a substantial cause of the sale of products abroad.

c. Remittitur

Defendants seek remittitur of the jury's award of $1,593,762 in royalties for sales of DNA-related products in the United States on the basis that the award is higher than any estimate provided by the parties' experts and appears to be the result of a mistake on the part of the jury. (D.I. 497 at 29-30). Defendants point out that both the lost profits award and the royalty award for foreign sales match Plaintiffs' expert's estimates to the dollar. (D.I. 497 at 30 (citing Tr. at 732:19 & 743:13-21)). The $1,593,762 U.S. royalty award, however, is much higher than the estimate provided by Plaintiffs' expert, and, in fact is equivalent to the expert's suggested award for lost profits ($841,756) *plus* U.S. royalties ($752,006). (Tr. at 743:15-21). Defendants thus request the Court remit the award to the highest amount presented at trial: $752,006.

Plaintiffs agree that "the jury intended to award Plaintiffs a 7% royalty for infringing DNA-related products in the United States." (D.I. 500 at 30). Plaintiffs, however, contend that Qiagen waived its objection by failing to object to the verdict before the jury was excused. (*See id.*). In support of this argument, Plaintiffs cite *Frank C. Pollara Grp., LLC v. Ocean View Inv. Holding, LLC*, 784 F.3d 177, 190-91 (3d Cir. 2015). In *Pollara*, however, the court held, "if a party fails to object to an inconsistency in a general verdict before the jury is excused, that party waives any objection in that regard." *Id.* In *Pollara*, the issue was that the verdict was facially inconsistent. *Id.* ("Appellants . . . complain that the jury's verdict was inconsistent because its finding of intentional misrepresentation by Cheng precludes a finding of negligent misrepresentation by OMEI."). Here, Defendants argue that remittitur is appropriate not because of an inconsistent

verdict but because the verdict is unsupported by the evidence and the result of mistake. Plaintiffs have not cited to any case law in which the rule in *Pollara* is applied to cases like the one before the Court. *Cf. Delaware & Hudson Ry. Co., Inc. v. Knoedler Mfrs., Inc.*, 805 F. App'x 149, 152 (3d Cir. 2020) (applying *Pollara* in the context of a claim of inconsistent verdict); *EMC Corp. v. Zerto, Inc.*, No. CV 12-956(GMS), 2016 WL 1291757, at *6 (D. Del. Mar. 31, 2016), *aff'd*, 691 F. App'x 623 (Fed. Cir. 2017) (same).

There simply is no evidentiary basis to support the $1,593,762 U.S. royalty award. And given that both parties agree the jury award was the result of a mistake, the Court will grant remittitur of the jury's award for U.S. royalties to the amount estimated by Plaintiffs' witness (the highest amount presented to the jury): $752,006. *See Smith v. Katz*, 696 F. App'x 582, 591 (3d Cir. 2017) ("[T]he remittitur is well established as a device employed when the trial judge finds that a decision of the jury is clearly unsupported and/or excessive.") (quoting *Cortez v. Trans Union, LLC*, 617 F.3d 688, 716 (3d Cir. 2010)); *see also Belardinelli v. Carroll*, 773 F. Supp. 657, 659 (D. Del. 1991) (granting remittitur because the "awards are not rationally related to the evidence and the evidence suggests that the jury may have mistakenly believed that compensation was sought for the decedent's death"); *Lewis v. United States*, No. 02-2958-STA-EGB, 2014 WL 12828838, at *3-4 (W.D. Tenn. Sept. 26, 2014) (granting remittitur due to jury's mistake in adding estimated damages shown at trial).

Plaintiffs may either accept the remittitur or receive a new trial on these damages. *See Cortez v. Trans Union, LLC*, 617 F.3d 688, 716 (3d Cir. 2010) ("[T]he court must offer a new trial as an alternative to a reduction in the award" when remittitur is due to insufficient evidence.).

B. Plaintiffs' Motions

Before the Court are Plaintiffs' motions for injunction, ongoing royalty, enhanced damages, supplemental damages and pre- and post-judgment interest.

1. Injunction

Plaintiffs ask the Court to permanently enjoin Defendants from "offering [infringing] products and services approved for clinical diagnosis by a regulatory body." (D.I. 496 at 1-2).

A permanent injunction does not automatically "issue once infringement and validity have been adjudged." *eBay Inc. v. MercExchange, L.L.C.*, 547 U.S. 388, 393-94 (2006). Rather, a plaintiff seeking a permanent injunction must satisfy a four-factor test before a court may grant such relief. A plaintiff must demonstrate: (1) that it has suffered an irreparable injury; (2) that remedies available at law, such as monetary damages, are inadequate to compensate for that injury; (3) that, considering the balance of hardships between the plaintiff and defendant, a remedy in equity is warranted; and (4) that the public interest would not be disserved by a permanent injunction. *Id.* at 391; *see also Apple, Inc. v. Samsung Elecs. Co.*, 695 F.3d 1370, 1374 (Fed. Cir. 2012).

In support of their arguments for a permanent injunction, Plaintiffs rely on evidence entered during trial as well as new evidence in the form of several declarations – one from an expert, Dr. Maria Fe Paz, who previously submitted an expert report, but who did not testify at trial, and two from Invitae employees.[10] (*See* D.I. 496 at 1-12 & D.I. 496, Ex. A-C; *see also* D.I. 501 at 17). Plaintiffs relied heavily on these declarations to support their arguments on direct competition, lost market share and the adequacy of monetary damages to compensate them. In response, Defendants attack the assertions and opinions in the declarations and request an evidentiary hearing on the injunction issue. (*See* D.I. 501 at 17). The Court agrees that a hearing at which witnesses can be cross-examined and their opinions tested would be helpful in making the credibility determinations

[10] In October of 2020, Invitae acquired Archer. (D.I. 496, Ex. B ¶ 1).

necessary to decide whether injunctive relief should issue. Thus, the Court will deny Plaintiffs' motion for an injunction subject to renewal after a hearing.

2. Ongoing Royalties

Plaintiffs request ongoing royalties for all infringing products not enjoined of 10% for sales of infringing DNA-related products and 15.5% for sales of infringing RNA-related products. (D.I. 496 at 14). Defendants ask the Court to defer its decision on whether to award an ongoing royalty until after appeal, or, in the case that the Court decides to award a royalty to do so at the rate applied by the jury. (D.I. 501 at 18).

It is within the Court's equitable discretion to determine whether an ongoing royalty need be imposed. *See Paice LLC v. Toyota Motor Corp.*, 504 F.3d 1293, 1314-15 (Fed. Cir. 2007). Although an ongoing royalty is not automatic, "the Federal Circuit has indicated that a prevailing patentee should receive compensation for any continuing infringement." *Apple, Inc. v. Samsung Elecs. Co.*, No. 12-CV-00630-LHK, 2014 WL 6687122, at *7 (N.D. Cal. Nov. 25, 2014) (citing *Telcordia Techs., Inc. v. Cisco Sys., Inc.*, 612 F.3d 1365, 1379 (Fed.Cir.2010)). Defendants dedicate just four words to arguing that an ongoing royalty would be inappropriate in this case. (*See* D.I. 501 at 18 ("[I]t should be denied.")). Rather, Defendants argue that the decision should be deferred until after any appeal is concluded. The Federal Circuit, however, has indicated that deferral of such a decision would render the judgment not "final" and thus not fit for appeal. *Warsaw Orthopedic, Inc. v. NuVasive, Inc.*, 515 F. App'x 882, 882 (Fed. Cir. 2012) ("[E]ven assuming Rule 54(b) would give this court jurisdiction over a claim that is 'final except for an accounting' within the meaning of 28 U.S.C. § 1292(c)(2), the case is not 'final' because the district court has not yet determined ongoing royalties. An ongoing royalty is not the same as an accounting for damages."); *Arctic Cat Inc. v. Bombardier Recreational Prod. Inc.*, No. 2016-2556, 2016 WL 11726241, at *1 (Fed. Cir. 2016) ("[T]his case is not 'final' for purposes of section

Tecan Ex. 2207, Page 143 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

1292(c)(2) until the district court renders a final determination on the ongoing royalties issue."). Therefore, the Court will not defer the decision. Defendants have provided no substantive reasons for why an ongoing royalty would be inappropriate in this case, thus the Court will use its equitable discretion to grant an ongoing royalty.

In determining the ongoing royalty rate, the Court must consider: (i) the "change in the parties' bargaining positions, and the resulting change in economic circumstances, resulting from the determination of liability,"; (ii) "changed economic circumstances, such as changes related to the market for the patented products,"; and (iii) any other "post-verdict factor" that would impact "what a hypothetical negotiation would look like after the prior infringement verdict." *XY, LLC v. Trans Ova Genetics, L.C.*, 890 F.3d 1282, 1297 (Fed. Cir. 2018). "Generally, the jury's damages award is a starting point for evaluating ongoing royalties." *Apple, Inc.*, 2014 WL 6687122, at *14 (citing *Bard Peripheral Vascular, Inc. v. W.L. Gore & Assocs., Inc.*, 670 F.3d 1171, 1193 (Fed. Cir. 2012), *vacated on other grounds*, 467 F. App'x 747. In addition, the Court may consider the *Georgia-Pacific* factors. *Arctic Cat Inc. v. Bombardier Recreational Prods. Inc.*, 876 F.3d 1350, 1370 (Fed. Cir. 2017) (citing *Georgia–Pac. Corp. v. U.S. Plywood Corp.*, 318 F. Supp. 1116, 1120 (S.D.N.Y. 1970)).

The jury awarded Plaintiffs a royalty of 7% for sales of DNA-related products in the United States and used 7% to calculate damages for the infringement that led to sale of products abroad. With respect to the DNA-related products, Plaintiffs argue that, post-verdict, the royalty should be enhanced because "(1) the significant value of additional consideration paid by Archer to MGH . . .; (2) the vast difference in the stages of the technology . . .; and (3) the jury's finding that Qiagen willfully infringed." (D.I. 496 at 16-17). In addition, Plaintiffs argue that the following *Georgia-Pacific* factors weigh in favor of enhancement: "Archer's unwillingness to license the

Infringed Patents (Factor 4); the parties' post-verdict commercial relationship (Factor 5); the profitability and commercial success of the products using the patented technology (Factors 8, 11); the benefits of the Infringed patents (Factors 9, 10); and the contribution of the patented features to the overall product (Factor 13)." (*Id.*).

The Court is not persuaded that the factors Plaintiffs point to indicate that they would be able to secure a higher royalty rate than the rate determined by the jury. First, the jury considered all of the factors noted by Plaintiffs in making its royalty determination. (*See* D.I. 462 at 49, D.I. 496, Ex. E ⁋⁋ 28-32). At trial, Plaintiffs' expert argued that these exact factors were why the jury should enhance the 4.5% rate (provided as a starting point based on the Archer/MGH license) to 7%. (*See* D.I. 496, Ex. E ⁋⁋ 28-32.). The jury awarded Plaintiffs 7% and thus appears to have taken these factors into account. Although Plaintiff would be in a stronger bargaining position post-verdict[11], Plaintiffs point to no other circumstances that indicate they would be able to negotiate a rate higher than the already-enhanced rate determined by the jury.

With respect to the RNA-related products, Plaintiffs argue that the jury's award of a 7% royalty rate should be enhanced to 15.5% given the jury's award of lost profits. (*See* D.I. 496 at 17-18). Plaintiffs, however, cite to no law that indicates an award of lost profits requires the Court to enhance the royalty rate. Plaintiffs point to no other factors other than the lost profits award to

---

[11] The Federal Circuit has stated that in determining an ongoing royalty rate, "[o]nce a judgment of validity and infringement has been entered . . . the calculus is markedly different [from the jury's calculus] because different economic factors are involved." *Amado v. Microsoft Corp.*, 517 F.3d 1353, 1362 (Fed. Cir. 2008). As other courts have noted, however, the logic behind this view in unclear given that the jury is required to award a rate negotiated by willing licensors and licensees who considered the patent(s) valid and infringed. *See, e.g.*, *Ventura Ltd. v. GlaxoSmithKline LLC*, No. CV 16-638-RGA, 2019 WL 4346502, at *7 n. 10 (D. Del. Sept. 12, 2019); *see also Apple, Inc. v. Samsung Elecs. Co.*, No. 12-CV-00630-LHK, 2014 WL 6687122, at *2 (N.D. Cal. Nov. 25, 2014). ‼

support their contention that the rate should be higher than 7%. Therefore, the Court sees no reason to disturb the jury's finding.

The Court will thus grant Plaintiffs' request for ongoing royalties at a 7% rate for both RNA- and DNA-related products.

3. Enhanced Damages

Plaintiffs request that the Court award enhanced damages under 35 U.S.C. § 284. (D.I. 496 at 20). Whether to award enhanced damages is committed to the Court's discretion. *Halo Elecs., Inc. v. Pulse Elecs., Inc.*, 579 U.S. 93, 103 (2016). Enhanced damages are "designed as a 'punitive' or 'vindictive' sanction" that is "generally reserved for egregious cases of culpable behavior." *Id.* at 103-04. "The sort of conduct warranting enhanced damages has been variously described in our cases as willful, wanton, malicious, bad-faith, deliberate, consciously wrongful, flagrant, or – indeed – characteristic of a pirate." *Id.* Enhanced damages should not be awarded in "garden-variety cases." *Id.* at 109. "A jury's finding of willful infringement is a prerequisite to enhancement of damages but is not by itself sufficient." *Nox Med. Ehf v. Natus Neurology Inc.*, No. 1:15-CV-00709-RGA, 2018 WL 6427686, at *1 (D. Del. Dec. 7, 2018) (citing *Halo*, 579 U.S. at 103-04).

As an initial matter, the Court does not find that this case involves conduct suggesting willful, wanton, malicious or other nefarious behavior on the part of Defendants. Plaintiffs sued Defendants on the '810 patent just one day after it issued in 2018 and they sued Defendants on the '597 patent after this litigation had begun. (*See* D.I. 1, D.I. 130; *see also* Tr. at 692:1-19). This was a hard-fought litigation and although Plaintiffs ultimately prevailed before the jury, the issues were close. And indeed, Defendants prevailed at summary judgment as to literal infringement of the '810 patent. (*See* D.I. 447).

That being said, Plaintiffs argue that applying the factors articulated in *Read Corp. v. Portec, Inc.*, 970 F.2d 816 (Fed. Cir. 1992) favors an award of enhanced damages in the present case. Although not required, courts may turn to the *Read* factors to guide their analysis. *Presidio Components, Inc. v. Am. Tech. Ceramics Corp.*, 875 F.3d 1369, 1382 (Fed. Cir. 2017). The *Read* factors are: "(1) whether the infringer deliberately copied the ideas or design of another; (2) whether the infringer, when he knew of the other's patent protection, investigated the scope of the patent and formed a good-faith belief that it was invalid or that it was not infringed"; (3) "the infringer's behavior as a party to the litigation"; (4) "defendant's size and financial condition"; (5) "closeness of the case"; (6) "duration of defendant's misconduct"; (7) "remedial action by the defendant"; (8) "defendant's motivation for harm"; and (9) "whether defendant attempted to conceal its misconduct." *Liquid Dynamics Corp. v. Vaughan Co.*, 449 F.3d 1209,1225 (Fed. Cir. 2006) (citing *Read*, 970 F.2d at 826-27).

Plaintiffs primarily rely on the first factor. That is, "Defendant had direct access to, and deliberately copied, Plaintiffs groundbreaking technology." (D.I. 496 at 21). The evidence Plaintiffs point to, however, is circumstantial at best and merely shows that Defendants had the opportunity to copy certain technology. (*See. id.* at 22-24 (citing evidence that Defendants had access to and tested Plaintiffs' technology)). Therefore, this factor is, at most, neutral.

Plaintiffs argue that the second factor weighs in favor of awarding enhanced damages because Qiagen knew of the patents and possible infringement yet "ignored these red flags." (*See* D.I. 496 at 24-25). Defendants counter by pointing to testimony that those who reviewed the patents "did not believe [they] were infringed." (D.I. 501 at 27). In the face of conflicting evidence, the Court finds this factor neutral.

Tecan Ex. 2207, Page 147 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

Regarding factors three, six, seven, eight and nine, Plaintiffs fail to cite any evidence of "egregious conduct" on the part of Defendants that rises to the level of "malicious" behavior required to support an award of enhanced damages. These factors thus weigh against Plaintiffs.

Turning to the fourth factor, Plaintiffs note that "[t]rebling the jury's award is less than 0.1% of Qiagen's 2020 net sales of $1.87 billion." (D.I. 496 at 27). This factor thus weighs in favor of enhancing damages. *See nCUBE Corp. v. SeaChange Int'l, Inc.*, 313 F. Supp. 2d 361, 390 (D. Del. 2004), *aff'd*, 436 F.3d 1317 (Fed. Cir. 2006) (finding this factor to weigh in favor of enhanced damages because defendant would "not be materially impacted by an award of enhanced damages"). !

Turning to the fifth factor, as Defendants point out, this was a close case. Defendants successfully argued for summary judgment that the '810 patent was not literally infringed. (*See* D.I. 447). In addition, Defendants put forth valid arguments regarding other infringement and invalidity theories at trial. This factor weighs heavily against Plaintiffs.

In sum, the fourth factor (defendant's size and financial condition) is the only one that clearly weighs in favor of granting enhanced damages. The fifth factor strongly weighs against enhancing the damages and the others are largely neutral or unsupported by evidence. Given the lack of evidence of any egregious conduct on the part of Defendants, the Court is not persuaded that this is a case that warrants enhanced damages.

4. Supplemental Damages

Plaintiffs request supplemental damages for sales that occurred before the verdict but were not reflected in the last accounting. (D.I. 496 at 18). At trial, the parties presented damages calculations through December 31, 2020. During the pretrial conference, the Court ruled that an accounting would be performed for damages incurred between January 1, 2021 and the date of the jury's verdict. (D.I. 520 (Aug 16, 2021 Pretrial Conference Tr.) at 90). Defendants produced the

required financial data on September 3, 2021 and October 22, 2021. (*See* D.I. 496 at 19). Applying the jury award to Qiagen's data, Plaintiffs request the following amounts: $308,395 as lost profits for sales of infringing RNA-related products in the United States, $194,690[12] as a reasonable royalty on sales of infringing DNA-related products in the United States and $722,928 as a reasonable royalty for sales abroad. (*Id.*)

Defendants do not dispute this request nor the amounts. (*See* D.I. 501 at 21). Therefore, the Court will grant Plaintiffs' request for supplemental damages.

5. Pre- and Post-Judgment Interest

Plaintiffs request prejudgment interest of $225,296 for all damages through the date of judgment using the U.S. prime rate. (D.I. 496 at 19). Defendants do not contest the award of prejudgment interest or Plaintiffs' calculation. Rather, Defendants' request that the Court apply the Treasury bill rate rather than prime rate. (*See* D.I. 501 at 21). For the reasons set forth below, the Court will grant Plaintiffs' request and apply the prime rate.

Prejudgment interest is awarded to restore a plaintiff to the position it would have been in had there been no wrongdoing. *See General Motors Corp. v. Devex Corp.*, 461 U.S. 648, 655-56 (1983); *see also Delaware River & Bay Auth. v. Kopacz*, 584 F.3d 622, 634 (3d Cir. 2009) (prejudgment interest awards "must be compensatory rather than punitive"). "The matter of prejudgment interest is left to the discretion of the district court." *Taxman v. Bd. of Educ. of Twp. of Piscataway*, 91 F.3d 1547, 1566 (3d Cir. 1996). This broad discretion, of course, extends to a determination of the appropriate interest rate to apply. *See, e.g., Sun Ship, Inc. v. Matson Navigation Co.*, 785 F.2d 59, 63 (3d Cir. 1986) ("In federal question cases, the rate of prejudgment

[12] Plaintiffs use what parties agree was the jury's intended royalty of $752,006. (D.I. 496, Ex. E ¶ 15).

interest is committed to the discretion of the district court."); *Uniroyal, Inc. v. Rudkin-Wiley Corp.*, 939 F.2d 1540, 1545 (Fed. Cir. 1991).

The Court will assess prejudgment interest compounded quarterly at the prime rate. *See, e.g.*, *In re Frescati Shipping Co.*, 886 F.3d 291, 315 (3d Cir. 2018) (district court within its discretion to award prejudgment interest either at the prime rate or post-judgment rate prescribed by 28 U.S.C. § 1961(a)); *see also Taxman*, 91 F.3d at 1566 ("The adjusted prime rate, established periodically by the Secretary of the Treasury and codified in 26 U.S.C. § 6621, has been used regularly by district courts to calculate prejudgment interest."); *see also Amgen Inc. v. Hospira, Inc.*, 336 F. Supp. 3d 333, 364 (D. Del. 2018) (for patent damages, awarding prejudgment interest at the prime rate compounded quarterly), *aff'd*, 944 F.3d 1327 (Fed. Cir. 2019). The prime rate is by far the most common practice in the District of Delaware. *See, e.g.*, *Bayer v. Baxalta*, No. 16-1122-RGA, 2019 WL 4016235, at *7 (D. Del. Aug. 26, 2019); *Idenix Pharm. LLC v. Gilead Scis., Inc.*, 271 F. Supp. 3d 694, 705 (D. Del. 2017); *Comcast IP Holdings I, LLC v. Sprint Commc'ns Co.*, No. 12-0205-RGA, 2015 WL 4730899, at *10 (D. Del. Aug. 10, 2015); *Ateliers de la Haute-Garonne v. Broetje Automation-USA Inc.*, 85 F. Supp. 3d 768, 783 (D. Del. 2015); *Dow Chem. Co. v. Nova Chems. Corp.*, No. 05-737-LPS, 2014 WL 1285508, at *11 (D. Del. Mar. 28, 2014); *XpertUniverse, Inc. v. Cisco Sys., Inc.*, No. 09-157-RGA, 2013 WL 6118447, at *11 (D. Del. Nov. 20, 2013).

Defendants assert that the Court should apply the Treasury bill rate because Plaintiffs "have not put forward any evidence of its borrowing rates at the relevant time." (D.I. 501 at 21-22). It is not necessary, however, for Plaintiffs to "demonstrate that it borrowed at the prime rate in order to be entitled to prejudgment interest at that rate." *Baxalta*, 2019 WL 4016235, at *7 (quoting *Uniroyal*, 939 F.2d at 1545); *see also In re Frescati Shipping Co.*, 886 F.3d at 314 ("[H]ad the

District Court chosen to use the prime rate, it would not have abused its discretion even without extensive proof of borrowing costs."); *Idenix*, 271 F. Supp. 3d at 705 (citing *XpertUniverse*, 2013 WL 6118447, at *11). Therefore, the Court is not persuaded that the Treasury bill rate should be applied in this case.

In sum, the judgment shall be amended to include prejudgment interest compounded quarterly at the prime rate for the $841,756 in lost profits in connection with Defendants' RNA-related products sold in the Unites States, $752,006[13] in royalties for sales of DNA-related products in the Unites States and $2,240,303 in royalties for sales abroad.

In addition, Plaintiffs request post-judgment interest in the amount of $10.14 per day.[14] Post-judgment interest is mandatory for damages awarded in civil cases. *See* 28 U.S.C. § 1961(a) ("Interest shall be allowed on any money judgment in a civil case recovered in a district court."). Consistent with § 1961(a), the rate proposed by Plaintiffs is the weekly average one-year constant maturity Treasury yield for the week preceding entry of judgment (*i.e.*, 0.07%). (*See* D.I. 496 at 20). Defendants do not dispute that Plaintiffs are entitled to post-judgment interest or that the rate used by Plaintiffs is proper. Therefore, Plaintiffs shall be awarded post-judgment interest calculated using that rate and starting from the date judgment on the jury verdict was entered (*i.e.*, August 27, 2021). Post-judgment interest shall be awarded for the entire amount included in the

---

13 Plaintiffs use what both parties agree was the jury's intended royalty of $752,006. (*See* D.I. 496, Ex. E ¶¶18-20 (citing Schedule 4.1D)). The Court will grant Defendants' remittitur regarding this award and thus uses the same amount in its award of pre- and post-judgment interest.

14 Plaintiffs apply the weekly average one-year constant maturity Treasury yield for the week preceding entry of judgment, which was 0.07%. (D.I. 496 at 20). This corresponds to a daily interest rate of 0.00019% and results in $4.21 on damages for sales of infringing products sold to customers for use in the United States and $5.93 on damages for sales of infringing products sold to customers for use abroad. (*Id.*).

judgment, including prejudgment interest. *See Skretvedt v. E.I. DuPont de Nemours*, 372 F.3d 193, 217 (3d Cir. 2004). To be clear, however, post-judgment interest on the prejudgment interest award does not begin to accrue until the amended judgment quantifying the prejudgment interest is entered. *See Travelers Cas. & Sur. Co. v. Ins. Co. of N. Am.*, 609 F.3d 143, 175 (3d Cir. 2010) ("[P]ost-judgment interest on Travelers' award of prejudgment interest did not begin to run until the December 5, 2007 order was entered quantifying the amount in prejudgment interest owed to Travelers.").

In sum, the Court will grant Plaintiffs' request for pre- and post-judgment interest.

IV. CONCLUSION

For the foregoing reasons, Defendants' renewed motion for judgment as a matter of law or in the alternative, motion for a new trial or altered judgment (D.I. 495) is DENIED, Defendants' motion for remittitur is GRANTED, Plaintiffs' motion for enhanced damages (D.I. 494) is DENIED, motion for an injunction is DENIED with leave to renew after an evidentiary hearing, motion for ongoing royalty is GRANTED-IN-PART, motion for supplemental damages is GRANTED and motion for pre- and post-judgment interest is GRANTED. An appropriate order will follow.

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

ARCHERDX, LLC and THE GENERAL HOSPITAL CORPORATION d/b/a MASSACHUSETTS GENERAL HOSPITAL,	)	
Plaintiffs,	)	
v.	)	C.A. No. 18-1019 (MN)
QIAGEN SCIENCES, LLC, QIAGEN LLC f/k/a QIAGEN, INC., QIAGEN BEVERLY, LLC F/K/A QIAGEN BEVERLY, INC., QIAGEN GAITHERSBURG, LLC f/k/a QIAGEN GAITHERSBURG, INC., QIAGEN GMBH, QIAGEN N.V. and JONATHAN ARNOLD,	)	
Defendants.	)	

ORDER

At Wilmington this 30th day of September 2022:

For the reasons set forth in the Memorandum Opinion issued this date, IT IS HEREBY ORDERED that:

1. Defendants' motion for judgment as a matter of law, or, in the alternative, a new trial (D.I. 495) is DENIED;

2. Defendants' motion for remittitur (D.I. 495) is GRANTED;

3. Plaintiffs' motion for injunction (D.I. 494) is DENIED with leave to renew after an evidentiary hearing;

4. Plaintiffs' motion for ongoing royalty (D.I. 494) is GRANTED-IN-PART;

5. Plaintiffs' motion for enhanced damages (D.I. 494) is DENIED; and

6. Plaintiffs' motion for supplemental damages and pre- and post-judgment interest (D.I. 494) is GRANTED.

IT IS FURTHER ORDERED that, on or before October 10, 2022, the parties shall meet and confer and submit a JOINT status report on the requested evidentiary hearing pertaining to injunctive relief.

Maryellen Noreika

The Honorable Maryellen Noreika
United States District Judge

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

ARCHERDX, LLC and THE GENERAL HOSPITAL CORPORATION d/b/a MASSACHUSETTS GENERAL HOSPITAL,	)	
Plaintiffs,	)	
v.	)	C.A. No. 18-1019-MN
QIAGEN SCIENCES, LLC, QIAGEN LLC f/k/a QIAGEN, INC., QIAGEN BEVERLY, LLC f/k/a QIAGEN BEVERLY, INC., QIAGEN GAITHERSBURG, LLC f/k/a QIAGEN GAITHERSBURG, INC., QIAGEN GMBH, QIAGEN N.V. and JONATHAN ARNOLD,	)	
Defendants.	)	

FINAL JUDGMENT

This 8th day of August 2023, the Court having held a jury trial and the jury having rendered a unanimous verdict on August 27, 2021 (*see* D.I. 465), pursuant to Federal Rule of Civil Procedure 58, IT IS HEREBY ORDERED that:

1. Judgment is entered in favor of Plaintiffs ArcherDX, LLC and The General Hospital Corporation (collectively, "Plaintiffs") that Defendants Qiagen Sciences, LLC, Qiagen LLC, Qiagen Beverly, LLC, Qiagen Gaithersburg, LLC, Qiagen GmbH, and Qiagen N.V. (collectively, "Defendants") directly infringed claims 16, 17, and 19 of U.S. Patent No.

10,017,810 (the "'810 Patent")[1] and claims 1, 5, and 19 of U.S. Patent No. 10,450,597 (the "'597 Patent")[2] (collectively, the "Asserted Claims").

2. Judgment is entered in favor of Plaintiffs that Defendants induced infringement of the Asserted Claims.

3. Judgment is entered in favor of Plaintiffs that Defendants contributed to infringement of the Asserted Claims.

4. Judgment is entered in favor of Plaintiffs that Defendants have not shown invalidity for obviousness, lack of written description, or indefiniteness of the Asserted Claims of the '810 Patent, nor have Defendants shown invalidity for anticipation, lack of written description, or indefiniteness of the Asserted Claims of the '597 Patent.

5. Judgment is entered in favor of Plaintiffs and against Defendants for damages in the amount of $841,756.00 caused by Defendants' sale of RNA-related accused products to customers within the United States and in the amount of $752,006.00 caused by Defendants' sale of other accused products to customers within the United States.

6. Judgment is entered in favor of Plaintiffs and against Defendants for damages in the amount of $2,240,303.00 caused by Defendants' sale of accused products to customers outside the United States.

7. Judgment is entered in favor of Plaintiffs that Defendants' infringement was willful.

---

[1] Defendants' accused products relating to the '810 Patent are: QIAseq Targeted DNA Panels, QIAseq Targeted RNAscan Panels, QIAseq Immune Repertoire RNA Library Kits, and QIAseq Index Kits for the Illumina platform.

[2] Defendants' accused products relating to the '597 Patent are: QIAseq Targeted DNA Panels, QIAseq Targeted RNAscan Panels, QIAseq Immune Repertoire RNA Library Kits, QIAseq Index Kits for the Illumina and Ion Torrent platforms, and GeneRead QIAact Kits.

8. Judgment is entered that Defendants' willful infringement does not warrant enhanced damages.

9. Judgment is entered in favor of Plaintiffs and against Defendants and Jonathan Arnold in the amount of $100.00 on Archer's claim of breach of fiduciary duty against Jonathan Arnold (Count VII) and its claim of aiding and abetting breach of fiduciary duty as to the QIAGEN defendants (Count VIII).[3]

10. Plaintiffs' claims of trade secret misappropriation (Counts I, IV-VI), false advertising (Count III), deceptive trade practices (Count IX), and tortious interference (Count X) are dismissed as moot.

11. Defendants are ordered to pay Plaintiffs an ongoing royalty of 7% on all sales of QIAseq Targeted DNA Panels, QIAseq Targeted RNAscan Panels, QIAseq Immune Repertoire RNA Library Kits, QIAseq Index Kits for the Illumina and Ion Torrent platforms, and GeneRead QIAact Kits (the "Ongoing Royalty Products") from August 27, 2021 through the date the '597 Patent expires. Payment of the ongoing royalty shall be made 90 days from the entry of this Final Judgment, with payments made on a quarterly basis thereafter within 60 days of the end of each calendar quarter (March 31, June 30, September 30, and December 31) except as otherwise stayed through the posting of a bond on appeal. Thirty days before each payment, Defendants shall provide to Plaintiffs quarterly sales

[3] Judgment on Counts VII and VIII is pursuant to a Rule 68 Offer of Judgment that provided that the offer "shall not be deemed an admission (a) of liability by Defendants; (b) that Counts VII and VIII have any merit; or (c) that Archer has suffered any damages" and that the offer "shall satisfy all of Archer's claims with respect to Counts VII and VIII, including but not limited to damages of any kind, prejudgment interest, injunctive relief, costs, expenses, attorneys' fees, and any other amounts to which Archer might claim to be entitled in connection with its Counts VII and VIII in this litigation." D.I. 423-1 at 3 of 5.

information for each Ongoing Royalty Product, including the number of infringing units and corresponding revenue.

12. Judgment is entered in favor of Plaintiffs and against Defendants for supplemental damages in the amount of $308,395.00 caused by Defendants' sale of RNA-related accused products to customers within the United States, $194,690.00 caused by Defendants' sale of other accused products to customers within the United States, and $722,928.00 caused by Defendants' sale of accused products to customers outside the United States.

13. Judgement is entered in favor of Plaintiffs and against Defendants for pre-judgment interest in the amount of $98,564.00 for damages caused by Defendants' sale of accused products to customers within the United States and $126,732 for damages caused by Defendants' sale of accused products to customers outside the United States.

14. Post-judgment interest on the lost profit damages, royalty damages, and supplemental damages (which together total $5,060,078.00) began accruing on August 27, 2021, the date the jury verdict was entered (D.I. 465). Judgment is entered in favor of Plaintiffs and against Defendants for post-judgment interest, at the rate of 0.07% compounded annually, on the $5,060,078.00 award from August 27, 2021 until the judgment is paid.

15. Judgment is entered in favor of Plaintiffs and against Defendants for post-judgment interest on the pre-judgment interest award of $225,296.00, at the rate of 0.07% compounded annually from the date this Judgment is entered until the judgment is paid.

16. Judgment is entered in favor of Defendants and against Plaintiffs with respect to Plaintiffs' request for an injunction.

IT IS FURTHER ORDERED that the deadline for any party to move for costs and attorneys' fees (including under 35 U.S.C. § 285) is extended fourteen (14) days after the time for appeal has

expired or, if an appeal has been timely filed, within fourteen (14) days after issuance of the mandate from the appellate court, and no party shall file any such motion before that time.

Maryellen Noreika

The Honorable Maryellen Noreika
United States District Judge

US010017810B2

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

(10) **Patent No.: US 10,017,810 B2**
(45) **Date of Patent: *Jul. 10, 2018**

(54) **METHODS FOR DETERMINING A NUCLEOTIDE SEQUENCE CONTIGUOUS TO A KNOWN TARGET NUCLEOTIDE SEQUENCE**

(71) Applicant: **THE GENERAL HOSPITAL CORPORATION**, Boston, MA (US)

(72) Inventors: **Anthony John Iafrate**, Newton, MA (US); **Long Phi Le**, Boston, MA (US); **Zongli Zheng**, Boston, MA (US)

(73) Assignee: **THE GENERAL HOSPITAL CORPORATION**, Boston, 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.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **15/269,448**

(22) Filed: **Sep. 19, 2016**

(65) **Prior Publication Data**

US 2017/0226575 A1 Aug. 10, 2017

Related U.S. Application Data

(63) Continuation of application No. 13/793,564, filed on Mar. 11, 2013, now Pat. No. 9,487,828.

(60) Provisional application No. 61/645,364, filed on May 10, 2012, provisional application No. 61/679,302, filed on Aug. 3, 2012.

(51) **Int. Cl.**

C12Q 1/68	(2018.01)
C12Q 1/6855	(2018.01)
C12Q 1/6874	(2018.01)
C12Q 1/6806	(2018.01)

(52) **U.S. Cl.**

CPC ***C12Q 1/6855*** (2013.01); ***C12Q 1/6806*** (2013.01); ***C12Q 1/6874*** (2013.01)

(58) **Field of Classification Search**

None

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,868,104	A	9/1989	Kurn et al.
6,087,101	A	7/2000	Gruelich et al.
6,172,214	B1	1/2001	Brenner
6,576,448	B2	6/2003	Weissman et al.
6,632,611	B2	10/2003	Su et al.
6,833,246	B2	12/2004	Balasubramanian
7,214,490	B2	5/2007	Su et al.
7,244,559	B2	7/2007	Rothberg et al.
7,273,730	B2	9/2007	Du Breuil
7,824,856	B2	11/2010	Montforte
9,487,828	B2 *	11/2016	Iafrate C12Q 1/6874
2002/0086317	A1	7/2002	Nagayama
2003/0143553	A1	7/2003	Sommer
2009/0203085	A1	8/2009	Kurn et al.
2010/0286143	A1	11/2010	Dias-Santagata et al.
2012/0122701	A1	5/2012	Ryan et al.
2013/0005585	A1	1/2013	Anderson et al.
2013/0231253	A1	9/2013	Amorese et al.
2013/0303461	A1	11/2013	Iafrate et al.
2015/0011396	A1	1/2015	Schroeder et al.
2015/0140553	A1	5/2015	Cushing et al.

FOREIGN PATENT DOCUMENTS

AU	2005203617	A1	9/2005
WO	97/23646	A1	7/1997
WO	97/23647	A1	7/1997
WO	98/28443	A1	7/1998
WO	99/42618	A1	8/1999
WO	00/43544	A1	7/2000
WO	00/70095	A2	11/2000
WO	01/12859	A2	2/2001
WO	01/20035	A2	3/2001
WO	02/00938	A2	1/2002
WO	02/29117	A3	4/2002
WO	02/48402	A2	6/2002
WO	02/072772	A2	9/2002
WO	03/078645	A2	9/2003
WO	03/083435	A2	10/2003
WO	2004/011665	A2	2/2004
WO	2004/092418	A2	10/2004
WO	2005/065321	A2	7/2005
WO	2007/030759	A2	3/2007
WO	2007/057652	A1	5/2007
WO	2007/136717	A1	11/2007
WO	2008/005459	A2	1/2008
WO	2008/093098	A2	8/2008
WO	2009/102878	A2	8/2009
WO	2009/102896	A2	8/2009
WO	2009/117698	A2	9/2009
WO	2009/133466	A2	11/2009
WO	2009/148617	A2	12/2009
WO	2010/077288	A2	7/2010
WO	2010/083046	A2	7/2010

(Continued)

OTHER PUBLICATIONS

Gould et al., "Analysis of sequence variation among smeDEF multi drug efflux pump genes and flanking DNA from defined 16S rRNA subgroups of clinical Stenotrophomonas maltophilia isolates." Journal of Antimicrobial chemotherapy 54(2):348-353 (2004).

Liu et al., "High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences." Biotechniques Rapid Dispatches 43(5):649-656 (2007).

Wang et al. "Fusion primer and nested integrated PCR (FPNI-PCR): a new high-efficiency strategy for rapid chromosome walking or flanking sequence cloning." BMC Biotechnology 11(1):109 (2011).

(Continued)

Primary Examiner Angela M. Bertagna

(74) *Attorney, Agent, or Firm* — Nixon Peabody LLP; David S. Resnick; Nicole D. Kling

(57) **ABSTRACT**

The technology described herein is directed to methods of determining oligonucleotide sequences, e.g. by enriching target sequences prior to sequencing the sequences.

20 Claims, 9 Drawing Sheets

ArcherDX v. Qiagen 18-1019-MN
Joint Trial Exhibit
JX-001

(56) **References Cited**

FOREIGN PATENT DOCUMENTS

WO	2011/019964	A1	2/2011
WO	2011/032053	A1	3/2011
WO	2011/053987	A1	5/2011
WO	2011/156529	A2	12/2011
WO	2012/003374	A2	1/2012
WO	2012/040387	A1	3/2012
WO	2012/044956	A1	4/2012
WO	2012/064739	A2	5/2012
WO	2012/013154	A1	8/2012
WO	2013/059746	A1	4/2013
WO	2013/074833	A1	5/2013
WO	2013/112923	A1	8/2013
WO	2013/191775	A2	12/2013
WO	2014/144092	A1	9/2014
WO	2014/150931	A1	9/2014
WO	2015/073711	A1	5/2015

OTHER PUBLICATIONS

Callaway et al., "A Sobemovirus coat protein gene complements long-distance movement of a coat protein-null Dianthovirus." Virology 330(1):186-195 (2004).

Natsoulis et al., "A Flexible Approach for Highly Multiplexed Candidate Gene Targeted R.esequencing", PLoS One. 6(6):e21088(2011).

Natsoulis et al., "Identification of Insertion Deletion Mutations from Deep Targeted Resequencing", J Data Mining Genomics Proteomics 4(3) (2013).

Newburger et al., "The Human OligoGenome Resource: a database of oligonucleotide capture probes for resequencing target regions across the human genome", Nucleic Acids Research 40(Database issue):D1137-D1143 (2012).

Scheinin et al., "CanGEM: mining gene copy number changes in cancer", Nucleic Acids Research 36(Database issue):D830-835 (2008).

Schiffman et al., "Molecular inversion probes reveal patterns of 9p21 deletion and copy number aberrations in childhood leukemia", Cancer Genet Cytogenet. 193(1):9-18 (2009).

Schiffman et al., "Oncogenic BRAF Mutation with CDKN2A Inactivation is Characteristic of a Subset of Pediatric Malignant Astrocytomas", Cancer Res. 70(2):512-519 (2010).

Shendure et al., "Next-generation DNA sequencing", Nature Biotechnology 26(10):1135-1145 (2008).

Singh et al., "Microarray-based comparison of three amplification methods for nanogram amounts of total RNA", Am J Physiol Cell Physiol. 288(5):C1179-C1189 (2005).

Sinkkonen et al., "Serial Analysis of Gene Expression in tire Chicken Otocyst", JARO 12(6):697-710 (2011).

Tong et al., "Genome-scale identification of conditionally essential genes in *E. coli* by DNA microarrays", Biochemical and Biophysical Research Communications 322(1):347-354 (2004).

Turner et al., "Gene Expression Profiling of RNA Extracted from FFPE Tissues: NuGEN Technologies' Whole-Transcriptome Amplification System", Methods Mol Biol. 724:269-280 (2011).

Xiao et al., "Sequential Amplification of Flanking Sequences by Y-shaped Adaptor Dependent Extension Using Multiple Templates". 33(1):85-90 (2007). Abstract Only.

Yuanxin et al., "T-linker-specific ligation PCT (T-linker PCR): an advanced PCR technique for chromosome walking or for isolation of tagged DNA ends", Nucleic Acids Research 31(12):e68 (2003).

Zhang et al., "Detecting simultaneous changepoints in multiple sequences", Biometrika 97(3):631-645 (2010).

Zhang et al., "Titration-free 454 sequencing using Y adapters", Nature Protocols 6(9):1367-1376 (2011).

Ali et al., "Sequence Analysis of TnphoA Insertion Sites in Vibrio cholera Mutants Defective in Rugose Polysaccharide Production", Infection and Immunity 687:6857-6864 (2000).

Baetens et al., "Applying Massive Parallel Sequencing to Molecular Diagnosis of Martin and Loeys-Dietz Syndromes", Human Mutation 32:1053-1062 (2011).

Bohmer, "Novel application for isothermal nucleic acid sequence-based amplification (NASBA)". Journal of Virological Methods 158(1-2)199-201 (2009).

Chang et al., "Identification of a biomarker panel using a multiplex proximity ligation assay improves accuracy of pancreatic cancer diagnosis". Journal of Translational Medicine 7:105 (2009).

Chen et al., "Allele-specific copy number profiling by next-generation DNA sequencing." Nucleic Acids Research 43(4):e23 (2015).

Chenchik et al., "Full-Length cDNA Cloning and Determination of mRNA 5' and 3' Ends by Amplification of Adaptor-Ligated cDNA", BioTechniques 21(3):526-534 (1996).

Compton, "Nucleic acid sequence-based amplification", Nature 350(6313):91-92 (1991).

Cushing et al., "RVD: a command-line program for ultrasensitive rare single nucleotide variant detection using targeted next-generation DNA resequencing". BMC Research Notes 6:206 (2013).

Dafforn et al., "Linear mRNA amplification from as little as 5 ng total RNA for global gene expression analysis", I BioRechniques 37(5):854-657 (2004).

Dahl et al.: "Multigene amplification and massively parallel sequencing for cancer mutation discovery", PNAS 104 (22):9387-9392 (2007).

Flaherty et al., "Ultrasensitive detection of rare mutations using next-generation targeted resequencing". Nucleic Acids Research 40(1):e2 (2012).

Fredriksson et al., "Multiplex amplification of all coding sequences within 10 cancer genes by Gene-Collector", Nucleic Acids Research 35(7):e47 (2007).

Fredriksson et al., "Multiplexed protein detection by proximity ligation for cancer biomarker validation", Nature Methods 4(4):327-329 (2007).

Grace et al., "Degradable dUMP Outer Primers in Merged Tandem (M/T)-Nested PCR: Low- and Single-Copy DNA Target Amplification". Analytical Biochemistry 263(1):85-92 (1998).

Green et al., "Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma", Blood 121(9):1604-1611 (2013).

Grimes et al., "MendeLIMS: a web-based laboratory information management system for clinical genome sequencing". BMC Bioinformatics 15:290 (2014).

Guled et al., "Array comparative genomic hybridization analysis of olfactory neuroblastoma", Modern Pathology 21 (6):770-778 (2008).

Head et al., "Method for improved Illumina sequencing library preparation using NuGEN Ovation RNA-Seg System". Biotechniques 50(3)1 77-181 (2011).

Hoeijmakers et al., "Linear amplification for deep sequencing", Nature Protocols 6(7)1 026-1036 (2011).

Hoon et al., "Aptamer selection by high-throughput sequencing and informatic analysis". Biotechniques 51(6):413-416 (2011).

Hopmans et al., "A programmable method for massively parallel targeted sequencing". Nucleic Acids Research 42 (10):e88 (2014).

Ji et al., "Data quality in genomics and microarrays". Nat Biotechnol. 24(9):1112-1113 (2006).

Ji, "Improving bioinformatic pipelines for exome variant calling", Genome Medicine 4(1):7 (2012).

Ji et al., "Identification of a novel deletion mutant strain in Saccharomyces cerevisiae that results in a microsatellite instability phenotype". Biodiscovery 1:4 (2012).

Ji et al., "Molecular inversion Probe Analysis of Gene Copy Alterations Reveals Distinct Categories of Colorectal Carcinoma", Cancer Res. 66(16):7910-7919 (2006).

Ji et al., "Molecular inversion probe assay for allelic quantitation". Methods Mol Biol. 556:67-87 (2009).

Kim et al., "Genetic-based biomarkers and next-generation sequencing: the future of personalized care in colorectal cancer", Per Med. 8(3):331-345 (2011).

Kinde et al., "Detection and quantification of rare mutations with massively parallel sequencing". PNAS 108 (23):9530-95355 (2011).

(56) **References Cited**

OTHER PUBLICATIONS

Kurn et al., "Novel Isothermal, Linear Nucleic Acid Amplification Systems for Highly Multiplexed Applications", Clinical Chemistry 51(10):1973-1981 (2005).
Lam et al., "Performance comparison of whole-genome sequencing platforms", Nat Biotechnol. 30(1):78-82 (2011).
Lamant et al., "A New Fusion Gene TPM3-Alk in Anaplastic Large Cell Lymphoma Created by a (1:2)(q25;p23) Translocation", Blood 93(9):3088-3095 (1999).
Lee et al., "Systematic genomic identification of colorectal cancer genes delineating advanced from early clinical stage and metastasis", BMC Medical Genomics 6:54 (2013).
Lin et al., "Reproducibility Probability Score--incorporating measurement variability across laboratories for gene selection", Nature Biotechnology 24(12):1476-1477 (2006).
Lishanski et al., "Branch migration inhibition in PCR-amplified DNA: homogeneous mutation detection". Nucleic Acids Research 28(9):e42 (2000).
MAQC Consortium, "The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements", Nat Biotechnol. 24(9):1151-1161 (2006)
Miotke et al., "Correction to High Sensitivity Detection and Quantitation of DNA Copy Number and Single Nucleotide Variants Single Color Droplet Digital PCR", Anal. Chem, 87:3114 (2015).
Miotke et al., "High Sensitivity Detection and Quantitation of DNA Copy Number and Single Nucleotide Variants with Single Color Droplet Digital PCR". Anal. Chem.86(5):2618-2624 (2014).
Mugasa et al., "Nucleic Acid Sequence-Based Amplification with Oligochromatography for Detection of Trypanosome brucei in Clinical Samples", Journal of Clinical Microbiology 47(3):630-635 (2009).
Muralidharan et al., "A cross-sample statistical model for SNP detection in short-read sequencing data", Nucleic Acids Research 40(1):e5 (2012).
Myllykangas et al., "Classification of human cancers based on DNA copy number amplification modeling", BMC Medical Genomics 1:15 (2008).
Myllykangas et al., "DNA copy number amplification profiling of human neoplasms", Oncogerre 25(55):7324-7332 (2006).
Myllykangas et al., "Efficient targeted resequencing of human germline and cancer genomes by oligonucleotide-selective sequencing", Nat Biotechnol. 29(11):1024-1027 (2011).
Myllykangas et al., "Integrated gene copy number and expression microarray analysis of gastric cancer highlights potential target genes", Int. J. Cancer 123(4):817-825 (2008).
Myllykangas et al., "Manifestation, mechanisms and mysteries of amplifications", Cancer Letters 232 (1):79-89 (2006).
Myllykangas et al., "Novel high-throughput sequencing strategies in genetic diagnostics", Duodecim. 129 (2):141-148 (2013). English Abstract [Review Finnish].
Myllykangas et al., "Specificity, selection and significance of gene amplifications in cancer", Seminars in Cancer Biology 17(1):42-55 (2007).
Myllykangas et al., "Targeted deep resequencing of the human cancer genome using next-generation technologies", Biotechnol Genet Eng Rev. 27:135-158 (2010).
Myllykangas et al., "Targeted sequencing library preparation by genomic DNA circularization", BMC Biotechnology 11:122 (2011).
Nadauld et al., "Quantitative and Sensitive Detection of Cancer Genome Amplifications from Formalin Fixed Paraffin Embedded Tumors with Droplet Digital PCR". Transl Med (Sunnyvale) 2(2)(2012).

* cited by examiner

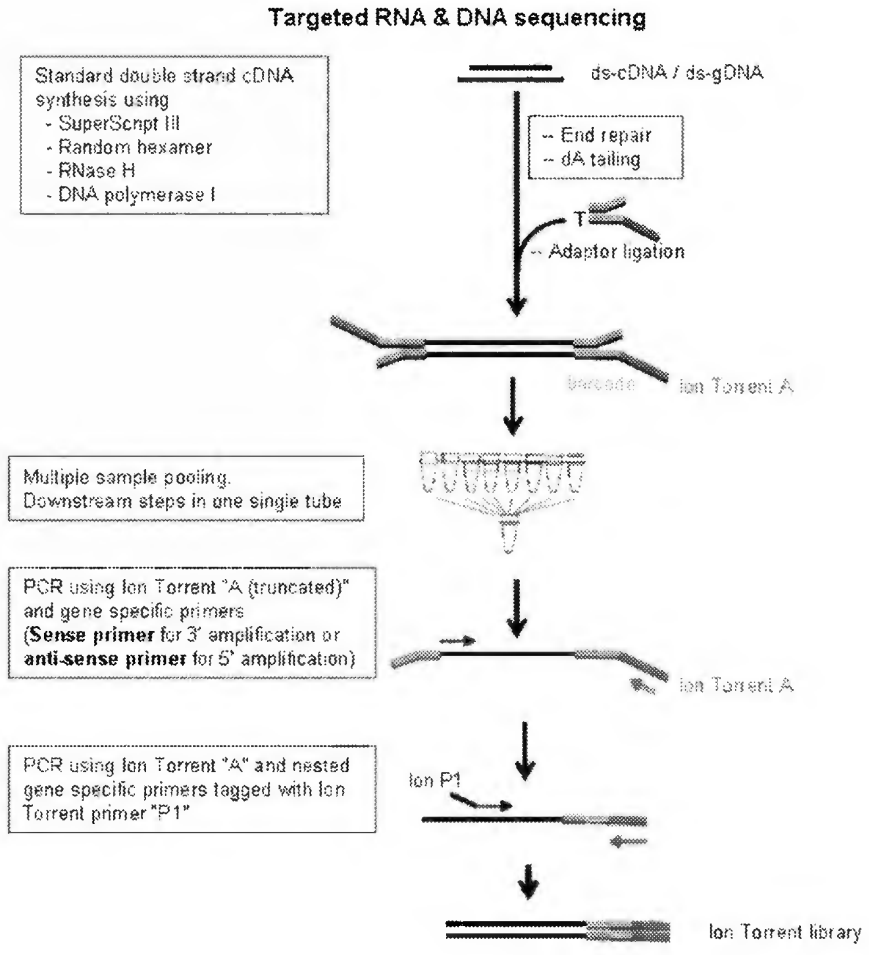


Figure 1

QIA-00689216

JX-001.4

Appx81

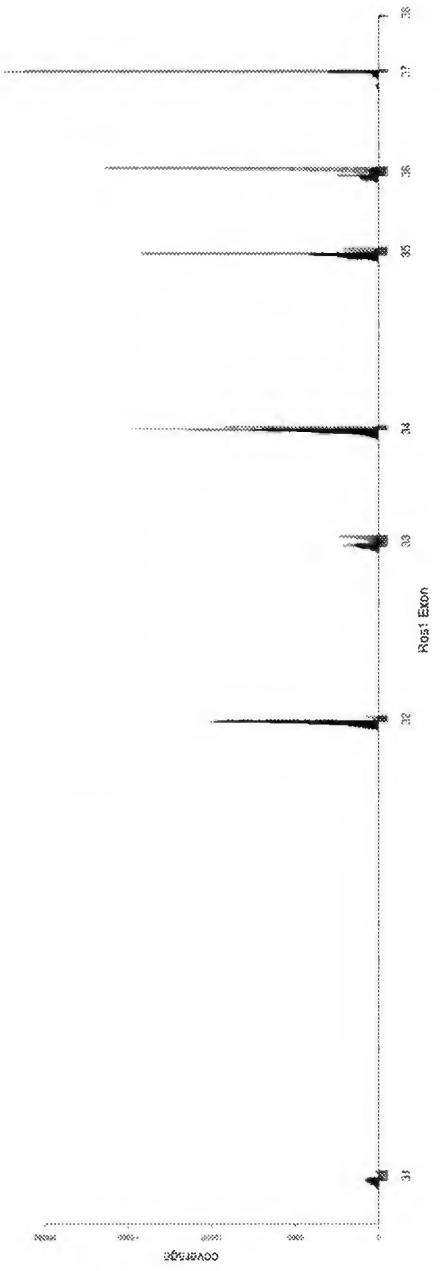


Figure 2

Figure 3A

Figure 3B

QIA-00689218

Appx83

Figure 4A

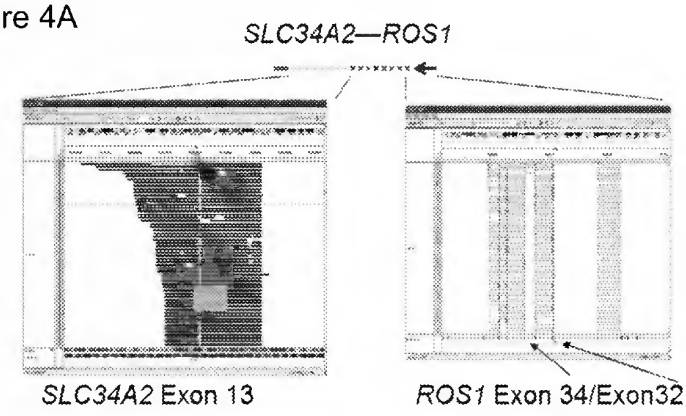


Figure 4B

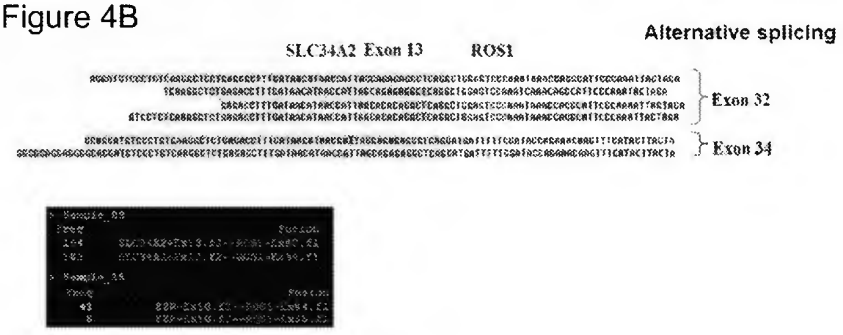


Figure 4C

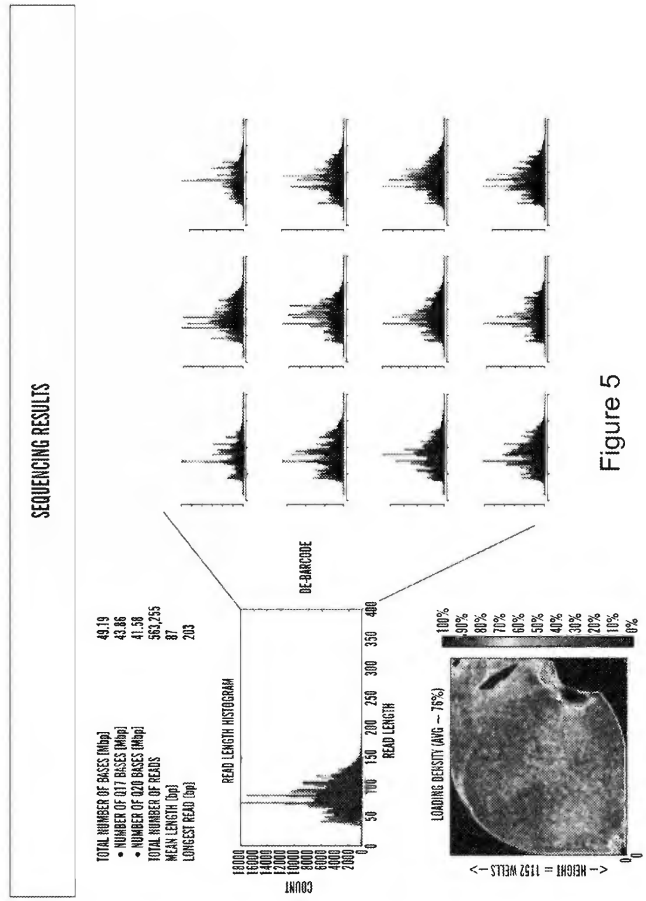


Figure 5

QIA-00689220

Appx85

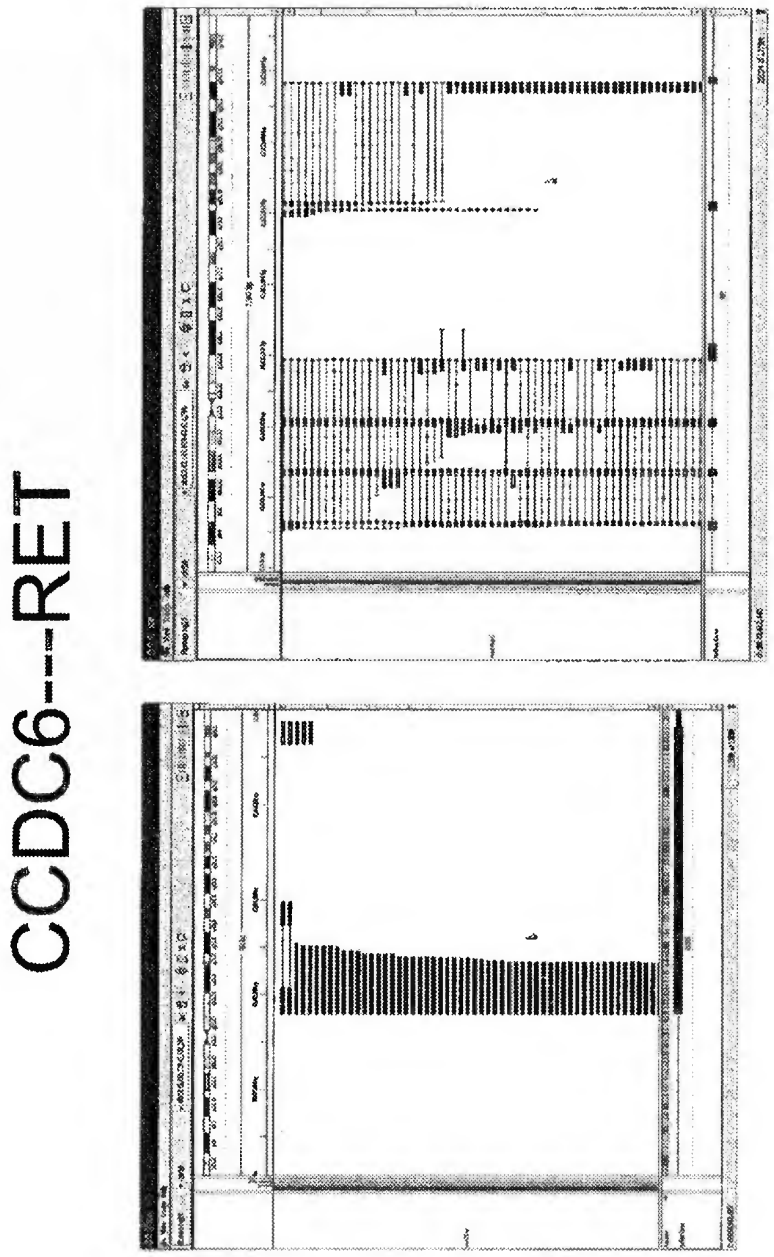


Figure 6

EML4--ALK

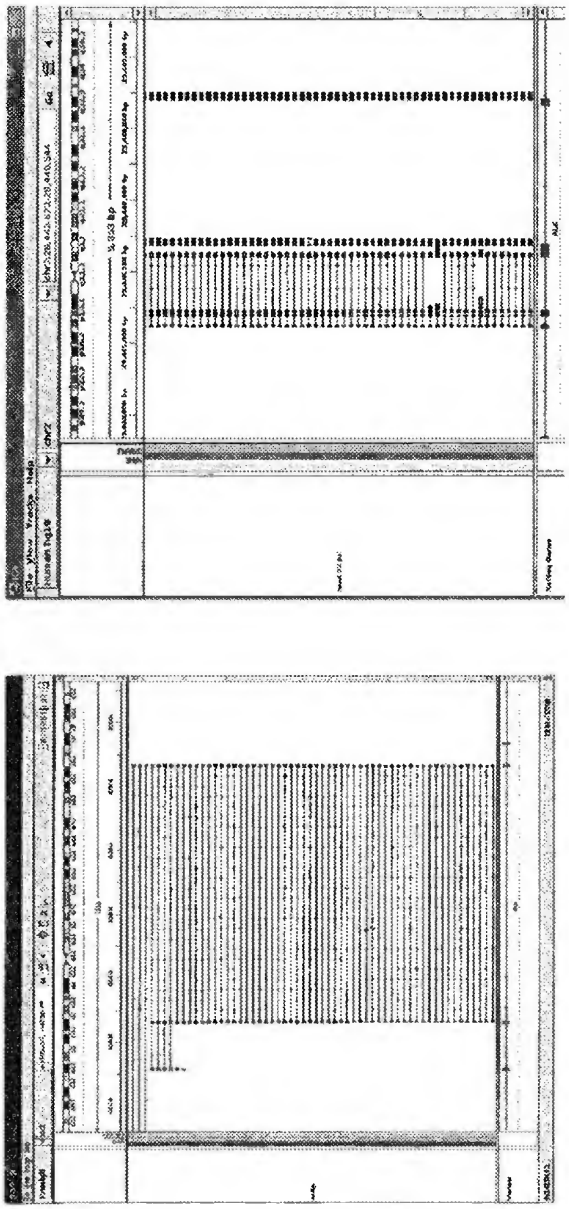


Figure 7

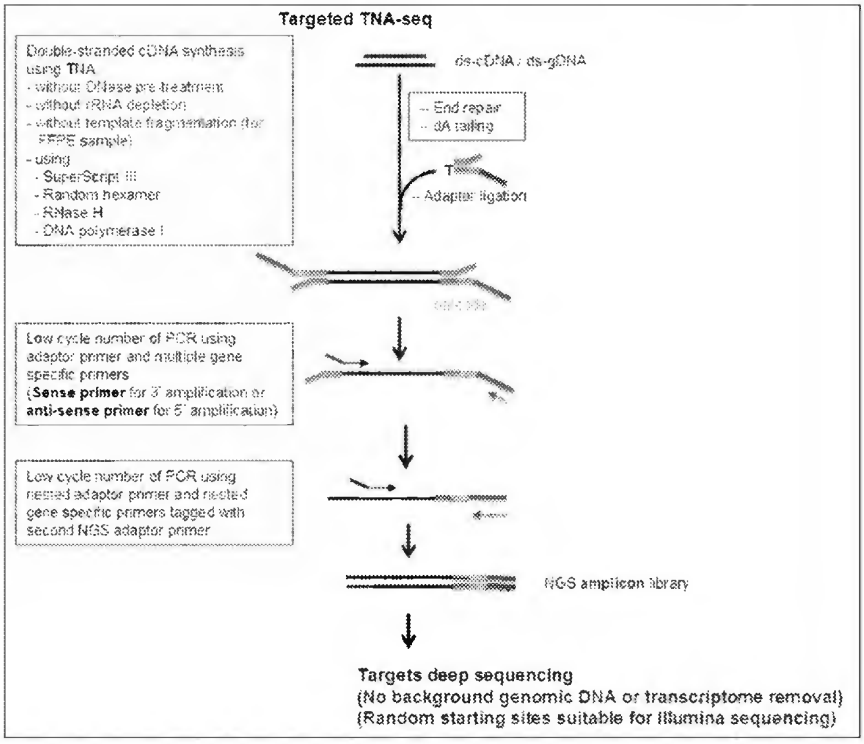


Figure 8

QIA-00689223
JX-001.11

Appx88

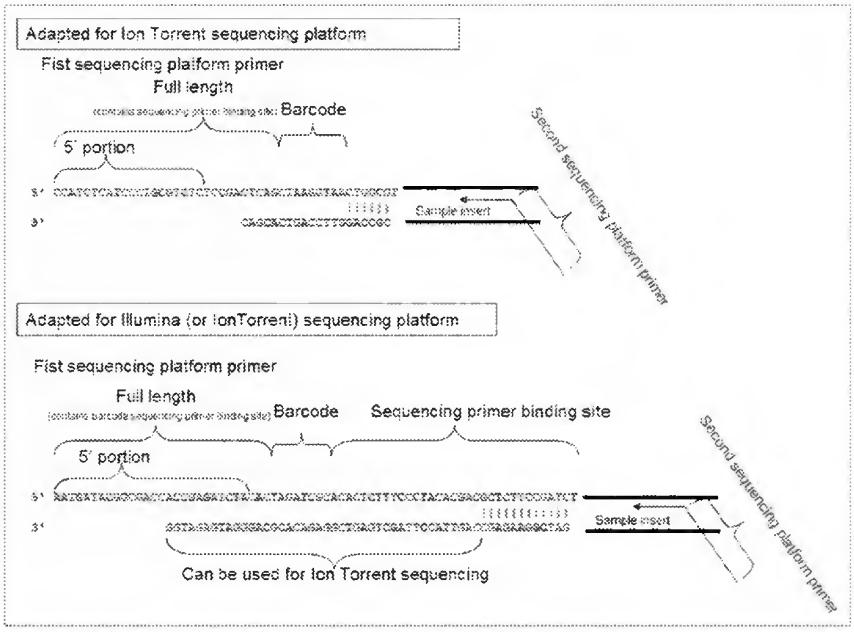


Figure 9

QIA-00689224
JX-001.12

Appx89

METHODS FOR DETERMINING A NUCLEOTIDE SEQUENCE CONTIGUOUS TO A KNOWN TARGET NUCLEOTIDE SEQUENCE

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation of U.S. application Ser. No. 13/793,564 (now U.S. Pat. No. 9,487,828B2) filed Mar. 11, 2013 and which claims benefit under 35 U.S.C. § 119(e) of U.S. Provisional Application Nos. 61/645,364 filed May 10, 2012 and 61/679,302 filed Aug. 3, 2012, the contents of which are incorporated herein by reference in their entirety.

GOVERNMENT SUPPORT

This invention was made with federal funding under Grant No. 5R21CA161590 awarded by the National Institutes of Health. The U.S. government has certain rights in the invention.

SEQUENCE LISTING

The instant application contains a Sequence Listing which has been submitted in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created on Mar. 8, 2013, is named 030258-074963-US_SL.txt and is 41,749 bytes in size.

TECHNICAL FIELD

The technology described herein relates to methods of determining oligonucleotide sequences.

BACKGROUND

Target enrichment prior to next-generation sequencing is more cost-effective than whole genome, whole exome, and whole transcriptome sequencing and therefore more practical for broad implementation; both for research discovery and clinical applications. For example, high coverage depth afforded by target enrichment approaches enables a wider dynamic range for allele counting (in gene expression and copy number assessment) and detection of low frequency mutations, a critical feature for evaluating somatic mutations in cancer. Examples of current enrichment protocols for next generation sequencing include hybridization-based capture assays (TruSeq Capture, Illumina; SureSelect Hybrid Capture, Agilent) and polymerase chain reaction (PCR)-based assays (HaloPlex, Agilent; AmpliSeq, Ion Torrent; TruSeq Amplicon, Illumina; emulsion/digital PCR, Raindance). Hybridization-based approaches capture not only the targeted sequences covered by the capture probes but also near off-target bases that consume sequencing capacity. In addition, these methods are relatively time-consuming, labor-intensive, and suffer from a relatively low level of specificity. A PCR amplification based approach is simpler and faster but by conventional design requires the use of both forward and reverse primers flanking the target loci. In particular, for detection of genomic rearrangements with unknown fusion partners, PCR is not applicable.

SUMMARY

The technology described herein is directed to methods of determining oligonucleotide sequences. In some embodiments, the methods described herein relate to enriching target sequences prior to sequencing the sequences.

In one aspect, the technology described herein relates to a method of determining the nucleotide sequence contiguous to a known target nucleotide sequence, the method comprising: (a) ligating a target nucleic acid comprising the known target nucleotide sequence with a universal oligonucleotide tail-adaptor; (b) amplifying a portion of the target nucleic acid and the amplification strand of the universal oligonucleotide tail-adaptor with a first adaptor primer and a first target-specific primer; (c) amplifying a portion of the amplicon resulting from step (b) with a second adaptor primer and a second target-specific primer; (d) sequencing the amplified portion from step (c) using a first and second sequencing primer; wherein the universal oligonucleotide tail-adaptor comprises a first ligatable duplex end and a second unpaired end; wherein the universal oligonucleotide tail-adaptor comprises a blocking strand and an amplification strand; wherein the blocking strand comprises a 5' duplex portion; wherein the amplification strand comprises an unpaired 5' portion, a 3' duplex portion, and a 3' T overhang; wherein the amplification strand comprises nucleic acid sequences identical to a first and second sequencing primers; wherein the duplex portions of the blocking strand and the amplification strand are substantially complementary and form the first ligatable duplex end comprising a 3' T overhang; wherein the duplex portion is of sufficient length to remain in duplex form at the ligation temperature; wherein the first target-specific primer comprises a nucleic acid sequence that can specifically anneal to the known target nucleotide sequence of the target nucleic acid at the annealing temperature; wherein the second target-specific primer comprises a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (b), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer and the second target-specific primer is nested with respect to the first target-specific primer; wherein the first adaptor primer comprises a nucleic acid sequence identical to a 5' portion of the first sequencing primer; and wherein the second adaptor primer comprises a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first adaptor primer.

In some embodiments, the blocking strand of the universal oligonucleotide tail-adaptor can further comprise a 3' unpaired portion which is not substantially complementary to the 5' unpaired portion of the amplification strand; and wherein the 3' unpaired portion of the blocking strand is not substantially complementary to or substantially identical to any of the primers. In some embodiments, the second adaptor primer can be nested with respect to the first adaptor primer by at least 3 nucleotides. In some embodiments, the portion of the amplification strand that comprises a nucleic acid sequence identical to a first and second sequencing primers can be comprised, at least in part, by the 5' unpaired portion of the amplification strand.

In some embodiments, the first target-specific primer can further comprise a 5' tag sequence portion comprising a nucleic acid sequence of high GC content which is not substantially complementary to or substantially identical to any other portion of any of the primers. In some embodiments, the first target-specific primer can further comprise a 5' tag sequence portion comprising a nucleic acid sequence of high GC content which will not specifically anneal to any other portion of any of the primers or their complements at the annealing temperature. In some embodiments, the sec-

ond adaptor primer can be identical to the full-length first sequencing primer. In some embodiments, the portions of the target-specific primers that specifically anneal to the known target can anneal specifically at a temperature of about 65° C. in a PCR buffer.

In some embodiments, the method can further, prior to step (a), the steps of: mechanically shearing the nucleic acid; subjecting the nucleic acid to end-repair; subjecting the nucleic acid to phosphorylation; and subjecting the nucleic acid to adenylation. In some embodiments, the sample can comprise genomic DNA. In some embodiments, the sample can comprise RNA and the method can further comprise a first step of subjecting the sample to a reverse transcriptase regimen. In some embodiments, the reverse transcriptase regimen can comprise the use of random hexamers.

In some embodiments, the known target sequence can be comprised by a gene rearrangement. In some embodiments, the gene rearrangement can be present in a nucleic acid selected from the group consisting of: genomic DNA; RNA; and cDNA. In some embodiments, the gene rearrangement can comprise an oncogene. In some embodiments, the gene rearrangement can comprise a fusion oncogene.

In some embodiments, the nucleic acid product can be sequenced by a next-generation sequencing method. In some embodiments, the next-generation sequencing method can comprise a method selected from the group consisting of: Ion Torrent, Illumina, SOLiD, 454; Massively Parallel Signature Sequencing solid-phase, reversible dye-terminator sequencing; and DNA nanoball sequencing. In some embodiments, the first and second sequencing primers are compatible with the selected next-generation sequencing method.

In some embodiments, the method can comprise contacting the sample, or separate portions of the sample, with a plurality of sets of first and second target-specific primers. In some embodiments, the method can comprise contacting a single reaction mixture comprising the sample with a plurality of sets of first and second target-specific primers. In some embodiments, the plurality of sets of first and second target-specific primers can specifically anneal to known target nucleotide sequences comprised by separate genes. In some embodiments, at least two sets of first and second target-specific primers can specifically anneal to different portions of a known target nucleotide sequence. In some embodiments, at least two sets of first and second target-specific primers can specifically anneal to different portions of a single gene comprising a known target nucleotide sequence. In some embodiments, at least two sets of first and second target-specific primers can specifically anneal to different exons of a gene comprising a known nucleotide target sequence. In some embodiments, the plurality of first target-specific primers can comprise identical 5' tag sequence portions. In some embodiments, the universal oligonucleotide tail-adaptor can further comprise a barcode portion. In some embodiments, multiple samples can each be contacted with a universal oligonucleotide tail-adaptor with a unique barcode portion and wherein the samples are pooled after step (a).

In some embodiments, each amplification step can comprise a set of cycles of a PCR amplification regimen from 5 cycles to 20 cycles in length. In some embodiments, the target-specific primers and the adaptor primers can be designed such that they will specifically anneal to their complementary sequences at an annealing temperature of from about 61 to 72° C. In some embodiments, the target-specific primers and the adaptor primers can be designed such that they will specifically anneal to their complementary sequences at an annealing temperature of about 65° C.

In some embodiments, the sample can comprise a biological sample obtained from a subject. In some embodiments, the sample can be obtained from a subject in need of treatment for a disease associated with a genetic alteration. In some embodiments, the disease can be cancer. In some embodiments, the sample can comprise a population of tumor cells. In some embodiments, the sample can comprise a tumor biopsy. In some embodiments, the cancer can be lung cancer.

In some embodiments, the known target sequence can be comprised by a disease-associated gene. In some embodiments, the known target sequence can be comprised by a gene rearrangement product in the sample. In some embodiments, the gene rearrangement product can be an oncogene.

In some embodiments, the known target sequence can comprise a sequence from a gene selected from the group of: ALK; ROS1; and RET. In some embodiments, at least one set of a first target-specific primer and a second target-specific primer can be selected from the group consisting of: SEQ ID NOs: 5 and 6; SEQ ID NOs: 7 and 8; SEQ ID NOs: 9 and 10; SEQ ID NOs: 11 and 12; SEQ ID NOs: 13 and 14; SEQ ID NOs: 15 and 16; SEQ ID NOs: 17 and 18; SEQ ID NOs: 19 and 20; SEQ ID NOs: 21 and 22; SEQ ID NOs: 23 and 24; SEQ ID NOs: 25 and 26; SEQ ID NOs: 27 and 28; SEQ ID NOs: 29 and 30; SEQ ID NOs: 31 and 32; SEQ ID NOs: 33 and 34; SEQ ID NOs: 35 and 36; and SEQ ID NOs: 37 and 38.

In some embodiments, the presence of a gene rearrangement of ALK in a sample obtained from a tumor in a subject can indicate that the tumor is susceptible to treatment with a treatment selected from the group consisting of: an ALK inhibitor; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; and NVP-TAE684.

In some embodiments, the presence of a gene rearrangement of ROS1 in a sample obtained from a tumor in a subject can indicate that the tumor is susceptible to treatment with a treatment selected from the group consisting of: a ROS inhibitor; an ALK inhibitor; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; and NVP-TAE684.

In some embodiments, the presence of a gene rearrangement of RET in a sample obtained from a tumor in a subject can indicate that the tumor is susceptible to treatment with a treatment selected from the group consisting of: a RET inhibitor; DP-2490; DP-3636; SU5416; BAY 43-9006; BAY 73-4506 (regorafenib); ZD6474; NVP-AST487; sorafenib; RPI-1; XL184; vandetanib; sunitinib; imatinib; pazopanib; axitinib; motesanib; gefitinib; and withaferin A.

In one aspect, the technology described herein relates to a method of treating cancer, the method comprising: detecting, in a tumor sample obtained from a subject in need of treatment for cancer, the presence of one or more oncogene rearrangements according to the method described herein; administering a cancer treatment which is effective against tumors having any of the detected oncogene rearrangements. In some embodiments, a treatment selected from the group consisting of: an ALK inhibitor; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; and NVP-TAE684; can be effective against tumors having an ALK oncogene rearrangement. In some embodiments, a treatment selected from the group consisting of: a ROS1 inhibitor; an ALK inhibitor; crizotinib (PF-02341066); AP26113;

LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; and NVP-TAE684; can be effective against tumors having an ROS1 oncogene rearrangement. In some embodiments, a treatment selected from the group consisting of: a RET inhibitor; DP-2490; DP-3636; SU5416; BAY 43-9006; BAY 73-4506 (regorafenib); ZD6474; NVP-AST487; sorafenib; RPI-1; XL184; vandetanib; sunitinib; imatinib; pazopanib; axitinib; motesanib; gefitinib; and withaferin A; can be effective against tumors having an RET oncogene rearrangement.

In one aspect, the technology described herein relates to a method of determining if a subject in need of treatment for cancer will be responsive to a given treatment, the method comprising; detecting, in a tumor sample obtained from the subject, the presence of an oncogene rearrangement according to the method as described herein; wherein the subject is determined to be responsive to a treatment targeting an oncogene rearrangement product if the presence of the oncogene rearrangement is detected.

In some embodiments, when the presence of an ALK oncogene rearrangement is detected, the subject can be responsive to a treatment selected from the group consisting of: an ALK inhibitor; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; and NVP-TAE684. In some embodiments, when the presence of an ROS1 oncogene rearrangement is detected, the subject can be responsive to a treatment selected from the group consisting of: an ALK inhibitor; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; and NVP-TAE684.

In some embodiments, when the presence of an RET oncogene rearrangement is detected, the subject will be responsive to a treatment selected from the group consisting of: a RET inhibitor; DP-2490; DP-3636; SU5416; BAY 43-9006; BAY 73-4506 (regorafenib); ZD6474; NVP-AST487; sorafenib; RPI-1; XL184; vandetanib; sunitinib; imatinib; pazopanib; axitinib; motesanib; gefitinib; and withaferin A.

In some embodiments, the cancer can be lung cancer.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** depicts a schematic illustration of an example of library construction for targeted RNA and DNA sequencing. 1) A standard procedure of double-stranded cDNA synthesis is applied using total nucleic acid from FFPE specimen as starting material. Alternatively, the starting material may also be sheared gDNA. 2) After cleanup, the double stranded cDNA or gDNA is subjected to end-repair and dA tailing, directly followed by ligation of a half-truncated Y adapter without cleanup in between. 3) After SPRI cleanup, the ligated sample is subjected to 14 cycles of PCR amplification using multiplex gene specific primers (GSP1s) and Ion Torrent short length Forward primer A 5' 20-mer (A20), with annealing temperature at 65° C. 4). After a second SPRI cleanup, the sample is subjected to an additional 14 cycles of PCR amplification using multiplex nested gene specific primers (3' downstream of GSP1 and in the same direction) tagged with Ion Torrent Reverse primer (P1_GSP2s) and Forward primer A (full length 30-mer), with annealing temperature at 65° C. 5) After a final third SPRI cleanup, the product is Ion Torrent library ready for downstream emulsion PCR and sequencing. Sample pooling may be carried out after Step 2.

FIG. **2** depicts the mapping results, demonstrating different primer (dots) extension products in one sequencing run of ROS1 sequences, including reads that corresponding to genomic DNA (spanning intron-exon boundaries), cDNA (spanning exons) and fusion cDNA (on exon 34, mapping on the fusion partner SLC34A2 not shown here). 91.8% specificity (127,446/138,787) was achieved.

FIG. **3**A depicts a schematic presentation of nested primer targeting strategy using ROS1 as an example. The assay target panel includes a total of 17 pairs of GSP1s and GSP2s for ROS1, ALK and RET. FIG. **3**B depicts a representation of the possible types of extension products using gDNA and cDNA templates.

FIG. **4**A depicts the visualization of reads mapping using the Integrative Genomics Viewer. FIG. **4**B depicts the sequence of two alternative spliced fusion sequences (SEQ ID NOS 39-44, respectively, in order of appearance). FIG. **4**C depicts a summary table reporting fusion transcripts in involved genes annotated with exon, frame, and fusion read coverage details.

FIG. **5** depicts the results of an example sequencing run.

FIGS. **6** and **7** depict examples of the results of sequencing runs for ALK and RET sequences, respectively.

FIG. **8** depicts a schematic presentation of the targeted sequencing approach described herein.

FIG. **9** depicts schematics of illustrative universal oligonucleotide tail-adaptors (SEQ ID NOS 45-48, respectively, in order of appearance).

DETAILED DESCRIPTION

Embodiments of the technology described herein relate to methods of determining (i.e sequencing) oligonucleotide sequences. In some embodiments, the methods described herein relate to methods of enriching target sequences prior to a sequencing step. In some embodiments, the sequence of one end of the target sequence to be enriched is not known prior to the sequencing step.

For convenience, the meaning of some terms and phrases used in the specification, examples, and appended claims, are provided below. Unless stated otherwise, or implicit from context, the following terms and phrases include the meanings provided below. The definitions are provided to aid in describing particular embodiments, and are not intended to limit the claimed invention, because the scope of the invention is limited only by the claims. Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. If there is an apparent discrepancy between the usage of a term in the art and its definition provided herein, the definition provided within the specification shall prevail.

For convenience, certain terms employed herein, in the specification, examples and appended claims are collected here.

The terms "decrease", "reduced", "reduction", or "inhibit" are all used herein generally to mean a decrease by a statistically significant amount. However, for avoidance of doubt, "reduced", "reduction", "decrease", or "inhibit" means a decrease by at least 10% as compared to a reference level, for example a decrease by at least about 20%, or at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90% or up to and including a 100% decrease (e.g. absent level or non-detectable level as compared to a reference level), or any decrease between 10-100% as compared to a reference level. In the context of a marker or symptom is meant a statistically significant decrease in such level. The decrease can be, for example, at

least 10%, at least 20%, at least 30%, at least 40% or more, and is preferably down to a level accepted as within the range of normal for an individual without such disorder.

The terms "increased", "increase", "enhance", or "activate" are all used herein to generally mean an increase by a statically significant amount; for the avoidance of doubt, the terms "increased", "increase", "enhance", or "activate" mean an increase of at least 10% as compared to a reference level, for example an increase of at least about 20%, or at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90% or up to and including a 100% increase or any increase between 10-100% as compared to a reference level, or at least about a 2-fold, or at least about a 3-fold, or at least about a 4-fold, or at least about a 5-fold or at least about a 10-fold increase, or any increase between 2-fold and 10-fold or greater as compared to a reference level.

As used herein, a "subject" means a human or animal. Usually the animal is a vertebrate such as a primate, rodent, domestic animal or game animal. Primates include chimpanzees, cynomologous monkeys, spider monkeys, and macaques, e.g., Rhesus. Rodents include mice, rats, woodchucks, ferrets, rabbits and hamsters. Domestic and game animals include cows, horses, pigs, deer, bison, buffalo, feline species, e.g., domestic cat, canine species, e.g., dog, fox, wolf, avian species, e.g., chicken, emu, ostrich, and fish, e.g., trout, catfish and salmon. In some embodiments, the subject is a mammal, e.g., a primate, e.g., a human. The terms, "individual," "patient" and "subject" are used interchangeably herein.

Preferably, the subject is a mammal. The mammal can be a human, non-human primate, mouse, rat, dog, cat, horse, or cow, but is not limited to these examples. Mammals other than humans can be advantageously used as subjects that represent animal models of, e.g. lung cancer. A subject can be male or female.

A subject can be one who has been previously diagnosed with or identified as suffering from or having a condition in need of treatment (e.g. cancer) or one or more complications related to such a condition, and optionally, have already undergone treatment for the condition or the one or more complications related to the condition. Alternatively, a subject can also be one who has not been previously diagnosed as having the condition (e.g. cancer) or one or more complications related to the condition. For example, a subject can be one who exhibits one or more risk factors for the condition or one or more complications related to the condition or a subject who does not exhibit risk factors.

A "subject in need" of treatment for a particular condition can be a subject having that condition, diagnosed as having that condition, or at risk of developing that condition.

As used herein, a "disease associated with a genetic alteration" refers to any disease which is caused by, at least in part, by an alteration in the genetic material of the subject as compared to a healthy wildtype subject, e.g. a deletion, an insertion, a SNP, a gene rearrangement. A disease can be caused by, at least in part, an alteration in the genetic material of the subject if the alteration increases the risk of the subject developing the disease, increases the subject's susceptibility to a disease (including infectious diseases, or diseases with an infectious component), causes the production of a disease-associated molecule, or causes cells to become diseased or abnormal (e.g. loss of cell cycle regulation in cancer cells). Diseases can be associated with multiple genetic alterations, e.g. cancers.

As used herein, the term "nucleic acid" or "nucleic acid sequence" refers to any molecule, preferably a polymeric molecule, incorporating units of ribonucleic acid, deoxyribonucleic acid or an analog thereof. The nucleic acid can be either single-stranded or double-stranded. A single-stranded nucleic acid can be one strand nucleic acid of a denatured double-stranded DNA. Alternatively, it can be a single-stranded nucleic acid not derived from any double-stranded DNA. In one aspect, the template nucleic acid is DNA. In another aspect, the template is RNA. Suitable nucleic acid molecules are DNA, including genomic DNA or cDNA. Other suitable nucleic acid molecules are RNA, including mRNA.

The term "isolated" or "partially purified" as used herein refers, in the case of a nucleic acid, to a nucleic acid separated from at least one other component (e.g., nucleic acid or polypeptide) that is present with the nucleic acid as found in its natural source and/or that would be present with the nucleic acid when expressed by a cell. A chemically synthesized nucleic acid or one synthesized using in vitro transcription/translation is considered "isolated."

The term "gene" means a nucleic acid sequence which is transcribed (DNA) to RNA in vitro or in vivo when operably linked to appropriate regulatory sequences. The gene can include regulatory regions preceding and following the coding region, e.g. 5' untranslated (5'UTR) or "leader" sequences and 3' UTR or "trailer" sequences, as well as intervening sequences (introns) between individual coding segments (exons).

As used herein, the term "complementary" refers to the hierarchy of hydrogen-bonded base pair formation preferences between the nucleotide bases G, A, T, C and U, such that when two given polynucleotides or polynucleotide sequences anneal to each other, A pairs with T and G pairs with C in DNA, and G pairs with C and A pairs with U in RNA. As used herein, "substantially complementary" refers to a nucleic acid molecule or portion thereof (e.g. a primer) having at least 90% complementarity over the entire length of the molecule or portion thereof with a second nucleotide sequence, e.g. 90% complementary, 95% complementary, 98% complementary, 99% complementary, or 100% complementary. As used herein, "substantially identical" refers to a nucleic acid molecule or portion thereof having at least 90% identity over the entire length of a the molecule or portion thereof with a second nucleotide sequence, e.g. 90% identity, 95% identity, 98% identity, 99% identity, or 100% identity.

As used herein, "specific" when used in the context of a primer specific for a target nucleic acid refers to a level of complementarity between the primer and the target such that there exists an annealing temperature at which the primer will anneal to and mediate amplification of the target nucleic acid and will not anneal to or mediate amplification of non-target sequences present in a sample.

As used herein, "amplified product", "amplification product", or "amplicon" refers to oligonucleotides resulting from a PCR reaction that are copies of a portion of a particular target nucleic acid template strand and/or its complementary sequence, which correspond in nucleotide sequence to the template oligonucleotide sequence and/or its complementary sequence. An amplification product can further comprise sequence specific to the primers and which flanks sequence which is a portion of the target nucleic acid and/or its complement. An amplified product, as described herein will generally be double-stranded DNA, although reference can be made to individual strands thereof.

As used herein, a "portion" of a nucleic acid molecule refers to contiguous set of nucleotides comprised by that molecule. A portion can comprise all or only a subset of the nucleotides comprised by the molecule. A portion can be double-stranded or single-stranded.

As used herein, the terms "treat," "treatment," "treating," or "amelioration" refer to therapeutic treatments, wherein the object is to reverse, alleviate, ameliorate, inhibit, slow down or stop the progression or severity of a condition associated with a disease or disorder, e.g. lung cancer. The term "treating" includes reducing or alleviating at least one adverse effect or symptom of a condition, disease or disorder associated with a condition. Treatment is generally "effective" if one or more symptoms or clinical markers are reduced. Alternatively, treatment is "effective" if the progression of a disease is reduced or halted. That is, "treatment" includes not just the improvement of symptoms or markers, but also a cessation of, or at least slowing of, progress or worsening of symptoms compared to what would be expected in the absence of treatment. Beneficial or desired clinical results include, but are not limited to, alleviation of one or more symptom(s), diminishment of extent of disease, stabilized (i.e., not worsening) state of disease, delay or slowing of disease progression, amelioration or palliation of the disease state, remission (whether partial or total), and/or decreased mortality, whether detectable or undetectable. The term "treatment" of a disease also includes providing relief from the symptoms or side-effects of the disease (including palliative treatment).

The term "statistically significant" or "significantly" refers to statistical significance and generally means a two standard deviation (2SD) below normal, or lower, concentration of the marker.

Other than in the operating examples, or where otherwise indicated, all numbers expressing quantities of ingredients or reaction conditions used herein should be understood as modified in all instances by the term "about." The term "about" when used in connection with percentages can mean±1%.

As used herein the term "comprising" or "comprises" is used in reference to compositions, methods, and respective component(s) thereof, that are essential to the method or composition, yet open to the inclusion of unspecified elements, whether essential or not.

The term "consisting of" refers to compositions, methods, and respective components thereof as described herein, which are exclusive of any element not recited in that description of the embodiment.

As used herein the term "consisting essentially of" refers to those elements required for a given embodiment. The term permits the presence of elements that do not materially affect the basic and novel or functional characteristic(s) of that embodiment.

The singular terms "a," "an," and "the" include plural referents unless context clearly indicates otherwise. Similarly, the word "or" is intended to include "and" unless the context clearly indicates otherwise. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of this disclosure, suitable methods and materials are described below. The abbreviation, "e.g." is derived from the Latin exempli gratia, and is used herein to indicate a non-limiting example. Thus, the abbreviation "e.g." is synonymous with the term "for example."

Definitions of common terms in cell biology and molecular biology can be found in "The Merck Manual of Diagnosis and Therapy", 19th Edition, published by Merck Research Laboratories, 2006 (ISBN 0-911910-19-0); Robert S. Porter et al. (eds.), The Encyclopedia of Molecular Biology, published by Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9). Definitions of common terms in molecular biology can also be found in Benjamin Lewin, Genes X, published by Jones & Bartlett Publishing, 2009 (ISBN-10: 0763766321); Kendrew et al. (eds.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8) and Current Protocols in Protein Sciences 2009, Wiley Intersciences, Coligan et al., eds.

Unless otherwise stated, the present invention was performed using standard procedures, as described, for example in Sambrook et al., Molecular Cloning: A Laboratory Manual (3 ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., USA (2001); and Davis et al., Basic Methods in Molecular Biology, Elsevier Science Publishing, Inc., New York, USA (1995) which are all incorporated by reference herein in their entireties.

Other terms are defined herein within the description of the various aspects of the invention.

Described herein are methods of determining the nucleotide sequence contiguous to a known target nucleotide sequence. Traditional sequencing methods generate sequence information randomly (e.g. "shotgun" sequencing) or between two known sequences which are used to design primers. In contrast, the methods described herein, in some embodiments, allow for determining the nucleotide sequence (e.g. sequencing) upstream or downstream of a single region of known sequence with a high level of specificity and sensitivity.

In some embodiments, the methods described herein relate to a method of enriching specific nucleotide sequences prior to determining the nucleotide sequence using a next-generation sequencing technology. In some embodiments, the methods of enriching specific nucleotide sequences do not comprise hybridization enrichment.

In some embodiments, the technology described herein can relate to a method of determining the nucleotide sequence contiguous to a known target nucleotide sequence, the method comprising: (a) ligating a target nucleic acid comprising the known target nucleotide sequence with a universal oligomcleotide tail-adaptor; (b) amplifying a portion of the target nucleic acid and the amplification strand of the universal oligonucleotide tail-adaptor with a first adaptor primer and a first target-specific primer; (c) amplifying a portion of the amplicon resulting from step (b) with a second adaptor primer and a second target-specific primer; and (d) sequencing the amplified portion from step (c) using a first and second sequencing primer. As used herein, the term "target nucleic acid" refers to a nucleic acid molecule comprising both the nucleic acid sequence which is to be determined and the known target nucleotide sequence. The target nucleic acid can be of any length and can be double-stranded or single-stranded. As used herein, the term "known target nucleotide sequence" refers to a portion of a target nucleic acid for which the sequence (e.g. the identity and order of the nucleotide bases comprises the nucleic acid) is known. The known target nucleotide sequence can be of any length of 10 or more nucleotides, preferably 30 or more nucleotides (e.g. 30 nucleotides, 40 nucleotides, 50 nucleotides or more). As used herein, the term "nucleotide sequence contiguous to" refers to a nucleotide sequence which is located on the same nucleic acid molecule (i.e. the target nucleic acid) as the known target nucleotide sequence and either upstream or downstream of the known target nucleotide sequence. The nucleotide sequence contiguous to

can comprise any length of nucleotide sequence. In some embodiments, the nucleotide sequence contiguous to the known target nucleotide sequence comprises 1 kb or less of nucleotide sequence, e.g. 1 kb or less of nucleotide sequence, 750 bp or less of nucleotide sequence, 500 bp or less of nucleotide sequence, 400 bp or less of nucleotide sequence, 300 bp or less of nucleotide sequence, 200 bp or less of nucleotide sequence, 100 bp or less of nucleotide sequence. Where a sample comprises different target nucleic acids comprising the known target nucleotide sequence (e.g. a cell where the known target nucleotide sequence occurs multiple times in the genome, or on separate, non-identical chromosomes), there can be multiple sequences which comprise "nucleotide sequence contiguous to" the known target nucleotide sequence. As used herein, the term "determining the nucleotide sequence", refers to determining the identity and relative positions of the nucleotide bases comprising a nucleic acid.

In step (a) of the method described herein, the universal oligonucleotide tail-adaptor can be ligated to the target nucleic acid. In some embodiments, the target nucleic acid can be comprised by a sample comprising a plurality of nucleic acids, some of which do not comprise the target nucleic acid. In some embodiments, the universal oligonucleotide tail-adaptor can be ligated to substantially all of the nucleic acids in a sample. In some embodiments, the universal oligonucleotide tail-adaptor can be ligated to both nucleic acids which comprise the target nucleic acid sequence and to nucleic acids which do not comprise the target nucleic acid sequence.

As used herein, the term "universal oligonucleotide tail-adaptor" refers to a nucleic acid molecule comprised of two strands (a blocking strand and an amplification strand) and comprising a first ligatable duplex end and a second unpaired end. The blocking strand of the universal oligonucleotide tail-adaptor comprises a 5' duplex portion. The amplification strand comprises an unpaired 5' portion, a 3' duplex portion, and a 3' T overhang and nucleic acid sequences identical to a first and second sequencing primers. The duplex portions of the blocking strand and the amplification strand are substantially complementary and form the first ligatable duplex end comprising a 3' T overhang and the duplex portion is of sufficient length to remain in duplex form at the ligation temperature.

In some embodiments, the portion of the amplification strand that comprises a nucleic acid sequence identical to a first and second sequencing primers can be comprised, at least in part, by the 5' unpaired portion of the amplification strand.

In some embodiments, the universal oligonucleotide tail-adaptor can comprise a duxplex portion and an unpaired portion, wherein the unpaired portion comprises only the 5' portion of the amplification strand, i.e. the entirety of the blocking strand is a duplex portion.

In some embodiments, the universal oligonucleotide tail-adaptor can have a "Y" shape, i.e. the unpaired portion can comprise portions of both the blocking strand and the amplification strand which are unpaired. The unpaired portion of the blocking strand can be shorter than, longer than, or equal in length to the unpaired portion of the amplification strand. In some embodiments, the unpaired portion of the blocking strand can be shorter than the unpaired portion of the amplification strand. Y shaped universal oligonucleotide tail-adaptors have the advantage that the unpaired portion of the blocking strand will not be subject to 3' extension during a PCR regimen.

In some embodiments, the blocking strand of the universal oligonucleotide tail-adaptor can further comprise a 3' unpaired portion which is not substantially complementary to the 5' unpaired portion of the amplification strand; and wherein the 3' unpaired portion of the blocking strand is not substantially complementary to or substantially identical to any of the primers. In some embodiments, the blocking strand of the universal oligonucleotide tail-adaptor can further comprise a 3' unpaired portion which will not specifically anneal to the 5' unpaired portion of the amplification strand at the annealing temperature; and wherein the 3' unpaired portion of the blocking strand will not specifically anneal to any of the primers or the complements thereof at the annealing temperature.

In some embodiments, the duplex portion of the universal oligonucleotide tail-adaptor (e.g. the duplex portions of either or both of the strands) is at least 7 base pairs in length, e.g. 7 bp or more, 8 bp or more, 9 bp or more, 10 bp or more, 11 bp or more, 12 bp or more, 13 bp or more, or 14 bp or more in length. In some embodiments, the duplex portion of a universal oligonucleotide tail-adaptor can be at least 30 bp or longer, e.g. 30 bp or more, 31 bp or more, 32 bp or more, 33 bp or more, 34 bp or more, 35 bp or more, 40 bp or more, or 50 bp or more in length. The duplex portion of a universal oligonucleotide tail-adaptor should not be so long as to suppress PCR amplification of the desired amplicons in the PCR amplification regimen being used. Some next-generation sequencing methods use Y-shaped adaptor molecules. These Y-shaped adaptor molecules require duplex portions that are of limited length, (e.g. 17 bp or less) to avoid the formation of intramolecular hairpins during several PCR steps (e.g. library enrichment PCR, bridge PCR, or emulsion PCR). The Y-shaped universal oligonucleotide tail-adaptors of the methods described herein are not subject to this limitation of the duplex end as the two This PCR suppression effect resulting from the duplex end is not applicable in this invention since the two target primers can access the target fragment internally. In some embodiments, the duplex portion of a universal oligonucleotide tail-adaptor can be at least 18 bp or longer, e.g. 18 bp or more, 19 bp or more, 20 bp or more, 21 bp or more, 22 bp or more, 23 bp or more, 24 bp or more, or 25 bp or more in length.

Illustrative examples of universal oligonucleotide tail-adaptors are shown in FIG. **9**.

Ligation of the universal oligonucleotide tail-adaptor can be accomplished by any method known in the art, e.g. blunt-end ligation or TA ligation. In some embodiments, prior to ligation of the Universal oligonucleotide tail-adaptor, the nucleic acids in a sample can be subjected to nucleic acid end-repair to blunt the ends of the nucleic acid. End-repair is well known in the art and relevant kits and/or enzymes are available commercially, (e.g. the NEBNEXT™ End Repair Module (Cat No. E6050L; New England Biolabs; Ipswich, Mass.).

In some embodiments, prior to ligation of the universal oligonucleotide tail-adaptor, the nucleic acids in a sample can be phosphorylated and/or adenylated. Adenylation can provide an adenosine overhang on the 3' end of a nucleic acid. A second nucleic acid with a thionine 3' overhang can then be ligated to the first nucleic acid by TA ligation. Methods of ligation are well known in the art and relevant kits and/or enzymes are available commercially, e.g. the NEBNEXT™ da-Tailing module (Cat No. E6053L: New England Biolabs; Ipswich, Mass.) can be used to adenylate a blunt end of a nucleic acid. In some embodiments, Universal oligonucleotide tail-adaptors can be provided with a thioninic 3' overhang.

Steps (b) and (c) of the methods described herein can each comprise a PCR amplification regimen, i.e. a set of polymerase chain reaction (PCR) amplification cycles. As used herein, the term "amplification regimen" refers to a process of specifically amplifying, i.e., increasing the abundance of, a nucleic acid sequence of interest, and more particularly, the exponential amplification occurring when the products of a previous polymerase extension serve as templates for the successive rounds of extension. A PCR amplification regimen according to the invention comprises at least one, and preferably at least 5 or more iterative cycles, where each cycle comprises the steps of: 1) strand separation (e.g., thermal denaturation); 2) oligonucleotide primer annealing to template molecules; and 3) nucleic acid polymerase extension of the annealed primers. Conditions and times necessary for each of these steps can be devised by one of ordinary skill in the art. An amplification regimen according to the methods described herein is preferably performed in a thermal cycler, many of which are commercially available.

PCR requires the use of a nucleic acid polymerase. As used herein, the phrase "nucleic acid polymerase" refers an enzyme that catalyzes the template-dependent polymerization of nucleoside triphosphates to form primer extension products that are complementary to the template nucleic acid sequence. A nucleic acid polymerase enzyme initiates synthesis at the 3' end of an annealed primer and proceeds in the direction toward the 5' end of the template. Numerous nucleic acid polymerases are known in the art and commercially available. One group of preferred nucleic acid polymerases are thermostable, i.e., they retain function after being subjected to temperatures sufficient to denature annealed strands of complementary nucleic acids, e.g. 94° C., or sometimes higher.

As understood in the art, PCR requires cycles including a strand separation step generally involving heating of the reaction mixture. As used herein, the term "strand separation" or "separating the strands" means treatment of a nucleic acid sample such that complementary double-stranded molecules are separated into two single strands available for annealing to an oligonucleotide primer. More specifically, strand separation according to the methods described herein is achieved by heating the nucleic acid sample above its T_m. Generally, for a sample containing nucleic acid molecules in buffer suitable for a nucleic acid polymerase, heating to 94° C. is sufficient to achieve strand separation. An exemplary buffer contains 50 mM KCl, 10 mM Tris-HCl (pH 8.8@25° C.), 0.5 to 3 mM $MgCl_2$, and 0.1% BSA.

As also understood in the art, PCR requires annealing primers to template nucleic acids. Any strand of a target nucleic acid can be a template nucleic acid, as the template nucleic acid is defined as a single-strand nucleic acid to which a given primer will specifically anneal. As used herein, "anneal" refers to permitting two complementary or substantially complementary nucleic acids strands to hybridize, and more particularly, when used in the context of PCR, to hybridize such that a primer extension substrate for a template-dependent polymerase enzyme is formed. Conditions for primer-target nucleic acid annealing vary with the length and sequence of the primer and are based upon the calculated T_m for the primer. Generally, an annealing step in an amplification regimen involves reducing the temperature following the strand separation step to a temperature based on the calculated T_m for the primer sequence, for a time sufficient to permit such annealing. T_m can be readily predicted by one of skill in the art using any of a number of widely available algorithms (e.g., OLIGO™ (Molecular Biology Insights Inc. Colorado) primer design software and VENTRO NTI™ (Invitrogen, Inc. California) primer design software and programs available on the internet, including Primer3, Oligo Calculator, and NetPrimer (Premier Biosoft: Palo Alto, Calif.; and freely available on the world wide web at www.premierbiosoft.com/netprimer/netprlaunch/Help/xnetprlaunch.html). For example, the T_m of a primer can be calculated using following formula, which is used by NetPrimer software and is described in more detail in Frieir et al. PNAS 1986 83:9373-9377 which is incorporated by reference herein in its entirety.

$$T_m = \frac{\Delta H/(\Delta S + R^* \ln(C/4)) + 16.6 \log([K^+]/(1+0.7[K^+]))}{273.15} -$$

wherein, ΔH is enthalpy for helix formation; ΔS is entropy for helix formation; R is molar gas constant (1.987 cal/° C.*mol); C is the nucleic acid concentration; and $[K^+]$ is salt concentration. For most amplification regimens, the annealing temperature is selected to be about 5° C. below the predicted T_m, although temperatures closer to and above the T_m (e.g., between 1° C. and 5° C. below the predicted T_m or between 1° C. and 5° C. above the predicted T_m) can be used, as can, for example, temperatures more than 5° C. below the predicted T_m (e.g., 6° C. below, 8° C. below, 10° C. below or lower). Generally, the closer the annealing temperature is to the T_m, the more specific is the annealing. The time allowed for primer annealing during a PCR amplification regimen depends largely upon the volume of the reaction, with larger volumes requiring longer times, but also depends upon primer and template concentrations, with higher relative concentrations of primer to template requiring less time than lower relative concentrations. Depending upon volume and relative primer/template concentration, primer annealing steps in an amplification regimen can be on the order of 1 second to 5 minutes, but will generally be between 10 seconds and 2 minutes, preferably on the order of 30 seconds to 2 minutes. As used herein, "substantially anneal" refers to a degree of annealing during a PCR amplification regimen which is sufficient to produce a detectable level of a specifically amplified product.

PCR also relies upon polymerase extension of annealed primers at each cycle. As used herein, the term "polymerase extension" means the template-dependent incorporation of at least one complementary nucleotide, by a nucleic acid polymerase, onto the 3' end of an annealed primer. Polymerase extension preferably adds more than one nucleotide, preferably up to and including nucleotides corresponding to the full length of the template. Conditions for polymerase extension vary with the identity of the polymerase. The temperature used for polymerase extension is generally based upon the known activity properties of the enzyme. Although, where annealing temperatures are required to be, for example, below the optimal temperatures for the enzyme, it will often be acceptable to use a lower extension temperature. In general, although the enzymes retain at least partial activity below their optimal extension temperatures, polymerase extension by the most commonly used thermostable polymerases (e.g., Taq polymerase and variants thereof) is performed at 65° C. to 75° C., preferably about 68-72° C.

Primer extension is performed under conditions that permit the extension of annealed oligonucleotide primers. As used herein, the term "conditions that permit the extension of an annealed oligonucleotide such that extension products are generated" refers to the set of conditions including, for example temperature, salt and co-factor concentrations, pH, and enzyme concentration under which a nucleic acid poly-

merase catalyzes primer extension. Such conditions will vary with the identity of the nucleic acid polymerase being used, but the conditions for a large number of useful polymerase enzymes are well known to those skilled in the art. One exemplary set of conditions is 50 mM KCl, 10 mM Tris-HCl (pH 8.8@25° C.), 0.5 to 3 mM $MgCl_2$, 200 uM each dNTP, and 0.1% BSA at 72° C., under which Taq polymerase catalyzes primer extension. The conditions for initiation and extension usually include the presence of at least one, but more preferably all four different deoxyribonucleoside triphosphates and a polymerization-inducing agent such as DNA polymerase or reverse transcriptase, in a suitable buffer (in this context "buffer" includes solvents (generally aqueous) plus necessary cofactors and reagents which affect pH, ionic strength, etc.) and at a suitable temperature.

In some embodiments, each amplification step can comprise a set of cycles of a PCR amplification regimen from 5 cycles to 20 cycles in length. In some embodiments, each amplification step can comprise a set of cycles of a PCR amplification regimen from 10 cycles to 20 cycles in length. In some embodiments, each amplification step can comprise a set of cycles of a PCR amplification regimen from 12 cycles to 16 cycles in length. In some embodiments, the annealing temperature can be less than 70° C. In some embodiments, the annealing temperature can be less than 72° C.

In various embodiments, the methods and compositions described herein relate to performing a PCR amplification regimen with one or more of the types of primers described herein. As used herein, "primer" refers to a DNA or RNA polynucleotide molecule or an analog thereof capable of specifically annealing to a polynucleotide template and providing a 3' end that serves as a substrate for a template-dependent polymerase to produce an extension product which is complementary to the polynucleotide template. A primer useful in the methods described herein is generally single-stranded, and a primer and its complement can anneal to form a double-stranded polynucleotide. Primers according to the methods and compositions described herein can be less than or equal to 300 nucleotides in length, e.g., less than or equal to 300, or 250, or 200, or 150, or 100, or 90, or 80, or 70, or 60, or 50, or 40, or 30 or fewer, or 20 or fewer, or 15 or fewer, but at least 10 nucleotides in length. Methods of making primers are well known in the art, and numerous commercial sources offer oligonucleotide synthesis services suitable for providing primers according to the methods and compositions described herein, e.g. INVITROGEN™ Custom DNA Oligos; Life Technologies; Grand Island, N.Y. or custom DNA Oligos from IDT; Coralville, Iowa).

In some embodiments, after the Universal oligonucleotide tail-adaptor is ligated to the nucleic acids in a sample (e.g. the target nucleic acids), the target nucleic acid can be amplified in a first amplification step (i.e. step (b)). The first amplification step can be a set of PCR amplification cycles using a first target-specific primer and a first tail-adaptor primer.

As used herein, the term "first target-specific primer" refers to a single-stranded oligonucleotide comprising a nucleic acid sequence that can specifically anneal to the target nucleic acid at the annealing temperature.

In some embodiments, the first-target specific primer can comprise a 5' tag sequence portion. In some embodiments, all first-target specific primers present in a reaction can comprise identical 5' tag sequence portions. In a multiplex PCR reaction, different primer species can interact with each other in an undesired off-target manner, leading to primer extension and subsequently amplification by DNA polymerase. These primer dimers tend to be short, and their efficient amplification can overtake the reaction and dominate resulting in poor amplification of desired target sequence. The inclusion of a 5' tag sequence on the first-target specific primer(s) causes any potential primer dimers that can result to contain the same complementary tails on both ends. In subsequent PCR cycles, the primer dimers would denature into single-stranded DNA primer dimers, each comprising complementary sequences on their two ends which are introduced by the 5' tag. Instead of primer annealing to these single stranded DNA primer dimers, an intra-molecular hairpin (a panhandle like structure) formation would preferentially occur due to the proximate accessibility of the complementary tags on the same primer dimer molecule instead of an inter-molecular interaction with new primers on separate molecules. As a result, these primer dimers are very inefficiently amplified, such that primers are not exponentially consumed by the undesired dimers for amplification. Instead the tagged primers can remain in high and sufficient concentration for desired specific amplification of target sequences. Accumulation of primer dimers can also be a detriment to multiplex PCR because they compete for and consume other reagents in the reaction. In some embodiments, the 5' tag sequence can be a GC-rich sequence, i.e. the tag sequence can comprise at least 50% GC content, at least 55% GC content, at least 60% GC content, at least 65% GC content, at least 70% GC content, at least 75% GC content, at least 80% GC content, or higher GC content. In some embodiments, the tag sequence can comprise at least 60% GC content. In some embodiments, the tag sequence can comprise at least 65% GC content.

As used herein, the term "second target-specific primer" refers to a single-stranded oligonucleotide comprising a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (b), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer. The second target-specific primer is nested with respect to the first target-specific primer. In some embodiments, the second target-specific primer is nested with respect to the first target-specific primer by at least 3 nucleotides, e.g. by 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more, or 15 or more nucleotides.

In some embodiments, all of the second target-specific primers present in a reaction comprise the same 5' portion. In some embodiments, the 5' portion the 5' portion of the can serve to suppress primer dimers as described for the 5' tag of the first target-specific primer described above herein.

In some embodiments, the first and second target-specific primers are substantially complementary to the same strand of the target nucleic acid. In some embodiments, the portions of the first and second target-specific primers that specifically anneal to the known target sequence can comprise a total of at least 20 unique bases of the known target nucleotide sequence, e.g. 20 or more unique bases, 25 or more unique bases, 30 or more unique bases, 35 or more unique bases, 40 or more unique bases, or 50 or more unique bases. In some embodiments, the portions of the first and second target-specific primers that specifically anneal to the known target sequence can comprise a total of at least 30 unique bases of the known target nucleotide sequence.

As used herein, the term "first adaptor primer" refers to a nucleic acid molecule comprising a nucleic acid sequence identical to a 5' portion of the first sequencing primer. As the first tail-adaptor primer is therefor identical to at least a

portion of the sequence of the amplification strand (as opposed to complementary), it will not be able to specifically anneal to any portion of the universal oligonucleotide tail-adaptor itself.

As used herein, the term "second adaptor primer" refers to a nucleic acid molecule comprising a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first adaptor primer. As the second tail-adaptor primer is therefor identical to at least a portion of the sequence of the amplification strand (as opposed to complementary), it will not be able to specifically anneal to any portion of the universal oligonucleotide tail-adaptor itself. In some embodiments, the second adaptor primer is identical to the first sequencing primer.

The second adaptor primer should be nested with respect to the first adaptor primer, that is, the first adaptor primer comprises a nucleic acid sequence identical to the amplification strand which is not comprised by the second adaptor primer and which is located closer to the 5' end of the amplification primer than any of the sequence identical to the amplification strand which is comprised by the second adaptor primer. In some embodiments, the second adaptor primer is nested by at least 3 nucleotides, e.g. by 3 nucleotides, by 4 nucleotides, by 5 nucleotides, by 6 nucleotides, by 7 nucleotides, by 8 nucleotides, by 9 nucleotides, by 10 nucleotides or more.

In some embodiments, the first adaptor primer can comprise a nucleic acid sequence identical to about the 20 5'-most bases of the amplification strand of the universal oligonucleotide tail-adaptor and the second adaptor primer can comprise a nucleic acid sequence identical to about 30 bases of the amplification strand of the universal oligonucleotide tail-adaptor, with a 5' base which is at least 3 nucleotides 3' of the 5' terminus of the amplification strand.

The use of nested adaptor primers eliminates the possibility of producing final amplicons that are amplifiable (e.g. during bridge PCR or emulsion PCR) but cannot be sequence, a situation that can arise during hemi-nested methods. In other situations, hemi-nested approaches using a primer identical to a sequencing primer can result in the carry-over of undesired amplification products from the first PCR step to the second PCR step and would ultimately yield artificial sequencing reads. The use of two adaptor primers, as described herein can reduce, and in some embodiments eliminate, these problems.

In the first PCR amplification cycle of the first amplification step, the first target-specific primer can specifically anneal to a template strand of any nucleic acid comprising the known target nucleotide sequence. Depending upon the orientation with which the first target-specific primer was designed, sequence upstream or downstream of the known target nucleotide sequence, and complementary to the template strand will be synthesized. If, during the extension phase of PCR, the 5' end of the template strand terminates in a ligated Universal oligonucleotide tail-adaptor, the 3' end of the newly synthesized product strand will comprise sequence complementary to the first tail-adaptor primer. In subsequent PCR amplification cycles, both the first target-specific primer and the first tail-adaptor primer will be able to specifically anneal to the appropriate strands of the target nucleic acid sequence and the sequence between the known nucleotide target sequence and the Universal oligonucleotide tail-adaptor can be amplified, (i.e. copied).

In the next step (i.e. step (c)) of the method described herein, a portion of the amplified portion resulting from step (b) is amplified in a second amplification step. The second amplification step can be a set of PCR amplification cycles using a second target-specific primer and a first sequencing primer. The second set of PCR amplification cycles can have PCR parameters identical to, or which differ from, those of the first set of PCR amplification cycles. E.g. the PCR amplification regimens of steps (b) and (c) can have the same or different annealing temperatures or the same or different extension step time lengths.

The methods described herein allow for determining the nucleotide sequence contiguous to a known target nucleotide sequence on either or both flanks of the known target nucleotide sequence. Regardless of whether the target nucleic acid normally exists as a single-stranded or double-stranded nucleic acid, sequence information is typically represented in a single-stranded format (Strand A), from 5' to 3'. If the sequence 5' of the known target nucleotide sequence of Strand A is to be determined, the gene-specific primers can be complementary (i.e. anneal to) Strand A. If the sequence 3' of the known target nucleotide sequence of Strand A is to be determined, the gene-specific primers can be identical to Strand A, such that they will anneal to the complementary strand of a double-stranded target nucleic acid. Such considerations of primer design are well known to those of ordinary skill in the art.

In some embodiments, the methods described herein, relating to the use of a first and second gene-specific primer can result in assays with a superior on-target rate, e.g. 70-90%. In some embodiments, the assays and methods described herein can have a target specificity rate of at least 85%.

In some embodiments, the four types of primers are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of from about 61 to 72° C., e.g. from about 61 to 69° C., from about 63 to 69° C., from about 63 to 67° C., from about 64 to 66° C. In some embodiments, the four types of primers are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of less than 72° C. In some embodiments, the four types of primers are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of less than 70° C. In some embodiments, the four types of primers are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of less than 68° C. In some embodiments, the four types of primers are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of about 65° C.

In some embodiments, the portions of the target-specific primers that specifically anneal to the known target nucleotide sequence will anneal specifically at a temperature of about 61 to 72° C., e.g. from about 61 to 69° C., from about 63 to 69° C., from about 63 to 67° C., from about 64 to 66° C. In some embodiments, the portions of the target-specific primers that specifically anneal to the known target nucleotide sequence will anneal specifically at a temperature of about 65° C. in a PCR buffer.

In some embodiments, the primers and/or adaptors described herein can not comprise modified bases (e.g. the primers and/or adaptors can not comprise a blocking 3' amine).

In the next step (i.e. step (d)) of the methods described herein, the amplified portion resulting from step (c) can be sequenced. In some embodiments, the sequencing can be performed by a next-generation sequencing method. As used herein "next-generation sequencing" refers to oligonucleotide sequencing technologies that have the capacity to sequence oligonucleotides at speeds above those possible

with conventional sequencing methods (e.g. Sanger sequencing), due to performing and reading out thousands to millions of sequencing reactions in parallel. Non-limiting examples of next-generation sequencing methods/platforms include Massively Parallel Signature Sequencing (Lynx Therapeutics); 454 pyro-sequencing (454 Life Sciences/Roche Diagnostics); solid-phase, reversible dye-terminator sequencing (Solexa/Illumina); SOLiD technology (Applied Biosystems); Ion semiconductor sequencing (ION Torrent); DNA nanoball sequencing (Complete Genomics); and technologies available from Pacific Biosciences, Intelligen Biosystems, Oxford Nanopore Technologies, and Helicos Biosciences. In some embodiments, the sequencing primers can comprise portions compatible with the selected next-generation sequencing method. Next-generation sequencing technologies and the constraints and design parameters of associated sequencing primers are well known in the art (see, e.g. Shendure, et al., "Next-generation DNA sequencing," Nature. 2008, vol. 26. No. 10, 1135-1145; Mardis, "The impact of next-generation sequencing technology on genetics," Trends in Genetics, 2007, vol. 24, No. 3, pp. 133-141; Su, et al., "Next-generation sequencing and its applications in molecular diagnostics" Expert Rev Mol Diagn, 2011, 11(3):333-43; Zhang et al., "The impact of next-generation sequencing on genomics", J Genet Genomics, 2011, 38(3):95-109; (Nyren, P. et al. Anal Biochem 208: 17175 (1993); Bentley, D. R. Curr Opin Genet Dev 16:545-52 (2006); Strausberg, R. L., et al. Drug Disc Today 13:569-77 (2008); U.S. Pat. No. 7,282,337; U.S. Pat. No. 7,279,563; U.S. Pat. No. 7,226,720; U.S. Pat. No. 7,220,549; U.S. Pat. No. 7,169,560; U.S. Pat. No. 6,818,395; U.S. Pat. No. 6,911,345; US Pub. Nos. 2006/0252077; 2007/0070349; and 20070070349; which are incorporated by reference herein in their entireties).

In some embodiments, the sequencing step relies upon the use of a first and second sequencing primers. In some embodiments, the first and second sequencing primers are selected to be compatible with a next-generation sequencing method as described herein.

Methods of aligning sequencing reads to known sequence databases of genomic and/or cDNA sequences are well known in the art and software is commercially available for this process. In some embodiments, reads (less the sequencing primer and/or adaptor nucleotide sequence) which do not map, in their entirety, to wild-type sequence databases can be genomic rearrangements or large indel mutations. In some embodiments, reads (less the sequencing primer and/or adaptor nucleotide sequence) comprising sequences which map to multiple locations in the genome can be genomic rearrangements.

In some embodiments, target nucleic acids and/or amplification products thereof can be isolated from enzymes, primers, or buffer components before and/or after any of steps a-d. Methods for isolating nucleic acids are well known in the art. In some embodiments, the isolation can comprise Solid Phase Reversible Immobilization (SPRI) cleanup. Methods for SPRI cleanup are well known in the art and kits are commercially available, e.g. Agencourt AMPure XP—PCR Purification (Cat No. A63880, Beckman Coulter; Brea, Calif.). In some embodiments, enzymes can be inactivated by heat treatment.

In some embodiments, the target nucleic acid can be comprised by genomic DNA. In some embodiments, the target nucleic acid can be comprised by ribonucleic acid (RNA), e.g. mRNA. In some embodiments, the target nucleic acid can be comprised by cDNA. Many of the sequencing methods suitable for use in the methods described herein provide sequencing runs with optimal read lengths of tens to hundreds of nucleotide bases (e.g. Ion Torrent technology can produce read lengths of 200-400 bp). Target nucleic acids comprised, for example, by genomic DNA or mRNA, can be comprised by nucleic acid molecules which are substantially longer than this optimal read length. In order for the amplified nucleic acid portion resulting from step (c) to be of a suitable length for use in a particular sequencing technology, the average distance between the known target nucleotide sequence and an end of the target nucleic acid to which the Universal oligonucleotide tail-adaptor can be ligated should be as close to the optimal read length of the selected technology as possible. For example, if the optimal read-length of a given sequencing technology is 200 bp, then the nucleic acid molecules amplified in accordance with the methods described herein should have an average length of about 400 bp or less. Target nucleic acids comprised by, e.g., genomic DNA or mRNA, can be sheared, e.g. mechanically or enzymatically sheared, to generate fragments of any desired size prior to step (a). Non-limiting examples of mechanical shearing processes include sonication, nebulization, and AFA™ shearing technology available from Covaris (Woburn, Mass.). In some embodiments, a target nucleic acid comprised by genomic DNA can be mechanically sheared by sonication. In some embodiments, when the target nucleic acid is comprised by RNA, the sample can be subjected to a reverse transcriptase regimen to generate DNA template and the DNA template can then be sheared. In some embodiments, target RNA can be sheared before performing the reverse transcriptase regimen. In some embodiments, a sample comprising target RNA can be used in the methods described herein using total nucleic acids extracted from either fresh or degraded specimens; without the need of genomic DNA removal for cDNA sequencing; without the need of ribosomal RNA depletion for cDNA sequencing; without the need of mechanical or enzymatic shearing in any of the steps; by subjecting the RNA for double-stranded cDNA synthesis using random hexamers; and by subjecting the nucleic acid to end-repair, phosphorylation, and adenylation in a single tube.

In some embodiments, the known target nucleotide sequence can be comprised by a gene rearrangement. The methods described herein are suited for determining the presence and/or identity of a gene rearrangement as the identity of only one half of the gene rearrangement must be previously known (i.e. the half of the gene rearrangement which is to be targeted by the gene-specific primers). In some embodiments, the gene rearrangement can comprise an oncogene. In some embodiments, the gene rearrangement can comprise a fusion oncogene.

In some embodiments, the target nucleic acid can be comprised by a sample. In some embodiments, the target nucleic acid can be comprised by a sample obtained from a subject. In some embodiments a sample can be a diagnostic sample obtained from a subject. In some embodiments, a sample can further comprise proteins, cells, fluids, biological fluids, preservatives, and/or other substances. By way of non-limiting example, a sample can be a cheek swab, blood, serum, plasma, sputum, cerebROS1pinal fluid, urine, tears, alveolar isolates, pleural fluid, pericardial fluid, cyst fluid, tumor tissue, tissue, a biopsy, saliva, an aspirate, or combinations thereof. In some embodiments, a sample can be obtained by resection or biopsy.

In some embodiments, the sample can be obtained from a subject in need of treatment for cancer. In some embodiments, the sample can comprise a population of tumor cells, e.g. at least one tumor cell. In some embodiments, the

sample can comprise a tumor biopsy, including but not limited to, untreated biopsy tissue or treated biopsy tissue (e.g. formalin-fixed and/or paraffin-embedded biopsy tissue).

In some embodiments, the sample can be freshly collected. In some embodiments, the sample can be stored prior to being used in the methods and compositions described herein. In some embodiments, the sample is an untreated sample. As used herein, "untreated sample" refers to a biological sample that has not had any prior sample pre-treatment except for dilution and/or suspension in a solution. In some embodiments, a sample can be obtained from a subject and preserved or processed prior to being utilized in the methods and compositions described herein. By way of non-limiting example, a sample can be embedded in paraffin wax, refrigerated, or frozen. A frozen sample can be thawed before determining the presence of a nucleic acid according to the methods and compositions described herein. In some embodiments, the sample can be a processed or treated sample. Exemplary methods for treating or processing a sample include, but are not limited to, centrifugation, filtration, sonication, homogenization, heating, freezing and thawing, contacting with a preservative (e.g. anti-coagulant or nuclease inhibitor) and any combination thereof. In some embodiments, the sample can be treated with a chemical and/or biological reagent. Chemical and/or biological reagents can be employed to protect and/or maintain the stability of the sample or nucleic acid comprised by the sample during processing and/or storage. In addition, or alternatively, chemical and/or biological reagents can be employed to release nucleic acids from other components of the sample. By way of non-limiting example, a blood sample can be treated with an anti-coagulant prior to being utilized in the methods and compositions described herein. The skilled artisan is well aware of methods and processes for processing, preservation, or treatment of samples for nucleic acid analysis. In some embodiments, the sample can be a clarified fluid sample, for example, by centrifugation. In some embodiments, the sample can be clarified by low-speed centrifugation (e.g. 3,000×g or less) and collection of the supernatant comprising the clarified fluid sample.

In some embodiments, the nucleic acid present in a sample can be isolated, enriched, or purified prior to being utilized in the methods and compositions described herein. Methods of isolating, enriching, or purifying nucleic acids from a sample are well known to one of ordinary skill in the art. By way of non-limiting example, kits for isolation of genomic DNA from various sample types are commercially available (e.g. Catalog Nos. 51104, 51304, 56504, and 56404; Qiagen; Germantown, Md.).

The methods described herein can be used in multiplex techniques. In embodiments of the methods described herein, multiplex applications can include determining the nucleotide sequence contiguous to one or more known target nucleotide sequences. As used herein, "multiplex PCR" refers to a variant of PCR where simultaneous amplification of more than one target nucleic acid in one reaction vessel and subsequent determination of the sequence of the amplification products by using more than one set of first and second gene-specific primers. Multiplex can refer to the detection of between about 2-1,000 different target sequences in a single reaction. As used herein, multiplex refers to the detection of any range between 2-1,000, e.g., between 5-500, 25-1000, or 10-100 different target sequences in a single reaction, etc. The term "multiplex" as applied to PCR implies that there are primers specific for at least two different target sequences in the same PCR reaction.

In some embodiments, the target nucleic acids in a sample, or separate portions of a sample, can be amplified with a plurality of first and second target-specific primers. In some embodiments, the plurality of first and second target-specific primers can be present in a single reaction mixture, e.g. multiple amplification products can be produced in the same reaction mixture. In some embodiments, the plurality of sets of first and second target-specific primers can specifically anneal to known target sequences comprised by separate genes. In some embodiments, at least two sets of first and second target-specific primers can specifically anneal to different portions of a known target sequence. In some embodiments, at least two sets of first and second target-specific primers can specifically anneal to different portions of a known target sequence comprised by a single gene. In some embodiments, at least two sets of first and second target-specific primers can specifically anneal to different exons of a gene comprising a known target sequence. In some embodiments, the plurality of first target-specific primers can comprise identical 5' tag sequence portions.

In embodiments of the methods described herein, multiplex applications can include determining the nucleotide sequence contiguous to one or more known target nucleotide sequences in multiple samples in one sequencing reaction or sequencing run. The multiple samples can be of different origins, e.g. from different tissues and/or different subjects. In such embodiments, the universal oligonucleotide tail-adaptor can further comprise a barcode portion. In some embodiments, a universal oligonucleotide tail-adaptor with a unique barcode portion can be added to each sample and ligated to the nucleic acids therein; the samples can then be pooled after step (a). Each resulting sequencing read of an amplification product will therefor comprise a barcode identifying which sample comprised the original template nucleic acid from which the amplification product is derived. The use of barcode portions in next-generation sequencing applications is well known in the art and described, for example, in Margulies, M. et al. "Genome Sequencing in Microfabricated High-Density Picoliter Reactors", Nature, 437, 376-80 (2005); Mikkelsen, T. et al. "Genome-Wide Maps of Chromatin State in Pluripotent and Lineage-Committed Cells", Nature, 448, 553-60 (2007); McLaughlin, S. et al. "Whole-Genome Resequencing With Short Reads: Accurate Mutation Discovery With Mate Pairs and Quality Values", ASHG Annual Meeting (2007); Shendure J. et al. "Accurate Multiplex Polony Sequencing of an Evolved Bacterial Genome", Science, 309, 1728-32 (2005); Harris, T. et al. "Single-Molecule DNA Sequencing of a Viral Genome", Science, 320, 106-9 (2008); Simen, B. et al. "Prevalence of LoW Abundance Drug Resistant Variants by Ultra Deep Sequencing in Chronically HIV-infected Antiretroviral (ARV) Naive Patients and the Impact on Virologic Outcomes", 16th International HIV Drug Resistance Workshop, Barbados (2007); Thomas, R. et al. "Sensitive Mutation Detection in Heterogeneous Cancer Specimens by Massively Parallel Picoliter Reactor Sequencing", Nature Med., 12, 852-855 (2006); Mitsuya, Y et al. "Minority Human Immunodeficiency Virus Type 1 Variants in Antiretroviral-Naive Persons With Reverse Transcriptase Codon 215 Revertant Mutations", J. Vir., 82, 10747-10755 (2008); Binladen, J. et al. "The Use of Coded PCR Primers Enables High-Throughput Sequencing of Multiple HomologAmplification Products by 454 Parallel Sequencing", PLoS ONE,

2, e197 (2007); and Hoffmann. C. et al. "DNA Bar Coding and PyROS1equencing to Identify Rare HIV Drug Resistance Mutations". Nuc. Acids Res., 35. e91 (2007). all of which are herein incorporated by reference.

In some embodiments of the technology described herein, determining the sequence contiguous to a known oligonucleotide target sequence can provide information relevant to treatment of disease, and/or can be comprised by a method of treating disease. In some embodiments, the sample can be from a subject in need of treatment for a disease associated with a genetic alteration. In some embodiments, the known oligonucleotide target sequence can be comprised by a disease-associated gene, e.g. an oncogene. In some embodiments, the sequence contiguous to a known oligonucleotide target sequence and/or the known oligonucleotide target sequence can comprise a mutation or genetic abnormality which is disease-associated, e.g. a SNP, an insertion, a deletion, and/or a gene rearrangement. In some embodiments, the sequence contiguous to a known oligonucleotide target sequence and/or the known oligonucleotide target sequence present in a sample can be comprised by a gene rearrangement product. In some embodiments, the gene rearrangement can be an oncogene, e.g. a fusion oncogene.

Certain treatments for cancer are particularly effective against tumors comprising certain oncogenes. e.g. a treatment agent which targets the action or expression of a given fusion oncogene can be effective against tumors comprising that fusion oncogene but not against tumors lacking the fusion oncogene. The methods described herein can allow the determination of specific sequences which reveal oncogene status (e.g. mutations. SNPs, and/or rearrangements). As described herein, the methods described herein can further allow the determination of specific sequences when the sequence of only one flank is known, e.g. the methods described herein can determine the presence and identity of gene rearrangements involving known oncogenes where the precise location and/or rearrangement partner are not known before the methods described herein are performed.

In some embodiments. the technology described herein relates to a method of treating cancer, the method comprising; detecting, in a tumor sample obtained from a subject in need of treatment for cancer, the presence of one or more oncogene rearrangements according to the method described herein; administering a cancer treatment which is effective against tumors having any of the detected oncogene rearrangements. In some embodiments, the technology described herein relates to a method of determining if a subject in need of treatment for cancer will be responsive to a given treatment. the method comprising; detecting, in a tumor sample obtained from the subject, the presence of an oncogene rearrangement according to the method as described herein; wherein the subject is determined to be responsive to a treatment targeting an oncogene rearrangement product if the presence of the oncogene rearrangement is detected.

In some embodiments, e.g. when the sample is obtained from a subject in need of treatment for lung cancer. the known oligonucleotide target sequence can comprise sequence from a gene selected from the group of ALK; ROS1; and RET. Gene rearrangements involving the ALK, ROS1, and RET genes and which result in fusion oncogenes are well known in the art (see, e.g. Soda et al. Nature 2007 448561-6: Rikova et al. Cell 2007 131:1190-1203; Kohno et al. Nature Medicine 2012 18:375-7; Takouchi et al. Nature Medicine 2012 18:378-81; which are incorporated by reference herein in their entireties. However, the precise location of the gene rearrangement (e.g. where in the ALK, ROS1, and/or RET gene the rearrangement has occurred), and the identity of the second gene involved in the rearrangement can vary. In the methods described herein, the presence and identity of such a rearrangement can be detected without having to know the location of the rearrangement or the identity of the second gene involved in the gene rearrangement.

In some embodiments, the known target sequence can comprise sequence from a gene selected from the group of: ALK; ROS1; and RET. In some embodiments, at least one set of a first target-specific primer and a second target-specific primer can be selected from the group consisting of: SEQ ID NOs: 5 and 6; SEQ ID NOs: 7 and 8; SEQ ID NOs: 9 and 10; SEQ ID NOs: 11 and 12; SEQ ID NOs: 13 and 14; SEQ ID NOs: 15 and 16; SEQ ID NOs: 17 and 18; SEQ ID NOs: 19 and 20; SEQ ID NOs: 21 and 22; SEQ ID NOs: 23 and 24; SEQ ID NOs: 25 and 26; SEQ ID NOs: 27 and 28; SEQ ID NOs: 29 and 30; SEQ ID NOs: 31 and 32; SEQ ID NOs: 33 and 34; SEQ ID NOs: 35 and 36; and SEQ ID NOs: 37 and 38.

In some embodiments, the presence of a gene rearrangement of ALK in a sample obtained from a tumor in a subject can indicate that the tumor is susceptible to treatment with a treatment selected from the group consisting of: an ALK inhibitor; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; diamino and aminopyrimidine inhibitors of ALK kinase activity such as NVP-TAE684 and PF-02341066 (see, e.g. Galkin et al. Proc Natl Acad Sci USA, 2007, 104:270-275; Zou et al. Cancer Res, 2007, 67:4408-4417; Hallberg and Palmer F1000 Med Reports 2011 3:21; and Sakamoto et al. Cancer Cell 2011 19:679-690) and molecules disclosed in WO 04/079326. All of the foregoing references are incorporated by reference herein in their entireties. An ALK inhibitor can include any agent that reduces the expression and/or kinase activity of ALK or a portion thereof, including, e.g. oligonucleotides, small molecules, and/or peptides that reduce the expression and/or activity of ALK or a portion thereof. As used herein "anaplastic lymphoma kinase" or "ALK" refers to a transmembrane tyROS1ine kinase typically involved in neuronal regulation in the wildtype form. The nucleotide sequence of the ALK gene and mRNA are known for a number of species, including human (e.g. SEQ ID NO: 2 (mRNA), NCBI Gene ID: 238).

In some embodiments, the presence of a gene rearrangement of ROS1 in a sample obtained from a tumor in a subject can indicate that the tumor is susceptible to treatment with a treatment selected from the group consisting of: a ROS1 inhibitor and an ALK inhibitor as described herein above (e.g. crizotinib). A ROS1 inhibitor can include any agent that reduces the expression and/or kinase activity of ROS1 or a portion thereof, including, e.g. oligonucleotides, small molecules, and/or peptides that reduce the expression and/or activity of ROS1 or a portion thereof. As used herein "c-ROS1 oncogene 1" or "ROS1" (also referred to in the art as ROS1-1) refers to a transmembrane tyROS1ine kinase of the sevenless subfamily and which interacts with PTPN6. The nucleotide sequence of the ROS1 gene and mRNA are known for a number of species, including human (e.g. SEQ ID NO: 1 (mRNA), NCBI Gene ID: 238).

In some embodiments, the presence of a gene rearrangement of RET in a sample obtained from a tumor in a subject can indicate that the tumor is susceptible to treatment with a treatment selected from the group consisting of: a RET inhibitor; DP-2490, DP-3636, SU5416; BAY 43-9006, BAY 73-4506 (regorafenib), ZD6474, NVP-AST487, sorafenib,

RPI-1, XL184, vandetanib, sunitinib, imatinib, pazopanib, axitinib, motesanib, gefitinib, and withaferin A (see, e.g. Samadi et al. Surgery 2010 148:1228-36; Cuccuru et al. JNCI 2004 13:1006-1014; Akeno-Stuart et al. Cancer Research 2007 67:6956; Grazma et al. J Clin Oncol 2010 28:15s 5559; Mologni et al. J Mol Endocrinol 2006 37:199-212; Calmomagno et al. Journal NCI 2006 98:326-334; Mologni. Curr Med Chem 2011 18:162-175 and the compounds disclosed in WO 06/034833; US Patent Publication 2011/0201598 and U.S. Pat. No. 8,067,434). All of the foregoing references are incorporated by reference herein in their entireties. A RET inhibitor can include any agent that reduces the expression and/or kinase activity of RET or a portion thereof, including, e.g. oligonucleotides, small molecules, and/or peptides that reduce the expression and/or activity of RET or a portion thereof. As used herein "rearranged during transfection" or "RET" refers to a receptor tyROS1ine kinase of the cadherein superfamily which is involved in neural crest development and recognizes glial cell line-derived neurotrophic factor family signaling molecules. The nucleotide sequence of the ROS1 gene and mRNA are known for a number of species, including human (e.g. SEQ ID NOs: 3-4 (mRNA), NCBI Gene ID: 5979).

Further non-limiting examples of applications of the methods described herein include detection of hematological malignancy markers and panels thereof (e.g. including those to detect genomic rearrangements in lymphomas and leukemias), detection of sarcoma-related genomic rearrangements and panels thereof; and detection of IGH/TCR gene rearrangements and panels thereof for lymphoma testing.

In some embodiments, the methods described herein relate to treating a subject having or diagnosed as having, e.g. cancer with a treatment for cancer. Subjects having cancer can be identified by a physician using current methods of diagnosing cancer. For example, symptoms and/or complications of lung cancer which characterize these conditions and aid in diagnosis are well known in the art and include but are not limited to, weak breathing, swollen lymph nodes above the collarbone, abnormal sounds in the lungs, dullness when the chest is tapped, and chest pain. Tests that may aid in a diagnosis of, e.g. lung cancer include, but are not limited to, x-rays, blood tests for high levels of certain substances (e.g. calcium), CT scans, and tumor biopsy. A family history of lung cancer, or exposure to risk factors for lung cancer (e.g. smoking or exposure to smoke and/or air pollution) can also aid in determining if a subject is likely to have lung cancer or in making a diagnosis of lung cancer.

Cancer can include, but is not limited to, carcinoma, including adenocarcinoma, lymphoma, blastoma, melanoma, sarcoma, leukemia, squamous cell cancer, small-cell lung cancer, non-small cell lung cancer, gastrointestinal cancer, Hodgkin's and non Hodgkin's lymphoma, pancreatic cancer, glioblastoma, basal cell carcinoma, biliary tract cancer, bladder cancer, brain cancer including glioblastomas and medulloblastomas; breast cancer, cervical cancer, choriocarcinoma; colon cancer, colorectal cancer, endometrial carcinoma, endometrial cancer; esophageal cancer, gastric cancer; various types of head and neck cancers, intraepithelial neoplasms including Bowen's disease and Paget's disease; hematological neoplasms including acute lymphocytic and myelogenous leukemia; Kaposi's sarcoma, hairy cell leukemia; chromic myelogenous leukemia, AIDS-associated leukemias and adult T-cell leukemia lymphoma; kidney cancer such as renal cell carcinoma, T-cell acute lymphoblastic leukemia/lymphoma, lymphomas including Hodgkin's disease and lymphocytic lymphomas; liver cancer such as hepatic carcinoma and hepatoma, Merkel cell carcinoma, melanoma, multiple myeloma; neuroblastomas; oral cancer including squamous cell carcinoma; ovarian cancer including those arising from epithelial cells, sarcomas including leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibROS1arcoma, and osteosarcoma; pancreatic cancer; skin cancer including melanoma, stromal cells, germ cells and mesenchymal cells; pROS1tate cancer, rectal cancer; vulval cancer, renal cancer including adenocarcinoma; testicular cancer including germinal tumors such as seminoma, non-seminoma (teratomas, choriocarcinomas), stromal tumors, and germ cell tumors; thyroid cancer including thyroid adenocarcinoma and medullar carcinoma; esophageal cancer, salivary gland carcinoma, and Wilms' tumors. In some embodiments, the cancer can be lung cancer.

In some embodiments, the methods described herein comprise administering an effective amount of compositions described herein, e.g. a treatment for cancer to a subject in order to alleviate a symptom of a cancer. As used herein, "alleviating a symptom of a cancer" is ameliorating any condition or symptom associated with the cancer. As compared with an equivalent untreated control, such reduction is by at least 5%, 10%, 20%, 40%, 50%, 60%, 80%, 90%, 95%, 99% or more as measured by any standard technique. A variety of means for administering the compositions described herein to subjects are known to those of skill in the art. Such methods can include, but are not limited to oral, parenteral, intravenous, intramuscular, subcutaneous, transdermal, airway (acROS1o1), pulmonary, cutaneous, topical, injection, or intratumoral administration. Administration can be local or systemic. The term "effective amount" as used herein refers to the amount of a treatment needed to alleviate at least one or more symptom of the disease or disorder, and relates to a sufficient amount of pharmacological composition to provide the desired effect. The term "therapeutically effective amount" therefore refers to an amount that is sufficient to effect a particular anti-cancer effect when administered to a typical subject. An effective amount as used herein, in various contexts, would also include an amount sufficient to delay the development of a symptom of the disease, alter the course of a symptom disease (for example but not limited to, slowing the progression of a symptom of the disease), or reverse a symptom of the disease. Thus, it is not generally practicable to specify an exact "effective amount". However, for any given case, an appropriate "effective amount" can be determined by one of ordinary skill in the art using only routine experimentation. The effects of any particular dosage can be monitored by a suitable bioassay. The dosage can be determined by a physician and adjusted, as necessary, to suit observed effects of the treatment.

Non-limiting examples of a treatment for cancer can include radiation therapy, surgery, gemcitabine, cisplastin, paclitaxel, carboplatin, bortezomib, AMG479, vorinostat, rituximab, temozolomide, rapamycin, ABT-737, PI-103; alkylating agents such as thiotepa and CYTOXAN® cyclosphosphamide; alkyl sulfonates such as busulfan, impROS1ulfan and piposulfan; aziridines such as benzodopa, carboquone, meturedopa, and uredopa; ethylenimines and methylamelamines including altretamine, triethylenemelamine, trietylenephosphoramide, triethiylenethiophosphoramide and trimethylolomelamine; acetogenins (especially bullatacin and bullatacinone); a camptothecin (including the synthetic analogue topotecan); bryostatin; callystatin; CC-1065 (including its adozelesin, carzelesin and bizelesin synthetic analogues); cryptophycins (particu-

larly cryptophycin 1 and cryptophycin 8); dolastatin; duocarmycin (including the synthetic analogues, KW-2189 and CB1-TM1); eleutherobin; pancratistatin; a sarcodictyin; spongistatin; nitrogen mustards such as chlorambucil, chlomaphazine, cholophosphamide, estramustine, ifosfamide, mechlorethamine, mechlorethamine oxide hydrochloride, melphalan, novembichin, phenesterine, prednimustine, trofosfamide, uracil mustard; nitROSUreas such as carmustine, chlorozotocin, fotemustine, lomustine, nimustine, and ranimnustine; antibiotics such as the enediyne antibiotics (e.g., calicheamicin, especially calicheamicin gammaII and calicheamicin omegaI1 (see, e.g., Agnew, Chem. Intl. Ed. Engl., 33: 183-186 (1994)); dynemicin, including dynemicin A; bisphosphonates, such as clodronate; an esperamicin; as well as neocarzinostatin chromophore and related chromoprotein enediyne antiobiotic chromophores), aclacinomysins, actinomycin, authramycin, azaserine, bleomycins, cactinomycin, carabicin, caminomycin, carzinophilin, chromomycinis, dactinomycin, daunorubicin, detorubicin, 6-diazo-5-oxo-L-norleucine, ADRIAMYCIN® doxorubicin (including morpholino-doxorubicin, cyanomorpholino-doxorubicin, 2-pyrrolino-doxorubicin and deoxydoxorubicin), epirubicin, esorubicin, idarubicin, marcellomycin, mitomycins such as mitomycin C, mycophenolic acid, nogalamycin, olivomycins, peplomycin, potfiromycin, puromycin, quelamycin, rodorubicin, streptonigrin, streptozocin, tubercidin, ubenimex, zinostatin, zorubicin; anti-metabolites such as methotrexate and 5-fluorouracil (5-FU); folic acid analogues such as denopterin, methotrexate, pteropterin, trimetrexate; purine analogs such as fludarabine, 6-mercaptopurine, thiamiprine, thioguanine; pyrimidine analogs such as ancitabine, azacitidine, 6-azauridine, carmofur, cytarabine, dideoxyuridine, doxifluridine, enocitabine, floxuridine; androgens such as calusterone, dromostanolone propionate, epitiostanol, mepitiostane, testolactone; anti-adrenals such as aminoglutethimide, mitotane, trilostane; folic acid replenisher such as frolinic acid; aceglatone; aldophosphamide glycoside; aminolevulinic acid; eniluracil; amsacrine; bestrabucil; bisantrene; edatraxate; defofamine; demecolcine; diaziquone; elformithine; elliptinium acetate; an epothilone; etoglucid; gallium nitrate; hydroxyurea; lentinan; lonidainine; maytansinoids such as maytansine and ansamitocins; mitoguazone; mitoxantrone; mopidanmol; nitracrine; pentostatin; phenamet; pirarubicin; losoxantrone; podophyllinic acid; 2-ethylhydrazide; procarbazine; PSK® polysaccharide complex (JHS Natural Products, Eugene, Oreg.); razoxane; rhizoxin; sizofuran; spirogermanium; tenuazonic acid; triaziquone; 2,2',2"-trichlorotriethylamine; trichothecenes (especially T-2 toxin, verracurin A, roridin A and anguidine); urethan; vindesine; dacarbazine; mannomustine; mitobronitol; mitolactol; pipobroman; gacytosine; arabinoside ("Ara-C"); cyclophosphamide; thiotepa; taxoids, e.g., TAXOL® paclitaxel (Bristol-Myers Squibb Oncology, Princeton, N.J.), ABRAXANE® Cremophor-free, albumin-engineered nanoparticle formulation of paclitaxel (American Pharmaceutical Partners, Schaumberg, Ill.), and TAXOTERE® doxetaxel (Rhone-Poulenc Rorer, Antony, France); chloranbucil; GEMZAR® gemcitabine; 6-thioguanine; mercaptopurine; methotrexate; platinum analogs such as cisplatin, oxaliplatin and carboplatin; vinblastine; platinum; etoposide (VP-16); ifosfamide; mitoxantrone; vincristine; NAVELBINE® vinorelbine; novantrone; teniposide; edatrexate; daunomycin; aminopterin; xeloda; ibandronate; irinotecan (Camptosar, CPT-11) (including the treatment regimen of irinotecan with 5-FU and leucovorin); topoisomerase inhibitor RFS 2000; difluoromethylomithine (DMFO); retinoids such as retinoic acid; capecitabine; combretastatin; leucovorin (LV); oxaliplatin, including the oxaliplatin treatment regimen (FOLFOX); lapatinib (Tykerb®); inhibitors of PKC-alpha, Raf, H-Ras, EGFR (e.g., erlotinib (Tarceva®)) and VEGF-A that reduce cell proliferation and pharmaceutically acceptable salts, acids or derivatives of any of the above. In addition, the methods of treatment can further include the use of radiation or radiation therapy. Further, the methods of treatment can further include the use of surgical treatments.

In some embodiments, the methods described herein can be applicable for resequencing, e.g. for confirming particularly relevant, low-quality, and/or complex sequences obtained by non-directed sequencing of a large amount of nucleic acids. By way of non-limiting examples, the methods described herein can allow the directed and/or targeted resequencing of targeted disease gene panels (e.g. 10-100 genes), resequencing to confirm variants obtained in large scale sequencing projects, whole exome resequencing, and/or targeted resequencing for detection of single nucleotide variants, multiple nucleotide variants, insertions, deletions, copy number changes, and methylation status.

In some embodiments, the methods described herein can allow microbiota sequencing, ancient sample sequencing, and/or new variant virus genotyping.

All patents and other publications: including literature references, issued patents, published patent applications, and co-pending patent applications; cited throughout this application are expressly incorporated herein by reference for the purpose of describing and disclosing, for example, the methodologies described in such publications that might be used in connection with the technology described herein. These publications are provided solely for their disclosure prior to the filing date of the present application. Nothing in this regard should be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention or for any other reason. All statements as to the date or representation as to the contents of these documents is based on the information available to the applicants and does not constitute any admission as to the correctness of the dates or contents of these documents.

The description of embodiments of the disclosure is not intended to be exhaustive or to limit the disclosure to the precise form disclosed. While specific embodiments of, and examples for, the disclosure are described herein for illustrative purposes, various equivalent modifications are possible within the scope of the disclosure, as those skilled in the relevant art will recognize. For example, while method steps or functions are presented in a given order, alternative embodiments may perform functions in a different order, or functions may be performed substantially concurrently. The teachings of the disclosure provided herein can be applied to other procedures or methods as appropriate. The various embodiments described herein can be combined to provide further embodiments. Aspects of the disclosure can be modified, if necessary, to employ the compositions, functions and concepts of the above references and application to provide yet further embodiments of the disclosure. These and other changes can be made to the disclosure in light of the detailed description. All such modifications are intended to be included within the scope of the appended claims.

Specific elements of any of the foregoing embodiments can be combined or substituted for elements in other embodiments. Furthermore, while advantages associated with certain embodiments of the disclosure have been described in the context of these embodiments, other embodiments may also exhibit such advantages, and not all

embodiments need necessarily exhibit such advantages to fall within the scope of the disclosure.

The technology described herein is further illustrated by the following examples which in no way should be construed as being further limiting.

Some embodiments of the technology described herein can be defined according to any of the following numbered paragraphs:

1. A method of determining the nucleotide sequence contiguous to a known target nucleotide sequence, the method comprising:
   (a) ligating a target nucleic acid comprising the known target nucleotide sequence with a universal oligonucleotide tail-adaptor;
   (b) amplifying a portion of the target nucleic acid and the amplification strand of the universal oligonucleotide tail-adaptor with a first adaptor primer and a first target-specific primer;
   (c) amplifying a portion of the amplicon resulting from step (b) with a second adaptor primer and a second target-specific primer;
   (d) sequencing the amplified portion from step (c) using a first and second sequencing primer;
   wherein the universal oligonucleotide tail-adaptor comprises a first ligatable duplex end and a second unpaired end;
   wherein the universal oligonucleotide tail-adaptor comprises a blocking strand and an amplification strand;
   wherein the blocking strand comprises a 5' duplex portion;
   wherein the amplification strand comprises an unpaired 5' portion, a 3' duplex portion, and a 3' T overhang;
   wherein the amplification strand comprises nucleic acid sequences identical to a first and second sequencing primers;
   wherein the duplex portions of the blocking strand and the amplification strand are substantially complementary and form the first ligatable duplex end comprising a 3' T overhang;
   wherein the duplex portion is of sufficient length to remain in duplex form at the ligation temperature;
   wherein the first target-specific primer comprises a nucleic acid sequence that can specifically anneal to the known target nucleotide sequence of the target nucleic acid at the annealing temperature;
   wherein the second target-specific primer comprises a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (b), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer and the second target-specific primer is nested with respect to the first target-specific primer;
   wherein the first adaptor primer comprises a nucleic acid sequence identical to a 5' portion of the first sequencing primer; and
   wherein the second adaptor primer comprises a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first adaptor primer.
2. The method of paragraph 1, wherein the blocking strand of the universal oligonucleotide tail-adaptor further comprises a 3' unpaired portion which is not substantially complementary to the 5' unpaired portion of the amplification strand; and
   wherein the 3' unpaired portion of the blocking strand is not substantially complementary to or substantially identical to any of the primers.
3. The method of any of paragraphs 1-2, wherein the second adaptor primer is nested with respect to the first adaptor primer by at least 3 nucleotides.
4. The method of any of paragraphs 1-3, wherein the portion of the amplification strand that comprises a nucleic acid sequence identical to a first and second sequencing primers is comprised, at least in part, by the 5' unpaired portion of the amplification strand.
5. The method of any of paragraphs 1-4, wherein the first target-specific primer further comprises a 5' tag sequence portion comprising a nucleic acid sequence of high GC content which is not substantially complementary to or substantially identical to any other portion of any of the primers.
6. The method of any of paragraphs 1-5, wherein the second adaptor primer is identical to the full-length first sequencing primer.
7. The method of any of paragraphs 1-6, wherein the portions of the target-specific primers that specifically anneal to the known target will anneal specifically at a temperature of about 65° C. in a PCR buffer.
8. The method of any of paragraphs 1-7, wherein the method further comprises; prior to step (a), the steps of:
   mechanically shearing the nucleic acid;
   subjecting the nucleic acid to end-repair;
   subjecting the nucleic acid to phosphorylation;
   and subjecting the nucleic acid to adenylation.
9. The method of any of paragraphs 1-8, wherein the sample comprises genomic DNA.
10. The method of any of paragraphs 1-9, wherein the sample comprises RNA and the method further comprises a first step of subjecting the sample to a reverse transcriptase regimen.
11. The method of any of paragraphs 1-10, wherein the reverse transcriptase regimen comprises the use of random hexamers.
12. The method of any of paragraphs 1-11, wherein the known target sequence is comprised by a gene rearrangement.
13. The method of paragraph 12 wherein the gene rearrangement is present in a nucleic acid selected from the group consisting of:
    genomic DNA; RNA; and cDNA.
14. The method of any of paragraphs 12-13, wherein the gene rearrangement comprises an oncogene.
15. The method of paragraph 14, wherein the gene rearrangement comprises a fusion oncogene.
16. The method of any of paragraphs 1-15, wherein the nucleic acid product is sequenced by a next-generation sequencing method.
17. The method of paragraph 16, wherein the next-generation sequencing method comprises a method selected from the group consisting of:
    Ion Torrent, Illumina, SOLiD, 454; Massively Parallel Signature Sequencing solid-phase, reversible dye-terminator sequencing; and DNA nanoball sequencing.
18. The method of any of paragraphs 1-17, wherein the first and second sequencing primers are compatible with the selected next-generation sequencing method.
19. The method of any of paragraphs 1-18, wherein the method comprises contacting the sample, or separate

portions of the sample, with a plurality of sets of first and second target-specific primers.

20. The method of any of paragraphs 1-19, wherein the method comprises contacting a single reaction mixture comprising the sample with a plurality of sets of first and second target-specific primers.

21. The method of any of paragraphs 1-20, wherein the plurality of sets of first and second target-specific primers specifically anneal to known target nucleotide sequences comprised by separate genes.

22. The method of any of paragraphs 1-21, wherein at least two sets of first and second target-specific primers specifically anneal to different portions of a known target nucleotide sequence.

23. The method of any of paragraphs 1-22, wherein at least two sets of first and second target-specific primers specifically anneal to different portions of a single gene comprising a known target nucleotide sequence.

24. The method of any of paragraphs 1-23, wherein at least two sets of first and second target-specific primers specifically anneal to different exons of a gene comprising a known nucleotide target sequence.

25. The method of any of paragraphs 19-24, wherein the plurality of first target-specific primers comprise identical 5' tag sequence portions.

26. The method of any of paragraphs 1-25, wherein the universal oligonucleotide tail-adaptor further comprises a barcode portion.

27. The method of paragraph 26, wherein multiple samples are each contacted with a universal oligonucleotide tail-adaptor with a unique barcode portion and wherein the samples are pooled after step (a).

28. The method of any of paragraphs 1-27, wherein each amplification step comprises a set of cycles of a PCR amplification regimen from 5 cycles to 20 cycles in length.

29. The method of any of paragraphs 1-28, wherein the target-specific primers and the adaptor primers are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of from about 61 to 72° C.

30. The method of any of paragraphs 1-29, wherein the target-specific primers and the adaptor primers are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of about 65° C.

31. The method of any of paragraphs 1-30, wherein the sample comprises a biological sample obtained from a subject.

32. The method of any of paragraphs 1-31, wherein the sample is obtained from a subject in need of treatment for a disease associated with a genetic alteration.

33. The method of paragraph 32, wherein the disease is cancer.

34. The method of any of paragraphs 1-33, wherein the sample comprises a population of tumor cells.

35. The method of any of paragraphs 1-34, wherein the sample comprises a tumor biopsy.

36. The method of any of paragraphs 1-35, wherein the cancer is lung cancer.

37. The method of any of paragraphs 1-36, wherein the known target sequence is comprised by a disease-associated gene.

38. The method of any of paragraphs 1-37, wherein the known target sequence is comprised by a gene rearrangement product in the sample.

39. The method of any of paragraphs 1-38, wherein gene rearrangement product is an oncogene.

40. The method of any of paragraphs 1-39, wherein the known target sequence comprises sequence from a gene selected from the group of:
ALK; ROS1; and RET.

41. The method of paragraph 40, wherein at least one set of a first target-specific primer and a second target-specific primer are selected from the group consisting of:
SEQ ID NOs: 5 and 6; SEQ ID NOs: 7 and 8; SEQ ID NOs: 9 and 10; SEQ ID NOs: 11 and 12; SEQ ID NOs: 13 and 14; SEQ ID NOs: 15 and 16; SEQ ID NOs: 17 and 18; SEQ ID NOs: 19 and 20; SEQ ID NOs: 21 and 22; SEQ ID NOs: 23 and 24; SEQ ID NOs: 25 and 26; SEQ ID NOs: 27 and 28; SEQ ID NOs: 29 and 30; SEQ ID NOs: 31 and 32; SEQ ID NOs: 33 and 34; SEQ ID NOs: 35 and 36; and SEQ ID NOs: 37 and 38.

42. The method of any of paragraphs 40-41, wherein the presence of a gene rearrangement of ALK in a sample obtained from a tumor in a subject indicates that the tumor is susceptible to treatment with a treatment selected from the group consisting of:
an ALK inhibitor; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; and NVP-TAE684.

43. The method of any of paragraphs 40-41, wherein the presence of a gene rearrangement of ROS1 in a sample obtained from a tumor in a subject indicates that the tumor is susceptible to treatment with a treatment selected from the group consisting of:
a ROS inhibitor; an ALK inhibitor; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; and NVP-TAE684.

44. The method of any of paragraphs 40-41, wherein the presence of a gene rearrangement of RET in a sample obtained from a tumor in a subject indicates that the tumor is susceptible to treatment with a treatment selected from the group consisting of:
a RET inhibitor; DP-2490; DP-3636; SU5416; BAY 43-9006; BAY 73-4506 (regorafenib); ZD6474; NVP-AST487; sorafenib; RPI-1; XL184; vandetanib; sunitinib; imatinib; pazopanib; axitinib; motesanib; gefitinib; and withaferin A.

45. A method of treating cancer, the method comprising;
detecting, in a tumor sample obtained from a subject in need of treatment for cancer, the presence of one or more oncogene rearrangements according to the method of any of paragraphs 1-44;
administering a cancer treatment which is effective against tumors having any of the detected oncogene rearrangements.

46. The method of paragraph 45, wherein a treatment selected from the group consisting of: an ALK inhibitor; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; and NVP-TAE684;
is effective against tumors having an ALK oncogene rearrangement.

47. The method of paragraph 45, wherein a treatment selected from the group consisting of:
a ROS1 inhibitor; an ALK inhibitor; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-

504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; and NVP-TAE684;

is effective against tumors having an ROS1 oncogene rearrangement.

48. The method of paragraph 45, wherein a treatment selected from the group consisting of:

a RET inhibitor; DP-2490; DP-3636; SU5416; BAY 43-9006; BAY 73-4506 (regorafenib); ZD6474; NVP-AST487; sorafenib; RPI-1; XL184; vandetanib; sunitinib; imatinib; pazopanib; axitinib; motesanib; gefitinib; and withaferin A;

is effective against tumors having an RET oncogene rearrangement.

49. A method of determining if a subject in need of treatment for cancer will be responsive to a given treatment, the method comprising:

detecting, in a tumor sample obtained from the subject, the presence of an oncogene rearrangement according to the method of any of paragraphs 1-44;

wherein the subject is determined to be responsive to a treatment targeting an oncogene rearrangement product if the presence of the oncogene rearrangement is detected.

50. The method of paragraph 49, wherein if the presence of an ALK oncogene rearrangement is detected, the subject will be responsive to a treatment selected from the group consisting of:

an ALK inhibitor; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; and NVP-TAE684.

51. The method of paragraph 49, wherein if the presence of an ROS1 oncogene rearrangement is detected, the subject will be responsive to a treatment selected from the group consisting of:

an ALK inhibitor; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; and NVP-TAE684.

52. The method of paragraph 49, wherein if the presence of an RET oncogene rearrangement is detected, the subject will be responsive to a treatment selected from the group consisting of:

a RET inhibitor; DP-2490; DP-3636; SU5416; BAY 43-9006; BAY 73-4506 (regorafenib); ZD6474; NVP-AST487; sorafenib; RPI-1; XL184; vandetanib; sunitinib; imatinib; pazopanib; axitinib; motesanib; gefitinib; and withaferin A.

53. The method of any of paragraphs 44-52, wherein the cancer is lung cancer.

EXAMPLES

Example 1

In one embodiment, described herein is an assay using a novel half-truncated "Y" shape adapter for next generation sequencing library construction followed by two single-end nested polymerase chain reactions enables rapid and efficient target enrichment using RNA or genomic DNA as template from fresh or formalin-fixed and paraffin-embedded (FFPE) specimens. This enrichment method enables targeted resequencing of cDNA or gDNA for potential detection of genetic alterations (single nucleotide variants, insertions/deletions, and copy number), epigenetic alterations (methylation), gene expression, and genomic rearrangements.

Target enrichment prior to next-generation sequencing is more cost-effective than whole genome, whole exome, and whole transcriptome sequencing and therefore more practical for broad implementation both for research discovery and clinical applications. High coverage depth afforded by target enrichment approaches enables wider dynamic range for allele counting (in gene expression and copy number assessment) and detection of low frequency mutations, a critical feature for evaluating somatic mutations in cancer. Before broad implementation of whole genome or whole exome sequencing is possible, the mainstay of clinical next generation sequencing will involve select disease panels with a discrete number of gene targets. Likewise, research studies for analysis of large sample sizes based on defined gene targets will also require an economical means for genotyping. The assays described herein would be applicable in both of these settings.

For the detection of inter and intra-chromosomal rearrangements, whole genome or transcriptome sequencing is useful for novel discovery of genomic rearrangements and does not require prior knowledge of the involved gene/chromosomal partners. However, these whole genome and transcriptome approaches are not practical currently in the clinical setting due to high cost, low sequencing depth resulting in poor sensitivity, and highly demanding bioinformatics analysis. Fluorescence in situ hybridization (FISH) has been the gold standard assay for the detection of genomic rearrangements in clinics; however, the assay is low-throughput and its implementation requires special equipment and expert knowledge/experience. RT-PCR is also effective at detecting such rearrangements but requires knowledge of both the 5' and 3' partners and is not scalable for a large number of targets/samples.

Examples of commonly used preparatory enrichment assays for next generation sequencing include hybridization-based capture assays (TruSeq Capture, Illumina; SureSelect Hybrid Capture, Agilent) and polymerase chain reaction (PCR)-based assays (HaloPlex, Agilent; AmpliSeq, Ion Torrent; TruSeq Amplicon, Illumina; emulsion/digital PCR, Raindance). Hybridization-based approaches capture not only the targeted sequences covered by the capture probes but also near off-target bases that consume additional sequencing capacity. In addition, these methods are relatively time-consuming, labor-intensive, and less specific. A PCR amplification based approach is simpler and faster but by conventional design requires both forward and reverse primers flanking target loci. In particular for detection of genomic rearrangements with unknown fusion partners, PCR is not applicable.

Described herein is a target enrichment assay using a novel half-truncated "Y" shape adapter for next generation sequencing library construction, enabling rapid sequencing of cDNA or gDNA from fresh or FFPE specimens. FIG. 1 outlines the library construction for target enrichment using a half-truncated Y adapter for Ion Torrent sequencing as an example. Importantly, this method is adaptable for any other next generation sequencing platform, including but not limited to Illumina, SOLiD, and 454. In short a randomly sheared, double-stranded cDNA or gDNA template can be end-repaired, adenylated, and then ligated on one or both ends with the universal Y adapter to create a universal sequence for initiating PCR and subsequent sequencing. An initial round of PCR using a target specific primer tagged with a stuffer tail (the stuff tail aids in multiplexing a high number of targets) and a primer same as the 20 bases at 5' Y adapter overhang. A second round of PCR is carried out with a primer same as the 30 bases at the 5' Y adapter

overhang and a second tandem nested target specific primer that anneals 3' downstream of the first target specific primer. The second tandem nested primer is 5' tagged with the second primer sequence required for downstream emulsion PCR or clustering depending on the sequencing technology. High specificity in this system is achieved with the unidirectional tandem. nested primers which will effectively cover a target sequence of 36-40 bases or more (depending on the spacing between the primers). Of note. the number of PCR cycles may be optimized depending on how much starting nucleic acid material is used. the number of pooled samples, and the number of targets.

In summary. the method utilizes a half-truncated Y adapter to universally tag all dsDNA fragments with a common 3' end which is utilized for two rounds of PCR using two unidirectional, nested target specific primers for specificity. The application of tagged nested primers also avoids the effects of primer homodimerization and heterodimerization when targeting many different sites in the genome or transcriptome. In addition to multiplexing targets, the methods described herein also allow for multiplex sample pooling after the Y adapter ligation (Step 2) during which individual samples are ligated with unique barcoded adapters. Once barcoded via Y adapter ligation, multiple samples may be pooled into one reaction tube for downstream target enrichment.

In contrast to detection of genomic rearrangements which is more tolerant of insertion and deletion sequencing errors such as those found on Ion Torrent and 454 sequencing as a result of homopolymers, detection of single base and multiple base mutations (including insertions and deletions) requires higher accuracy sequencing achievable with the Illumina sequencing platform. For this purpose, the half-truncated Y adapted library can be converted to an Illumina library, e.g. by amplifying the overhanging, truncated 3' arm of the Y adapter using a primer tagged with the Illumina forward adapter sequence (barcoded or non-barcoded), and tagged gene specific nested primers (GSP2s tagged with Illumina reversed primer). Similarly, Bi-directional sequencing is also achievable by introducing a sequencing primer via the overhanging, truncated 3' Y adapter arm using a tagged primer during the first PCR. Step 3.

The assays and methods described herein have been applied on cDNA derived from FFPE specimens for the detection of ROS1 gene fusions using seven gene specific primers targeting seven exons comprising the ROS1 kinase domain. ROS1 genomic rearrangements were detected with known and previously unknown fusion partners such as "SLC34A2 Exon 12—ROS1 Exon 34" and "EZR Exon 9—ROS1 Exon 34" fusions. The assay achieved high on-target specificity (~85-95% when mapping using human genome hg19 reference), which enables the sequencing of multiple samples even with the smallest scale sequencing platforms (and thus least expensive) such as the Ion Torrent PGM 314 chip with high coverage results (7 targeted loci, 5 samples, >1000x coverage for each target per sample). FIG. 2 shows mapping of sequencing reads to the target loci in gene ROS1 kinase domain.

Advantages of the methods described herein include:

1. The methods described herein can be utilized for target enrichment of double stranded cDNA or gDNA from fresh or FFPE samples, allowing mapping of both 5' and 3' ends of target cDNA or gDNA fragments, using next-generation sequencing. Current hybridization-based approaches require days of preparation and hybridization, and show lower specificity with capture of near off-target sequences. Current amplification based approaches require known forward and reverse primers by design and are not amenable for high scale multiplexing of targets.

2. Simple Bioinformatics analysis. Depending on the number of selected targets, analysis of data generated from this assay is simple and fast. In addition, it may be set up to be compatible with current existing bioinformatics tools. Thus, small-size clinical labs may be able to perform data analysis without the need for significant investment in bioinformatics, which will be a limitation for broad clinical implementation of next generation sequencing.

3. High specificity (~85-95%). A conventional Y adapter construct having both components of forward and reverse primers (5' and 3' overhangs) for routine library construction will introduce a high level of background off-target sequencing due to carry over of "sequenceable" starting genomic or transcriptome library material. For instance, if the target is of size of 100 bp and present in two copies in one genome, the ratio of target-to-genome is about 1:3×10^7. Hybridization-based approaches normally employ biotinylated oligonucleotide baits to pull out hybridized target fragments by streptavidin coated magnetic beads. Just one non-specific binding event out of the possible 3×10^7 possibilities could dramatically result in a 50% off-target rate. The use of a half-truncated Y adapter effectively avoids this background carry over issue, since the starting library material cannot be amplified in subsequent preparatory steps for sequencing. Additional specificity is achieved by using two unidirectional primers, with GSP2 3' downstream of GSP1. Effectively, the use of the two tandem primers yields a target priming site of 40 (assume two 20 base pair length primers) and higher specificity than with just one primer. The use of 5' 20-mer primer in the first PCR and the full 30-mer primer (serves as a nested primer for the universal priming site) in the second PCR further increases specificity. Finally, additional specificity is achieved with the use of tagged primers for both PCR steps. Tagged primers, through intramolecular hairpin formation, prevent the propagation of primer homodimers and heterodimers which could overtake the PCR reactions and lead to nonspecific and undesired artifactual products; therefore, tagged primers allow the ability to multiplex many targets while avoiding primer dimers.

4. Economical cost. The key components in the methods described herein are conventional, unmodified tagged primers, standard PCR reagents, and routine thermocycling. Unlike hybridization based capture methods or microfluidic digital PCR setups, the described target enrichment protocol avoids the use of relatively expensive biotinylated oligonucleotides, streptavidin coated magnetic beads, and special equipment. Once manufactured, pooled GSP1 and GSP2 primer mixes may be used for thousands of reactions.

5. Automation. Because the method described herein relies on standard PCR techniques and SPRI cleanup, there is potential for facile automation in high throughput applications. Volumes may be adjusted for 384 well plates for ultra-highthroughput implementation.

Applications of the methods and assays described herein include, but are not limited to: 1. A lung cancer translocation panel consisting of known therapeutic targets including genes ALK, ROS1 and RET; 2. hematological malignancy panels including those to detect genomic rearrangements in lymphomas and leukemias; 3. Sarcoma genomic rearrangement panel; 4. IGH/TCR gene rearrangement panel for lymphoma testing; 5. Targeted disease gene panel for resequencing (10-100 genes); 6. Targeted resequencing for confirming variants from large scale sequencing projects; 7. Potential for whole exome targeted resequencing; 8. Tar-

geted resequencing for detection of single nucleotide variants, multiple nucleotide variants, insertions, deletions, copy number changes, and methylation status; 9. Microbiota sequencing; 10. Ancient sample sequencing; 11. New variant virus genotyping.

Example 2: Targeted Next-Generation Sequencing Assay for Simultaneous Detection of ALK, ROS1 and RET Gene Rearrangements

Knowledge of chromosomal rearrangement status in cancer is important for individualized targeted therapy. Recently, three major receptor tyROS1ine kinases involved in rearrangements in lung cancer have been described. Gene rearrangements involving the ALK gene has been established as a therapeutic target. Early clinical trial data also suggest that a ROS1 inhibitor is effective in treating patients testing positive for the ROS1 rearrangement. In vitro evidence has shown that tumor cells harbouring the RET rearrangement are responsive to a RET inhibitor. Thus, ALK, ROS1, and RET currently represent three important therapeutic targets in lung cancer.

Current clinical assays for the detection of gene rearrangements include fluorescence in situ hybridization (FISH), immunohistochemistry (IHC), and reverse transcription polymerase chain reaction (RT-PCR). FISH and IHC may not be able to accommodate the increasing demand for high volume testing/screening due to their low throughput, high cost, and complex interpretation. RT-PCR assays require full knowledge of both fusion partners which are sometimes unavailable and effectively impact clinical sensitivity. Multiplex RT-PCR has been used for detection of variant fusions involving different exons between two known fusion partners. In general, these assays are limited to a one gene target at a time approach requiring multiple reaction setups and analysis.

The latest next-generation sequencing-based assays for transcriptome or target capture sequencing have been applied largely for the purpose of research and discovery in large sequencing facilities. These assays in the research setting generally achieve low sequencing depth due to the vast number of targets, thus yielding low analytical sensitivity and subsequent poor clinical sensitivity. Though affordable for large sequencing facilities, these assays are not yet within reach for many clinical laboratories.

Described above are methods and assays related to a targeted sequencing method which applies a novel half-truncated "Y" shape adapter for next generation sequencing library construction followed by two single-end nested polymerase chain reactions enabling rapid and efficient target enrichment using RNA or genomic DNA as template from fresh or formalin-fixed and paraffin-embedded (FFPE) specimens.

Based on this method, described in this Example is a specific assay for the detection of ALK, ROS1, and RET gene rearrangements (FIGS. 3A-3B). Gene specific primers (GSP1) are designed to prime the exons near or on the kinase domain. Nested primers (GSP2) are designed to prime downstream of GSP1 but within the same exons and proximity to their paired GSP1s. The panel currently includes seven pairs of primers for targeting ROS1 exons 31 to 37; four pairs of primers for targeting ALK exons 19 to 22; and six pairs of primers for targeting RET exons 8 to 13.

This assay can be adapted to different NGS sequencing platforms by using platform specific adaptor sequence following the half-functional "Y" adaptor configuration and the GSP2 primers. After sequencing, the reads are mapped to the human genome reference allowing the identification of fusion partner genes, alternative splicing, as well as frame status of the fusions using a bioinformatics algorithm developed by the inventors (e.g. FIGS. **4**A and **4**B). The resulting output is a simple annotated table for quick reporting (FIG. **4**C). Using this three gene target panel assay with multi-sample barcoding, ROS1, ALK and RET rearrangement positive samples have been detected using one Ion Torrent sequencing run.

This assay is applicable for degraded RNA, such as RNA extracted from formalin-fixed and paraffin-embedded (FFPE) specimen, which is the most widely available clinical material for molecular diagnostic testing. This assay takes advantage of bench-top NGS platforms, a simple informatics analysis pipeline, and therefore relatively easy implementation for clinical laboratories. Of note, this assay does not require prior knowledge of the fusion partners and would yield high clinical sensitivity in the detection of various gene partners (and corresponding multiple exons) associated with the principle target gene. In addition, the assay will be functional for both unknown 5' upstream or 3' downstream fusion partners as long as one of the partners is known. Multiplexing and deep sequencing allows testing of samples with low tumor cellularity, rare fusions, and rare alternative splicing events. The detailed, patient specific rearrangement information afforded by this assay would be useful for evaluating genotype specific therapeutic response and perhaps a patient specific tumor marker for minimal residual disease monitoring. Based on conventional oligo synthesis reagents, off the shelf enzymes, and the ability to multiplex many samples in one run, this assay will be a cost-effective clinical assay for detecting gene rearrangements.

TABLE 1

Primers adapted for IonTorrent platform, v1. Primer names indicate target gene and target gene exon comprising the known target nucleotide sequence. R1 designates a first target-specific primer and R2 designates a second target-specific primer. Detailed in this table is one set of a first target specific primer and one second target-specific for each listed exon of each gene.

Primer Name	Sequence	SEQ ID NO
ALKex19_P1R2	CCTCTCTATGGGCAGTCGGTGATGCGAGAGTGGCAGGTGTGG	5
ALKex19_tag.R1	GGATCTCGACGCTCTCCCTCAGAGGTCACCACAGAGAGGATCAG	6
ALKex20_P1R2	CCTCTCTATGGGCAGTCGGTGATCATGGCTTGCAGCTCCTGGT	7
ALKex20_tag.R1	GGATCTCGACGCTCTCCCTGCAGCTCCATCTGCATGGCTTG	8
ALKex21_P1R2	CCTCTCTATGGGCAGTCGGTGATGGCCTTCATACACCTCCCCAAA	9

TABLE 1-continued

Primers adapted for IonTorrent platform, v1. Primer names indicate target gene and target gene exon comprising the known target nucleotide sequence. R1 designates a first target-specific primer and R2 designates a second target-specific primer. Detailed in this table is one set of a first target specific primer and one second target-specific for each listed exon of each gene.

Primer Name	Sequence	SEQ ID NO
ALKex21_tag.R1	GGATCTCGACGCTCTCCCTTTGGGCATTCCGGACACC	10
ALKex22_P1R2	CCTCTCTATGGGCAGTCGGTGATAGGAAATCCAGTTCGTCCTGTTCAGA	11
ALKex22_tag.R1	GGATCTCGACGCTCTCCCTGATCAGGGCTTCCATGAGGAAATC	12
RETex10_P1R2	CCTCTCTATGGGCAGTCGGTGATGGCTCCCCAGGCTCGTGT	13
RETex10_tag.R1	GGATCTCGACGCTCTCCCTAGGTGCCATAGCCAGCTTTAATCC	14
RETex11_P1R2	CCTCTCTATGGGCAGTCGGTGATATCACCGTGCGGCACAGCTC	15
RETex11_tag.R1	GGATCTCGACGCTCTCCCTGAGGACAGCGGCTGCGATCA	16
RETex12_P1R2	CCTCTCTATGGGCAGTCGGTGATAGAACCAAGTTCTTCCGAGGGAAT	17
RETex12_tag.R1	GGATCTCGACGCTCTCCCTTCCAAATTCGCCTTCTCCTAGAGTT	18
RETex13_P1R2	CCTCTCTATGGGCAGTCGGTGATACAGCAGGTCTCGCAGCTCAC	19
RETex13_tag.R1	GGATCTCGACGCTCTCCCTTGACCTGCTTCAGGACGTTGAA	20
RETex8_P1R2	CCTCTCTATGGGCAGTCGGTGATCTTGCTGACTGCACAGGACAGG	21
RETex8_tag.R1	GGATCTCGACGCTCTCCCTTCCTCACACTCCAGCCGTCTC	22
RETex9_P1R2	CCTCTCTATGGGCAGTCGGTGATTGGTGCTGGGAGAGCAGGT	23
RETex9_tag.R1	GGATCTCGACGCTCTCCCTCCGTCGGGGCAGGTCTTG	24
ROS1Ex31_P1R2	CCTCTCTATGGGCAGTCGGTGATGGCTGCATGAAGTTTTAACATGG	25
ROS1Ex31_tag.R1	GGATCTCGACGCTCTCCCTTGATATTACAGACATAAGCAGGACCTTGG	26
ROS1Ex32_P1R2	CCTCTCTATGGGCAGTCGGTGATCTAGTAATTTGGGAATGCCTGGTTT	27
ROS1Ex32_tag.R1	GGATCTCGACGCTCTCCCTTTCAGCTTTCTCCCACTGTATTGAA	28
ROS1Ex33_P1R2	CCTCTCTATGGGCAGTCGGTGATCATCTTCCACCTTAAATTCTGGTTCTGTA	29
ROS1Ex33_tag.R1	GGATCTCGACGCTCTCCCTCAGGATCCATTAAATGTCATCTTCC	30
ROS1Ex34_P1R2	CCTCTCTATGGGCAGTCGGTGATAGTAAGTATGAAACTTGTTTCTGGTATCC	31
ROS1Ex34_tag.R1	GGATCTCGACGCTCTCCCTGGTCAGTGGGATTGTAACAACCAG	32
ROS1Ex35_P1R2	CCTCTCTATGGGCAGTCGGTGATCACCCCTTCCTTGGCACTTT	33
ROS1Ex35_tag.R1	GGATCTCGACGCTCTCCCTTCTTTGTCTTCGTTTATAAGCACTGTC	34
ROS1Ex36_P1R2	CCTCTCTATGGGCAGTCGGTGATTTCAATCTCCTCTTGGGTTGGA	35
ROS1Ex36_tag.R1	GGATCTCGACGCTCTCCCTCCGAGGGAAGGCAGGAAGATT	36
ROS1Ex37_P1R2	CCTCTCTATGGGCAGTCGGTGATCAGGAATTCAATCTTCTCCTGGTC	37
ROS1Ex37_tag.R1	GGATCTCGACGCTCTCCCTCTCATCAGATGTGCCTCCTTCAG	38

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 48

<210> SEQ ID NO 1
<211> LENGTH: 6267
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

QIA-00689244

JX-001.32

Appx109

-continued

<400> SEQUENCE: 1

agctgcaagt ggcgggcgcc caggcagatg cgatccagcg gctctggggg cggcagcggt 60

ggtagcagct ggtacctccc gccgcctctg ttcggagggt cgcggggcac cgaggtgctt 120

tccggccgcc ctctggtcgg ccacccaaag ccgcgggcgc tgatgatggg tgaggagggg 180

gcggcaagat ttcgggcgcc cctgccctga acgccctcag ctgctgccgc cggggccgct 240

ccagtgcctg cgaactctga ggagccgagg cgccggtgag agcaaggacg ctgcaaactt 300

gcgcagcgcg ggggctggga ttcacgccca gaagttcagc aggcagacag tccgaagcct 360

tcccgcagcg gagagatagc ttgagggtgc gcaagacggc agcctccgcc ctcggttccc 420

gcccagaccg ggcagaagag cttggaggag ccaaaaggaa cgcaaaaggc ggccaggaca 480

gcgtgcagca gctgggagcc gccgttctca gccttaaaag ttgcagagat tggaggctgc 540

cccgagaggg gacagacccc agctccgact gcgggggcga ggagaggacg gtacccaact 600

gccacctccc ttcaaccata gtagttcctc tgtaccgagc gcagcgagct acagacgggg 660

gcgcggcact cggcgcggag agcgggaggc tcaaggtccc agccagtgag cccagtgtgc 720

ttgagtgtct ctggactcgc ccctgagctt ccaggtctgt ttcatttaga ctcctgctcg 780

cctccgtgca gttgggggaa agcaagagac ttgcgcgcac gcacagtcct ctggagatca 840

ggtggaagga gccgctgggt accaaggact gttcagagcc tcttcccatc tcggggagag 900

cgaagggtga ggctgggccc ggagagcagt gtaaacggcc tcctccggcg ggatgggagc 960

catcgggctc ctgtggctcc tgccgctgct gctttccacg gcagctgtgg gctccgggat 1020

ggggaccggc cagcgcgcgg gctccccagc tgcggggccg ccgctgcagc cccgggagcc 1080

actcagctac tcgcgcctgc agaggaagag tctggcagtt gacttcgtgg tgccctcgct 1140

cttccgtgtc tacgcccggg acctactgct gccaccatcc tcctcggagc tgaaggctgg 1200

caggcccgag gcccgcggct cgctagctct ggactgcgcc ccgctgctca ggttgctggg 1260

gccggcgccg ggggtctcct ggaccgccgg ttcaccagcc ccggcagagg cccggacgct 1320

gtccagggtg ctgaagggcg gctccgtgcg caagctccgg cgtgccaagc agttggtgct 1380

ggagctgggc gaggaggcga tcttggaggg ttgcgtcggg ccccccgggg aggcggctgt 1440

ggggctgctc cagttcaatc tcagcgagct gttcagttgg tggattcgcc aaggcgaagg 1500

gcgactgagg atccgcctga tgcccgagaa gaaggcgtcg gaagtgggca gagagggaag 1560

gctgtccgcg gcaattcgcg cctcccagcc ccgccttctc ttccagatct tcgggactgg 1620

tcatagctcc ttggaatcac caacaaacat gccttctcct tctcctgatt attttacatg 1680

gaatctcacc tggataatga aagactcctt ccctttcctg tctcatcgca gccgatatgg 1740

tctggagtgc agctttgact tcccctgtga gctggagtat tcccctccac tgcatgacct 1800

caggaaccag agctggtcct ggcgccgcat ccgctccgag gaggcctccc agatggactt 1860

gctggatggg cctggggcag agcgttctaa ggagatgccc agaggctcct ttctccttct 1920

caacacctca gctgactcca agcacaccat cctgagtccg tggatgagga gcagcagtga 1980

gcactgcaca ctggccgtct cggtgcacag gcacctgcag ccctctggaa ggtacattgc 2040

ccagctgctg ccccacaacg aggctgcaag agagatcctc ctgatgccca ctccagggaa 2100

gcatggttgg acagtgctcc agggaagaat cgggcgtcca gacaacccat ttcgagtggc 2160

cctggaatac atctccagtg gaaaccgcag cttgtctgca gtggacttct ttgccctgaa 2220

gaactgcagt gaaggaacat ccccaggctc caagatggcc ctgcagagct ccttcacttg 2280

ttggaatggg acagtcctcc agcttgggca ggcctgtgac ttccaccagg actgtgccca 2340

-continued

gggagaagat gagagccaga tgtgccggaa actgcctgtg ggtttttact gcaactttga 2400

agatggcttc tgtggctgga cccaaggcac actgtcaccc cacactcctc aatggcaggt 2460

caggacccta aaggatgccc ggttccagga ccaccaagac catgctctat tgctcagtac 2520

cactgatgtc cccgcttctg aaagtgctac agtgaccagt gctacgtttc ctgcaccgat 2580

caagagctct ccatgtgagc tccgaatgtc ctggctcatt cgtggagtct tgaggggaaa 2640

cgtgtccttg gtgctagtgg agaacaaaac cgggaaggag caaggcagga tggtctggca 2700

tgtcgccgcc tatgaaggct tgagcctgtg gcagtggatg gtgttgcctc tcctcgatgt 2760

gtctgacagg ttctggctgc agatggtcgc atggtgggga caaggatcca gagccatcgt 2820

ggcttttgac aatatctcca tcagcctgga ctgctacctc accattagcg gagaggacaa 2880

gatcctgcag aatacagcac ccaaatcaag aaacctgttt gagagaaacc caaacaagga 2940

gctgaaaccc ggggaaaatt caccaagaca gacccccatc tttgaccctа cagttcattg 3000

gctgttcacc acatgtgggg ccagcgggcc ccatggcccc acccaggcac agtgcaacaa 3060

cgcctaccag aactccaacc tgagcgtgga ggtggggagc gagggccccc tgaaaggcat 3120

ccagatctgg aaggtgccag ccaccgacac ctacagcatc tcgggctacg gagctgctgg 3180

cgggaaaggc gggaagaaca ccatgatgcg gtcccacggc gtgtctgtgc tgggcatctt 3240

caacctggag aaggatgaca tgctgtacat cctggttggg cagcagggag aggacgcctg 3300

ccccagtaca aaccagttaa tccagaaagt ctgcattgga gagaacaatg tgatagaaga 3360

agaaatccgt gtgaacagaa gcgtgcatga gtgggcagga ggcggaggag gaggggtgg 3420

agccacctac gtatttaaga tgaaggatgg agtgccggtg cccctgatca ttgcagccgg 3480

aggtggtggc agggcctacg gggccaagac agacacgttc cacccagaga gactggagaa 3540

taactcctcg gttctagggc taaacggcaa ttccggagcc gcaggtggtg gaggtggctg 3600

gaatgataac acttccttgc tctgggccgg aaaatctttg caggagggtg ccaccggagg 3660

acattcctgc ccccaggcca tgaagaagtg ggggtgggag acaagagggg gtttcggagg 3720

gggtggaggg gggtgctcct caggtggagg aggcggagga tatataggcg gcaatgcagc 3780

ctcaaacaat gaccccgaaa tggatgggga agatggggtt tccttcatca gtccactggg 3840

catcctgtac accccagctt taaaagtgat ggaaggccac ggggaagtga atattaagca 3900

ttatctaaac tgcagtcact gtgaggtaga cgaatgtcac atggaccctg aaagccacaa 3960

ggtcatctgc ttctgtgacc acgggacggt gctggctgag gatggcgtct cctgcattgt 4020

gtcacccacc ccggagccac acctgccact ctcgctgatc ctctctgtgg tgacctctgc 4080

cctcgtggcc gccctggtcc tggctttctc cggcatcatg attgtgtacc gccggaagca 4140

ccaggagctg caagccatgc agatggagct gcagagccct gagtacaagc tgagcaagct 4200

ccgcacctcg accatcatga ccgactacaa ccccaactac tgctttgctg gcaagacctc 4260

ctccatcagt gacctgaagg aggtgccgcg gaaaaacatc accctcattc ggggtctggg 4320

ccatggcgcc tttggggagg tgtatgaagg ccaggtgtcc ggaatgccca acgacccaag 4380

ccccctgcaa gtggctgtga agacgctgcc tgaagtgtgc tctgaacagg acgaactgga 4440

tttcctcatg gaagccctga tcatcagcaa attcaaccac cagaacattg ttcgctgcat 4500

tggggtgagc ctgcaatccc tgccccggtt catcctgctg gagctcatgg cgggggaga 4560

cctcaagtcc ttcctccgag agacccgccc tcgcccgagc cagccctcct ccctggccat 4620

gctggacctt ctgcacgtgg ctcgggacat tgcctgtggc tgtcagtatt tggaggaaaa 4680

ccacttcatc caccgagaca ttgctgccag aaactgcctc ttgacctgtc caggccctgg 4740

-continued

aagagtggcc aagattggag acttcgggat ggcccgagac atctacaggg cgagctacta 4800

tagaaaggga ggctgtgcca tgctgccagt taagtggatg cccccagagg ccttcatgga 4860

aggaatattc acttctaaaa cagacacatg gtcctttgga gtgctgctat gggaaatctt 4920

ttctcttgga tatatgccat accccagcaa aagcaaccag gaagttctgg agtttgtcac 4980

cagtggaggc cggatggacc cacccaagaa ctgccctggg cctgtatacc ggataatgac 5040

tcagtgctgg caacatcagc ctgaagacag gcccaacttt gccatcattt tggagaggat 5100

tgaatactgc acccaggacc cggatgtaat caacaccgct ttgccgatag aatatggtcc 5160

acttgtggaa gaggaagaga aagtgcctgt gaggcccaag gaccctgagg gggttcctcc 5220

tctcctggtc tctcaacagg caaaacggga ggaggagcgc agcccagctg ccccaccacc 5280

tctgcctacc acctcctctg gcaaggctgc aaagaaaccc acagctgcag agatctctgt 5340

tcgagtccct agagggccgg ccgtggaagg gggacacgtg aatatggcat tctctcagtc 5400

caaccctcct tcggagttgc acaaggtcca cggatccaga aacaagccca ccagcttgtg 5460

gaacccaacg tacggctcct ggtttacaga gaaacccacc aaaaagaata atcctatagc 5520

aaagaaggag ccacacgaca ggggtaacct ggggctggag ggaagctgta ctgtcccacc 5580

taacgttgca actgggagac ttccgggggc ctcactgctc ctagagccct cttcgctgac 5640

tgccaatatg aaggaggtac ctctgttcag gctacgtcac ttcccttgtg ggaatgtcaa 5700

ttacggctac cagcaacagg gcttgccctt agaagccgct actgcccctg gagctggtca 5760

ttacgaggat accattctga aaagcaagaa tagcatgaac cagcctgggc cctgagctcg 5820

gtcgcacact cacttctctt ccttgggatc cctaagaccg tggaggagag agaggcaatg 5880

gctccttcac aaaccagaga ccaaatgtca cgttttgttt tgtgccaacc tattttgaag 5940

taccaccaaa aaagctgtat tttgaaaatg ctttagaaag gttttgagca tgggttcatc 6000

ctattctttc gaaagaagaa aatatcataa aaatgagtga taaatacaag gcccagatgt 6060

ggttgcataa ggtttttatg catgtttgtt gtatacttcc ttatgcttct ttcaaattgt 6120

gtgtgctctg cttcaatgta gtcagaatta gctgcttcta tgtttcatag ttggggtcat 6180

agatgtttcc ttgccttgtt gatgtggaca tgagccattt gaggggagag ggaacggaaa 6240

taaaggagtt atttgtaatg actaaaa 6267

<210> SEQ ID NO 2
<211> LENGTH: 7368
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2

caagctttca agcattcaaa ggtctaaatg aaaaaggcta agtattattt caaaaggcaa 60

gtatatccta atatagcaaa acaaacaaag caaaatccat cagctactcc tccaattgaa 120

gtgatgaagc ccaaataatt catatagcaa aatggagaaa attagaccgg ccatctaaaa 180

atctgccatt ggtgaagtga tgaagaacat ttactgtctt attccgaagc ttgtcaattt 240

tgcaactctt ggctgcctat ggatttctgt ggtgcagtgt acagttttaa atagctgcct 300

aaagtcgtgt gtaactaatc tgggccagca gcttgacctt ggcacaccac ataatctgag 360

tgaaccgtgt atccaaggat gtcacttttg gaactctgta gatcagaaaa actgtgcttt 420

aaagtgtcgg gagtcgtgtg aggttggctg tagcagcgcg gaaggtgcat atgaagagga 480

agtactggaa aatgcagacc taccaactgc tccctttgct tcttccattg gaagccacaa 540

tatgacatta cgatggaaat ctgcaaactt ctctggagta aaatacatca ttcagtggaa 600

-continued

```
atatgcacaa cttctgggaa gctggactta tactaagact gtgtccagac cgtcctatgt   660
ggtcaagccc ctgcacccct tcactgagta cattttccga gtggtttgga tcttcacagc   720
gcagctgcag ctctactccc ctccaagtcc cagttacagg actcatcctc atggagttcc   780
tgaaactgca cctttgatta ggaatattga gagctcaagt cccgacactg tggaagtcag   840
ctgggatcca cctcaattcc caggtggacc tattttgggt tataacttaa ggctgatcag   900
caaaaatcaa aaattagatg cagggacaca gagaaccagt ttccagtttt actccacttt   960
accaaatact atctacaggt tttctattgc agcagtaaat gaagttggtg agggtccaga  1020
agcagaatct agtattacca cttcatcttc agcagttcaa caagaggaac agtggctctt  1080
tttatccaga aaaacttctc taagaaagag atctttaaaa catttagtag atgaagcaca  1140
ttgccttcgg ttggatgcta tataccataa tattacagga atatctgttg atgtccacca  1200
gcaaattgtt tatttctctg aaggaactct catatgggcg aagaaggctg ccaacatgtc  1260
tgatgtatct gacctgagaa ttttttacag aggttcagga ttaatttctt ctatctccat  1320
agattggctt tatcaaagaa tgtatttcat catggatgaa ctggtatgtg tctgtgattt  1380
agagaactgc tcaaacatcg aggaaattac tccaccctct attagtgcac ctcaaaaaat  1440
tgtggctgat tcatacaatg ggtatgtctt ttacctcctg agagatggca tttatagagc  1500
agaccttcct gtaccatctg gccggtgtgc agaagctgtg cgtattgtgg agagttgcac  1560
gttaaaggac tttgcaatca agccacaagc caagcgaatc atttacttca atgacactgc  1620
ccaagtcttc atgtcaacat ttctggatgg ctctgcttcc catctcatcc tacctcgcat  1680
ccccttgct gatgtgaaaa gttttgcttg tgaaaacaat gactttcttg tcacagatgg  1740
caaggtcatt ttccaacagg atgctttgtc ttttaatgaa ttcatcgtgg gatgtgacct  1800
gagtcacata gaagaatttg ggtttggtaa cttggtcatc tttggctcat cctcccagct  1860
gcaccctctg ccaggccgcc cgcaggagct ttcggtgctg tttggctctc accaggctct  1920
tgttcaatgg aagcctcctg cccttgccat aggagccaat gtcatcctga tcagtgatat  1980
tattgaactc tttgaattag gcccttctgc ctggcagaac tggacctatg aggtgaaagt  2040
atccacccaa gaccctcctg aagtcactca tattttcttg aacataagtg gaaccatgct  2100
gaatgtacct gagctgcaga gtgctatgaa atacaaggtt tctgtgagag caagttctcc  2160
aaagaggcca ggcccctggt cagagccctc agtgggtact accctggtgc cagctagtga  2220
accaccattt atcatggctg tgaaagaaga tgggctttgg agtaaaccat taaatagctt  2280
tggcccagga gagttcttat cctctgatat aggaaatgtg tcagacatgg attggtataa  2340
caacagcctc tactacagtg acacgaaagg cgacgttttt gtgtggctgc tgaatgggac  2400
ggatatctca gagaattatc acctacccag cattgcagga gcaggggctt tagcttttga  2460
gtggctgggt cactttctct actgggctgg aaagacatat gtgatacaaa ggcagtctgt  2520
gttgacggga cacacagaca ttgttaccca cgtgaagcta ttggtgaatg acatggtggt  2580
ggattcagtt ggtggatatc tctactggac cacactctat tcagtggaaa gcaccagact  2640
aaatggggaa agttcccttg tactacagac acagccttgg ttttctggga aaaaggtaat  2700
tgctctaact ttagacctca gtgatgggct cctgtattgg ttggttcaag acagtcaatg  2760
tattcacctg tacacagctg ttcttcgggg acagagcact ggggatacca ccatcacaga  2820
atttgcagcc tggagtactt ctgaaatttc ccagaatgca ctgatgtact atagtggtcg  2880
gctgttctgg atcaatggct ttaggattat cacaactcaa gaaataggtc agaaaaccag  2940
tgtctctgtt ttggaaccag ccagatttaa tcagttcaca attattcaga catcccttaa  3000
```

-continued

gcccctgcca gggaactttt cctttacccc taaggttatt ccagattctg ttcaagagtc 3060

ttcatttagg attgaaggaa atgcttcaag ttttcaaatc ctgtggaatg gtcccctgc 3120

ggtagactgg ggtgtagttt tctacagtgt agaatttagt gctcattcta agttcttggc 3180

tagtgaacaa cactctttac ctgtatttac tgtggaagga ctggaacctt atgccttatt 3240

taatctttct gtcactcctt atacctactg gggaaagggc cccaaaacat ctctgtcact 3300

tcgagcacct gaaacagttc catcagcacc agagaacccc agaatattta tattaccaag 3360

tggaaaatgc tgcaacaaga atgaagttgt ggtggaattt aggtggaaca aacctaagca 3420

tgaaaatggg gtgttaacaa aatttgaaat tttctacaat atatccaatc aaagtattac 3480

aaacaaaaca tgtgaagact ggattgctgt caatgtcact ccctcagtga tgtcttttca 3540

acttgaaggc atgagtccca gatgctttat tgccttccag gttagggcct ttacatctaa 3600

ggggccagga ccatatgctg acgttgtaaa gtctacaaca tcagaaatca acccatttcc 3660

tcacctcata actcttcttg gtaacaagat agtttttta gatatggatc aaaatcaagt 3720

tgtgtggacg ttttcagcag aaagagttat cagtgccgtt tgctacacag ctgataatga 3780

gatgggatat tatgctgaag gggactcact ctttcttctg cacttgcaca atcgctctag 3840

ctctgagctt ttccaagatt cactggtttt tgatatcaca gttattacaa ttgactggat 3900

ttcaaggcac ctctactttg cactgaaaga atcacaaaat ggaatgcaag tatttgatgt 3960

tgatcttgaa cacaaggtga aatatcccag agaggtgaag attcacaata ggaattcaac 4020

aataatttct ttttctgtat atcctctttt aagtcgcttg tattggacag aagtttccaa 4080

ttttggctac cagatgttct actacagtat tatcagtcac accttgcacc gaattctgca 4140

acccacagct acaaaccaac aaaacaaaag gaatcaatgt tcttgtaatg tgactgaatt 4200

tgagttaagt ggagcaatgg ctattgatac ctctaaccta gagaaaccat tgatatactt 4260

tgccaaagca caagagatct gggcaatgga tctggaaggc tgtcagtgtt ggagagttat 4320

cacagtacct gctatgctcg caggaaaaac ccttgttagc ttaactgtgg atggagatct 4380

tatatactgg atcatcacag caaaggacag cacacagatt tatcaggcaa agaaaggaaa 4440

tggggccatc gtttcccagg tgaaggccct aaggagtagg catatcttgg cttacagttc 4500

agttatgcag ccttttccag ataaagcgtt tctgtctcta gcttcagaca ctgtggaacc 4560

aactatactt aatgccacta acactagcct cacaatcaga ttacctctgg ccaagacaaa 4620

cctcacatgg tatggcatca ccagccctac tccaacatac ctggtttatt atgcagaagt 4680

taatgacagg aaaaacagct ctgacttgaa atatagaatt ctggaatttc aggacagtat 4740

agctcttatt gaagatttac aaccattttc aacatacatg atacagatag ctgtaaaaaa 4800

ttattattca gatcctttgg aacatttacc accaggaaaa gagatttggg gaaaaactaa 4860

aaatggagta ccagaggcag tgcagctcat taatacaact gtgcggtcag acaccagcct 4920

cattatatct tggagagaat ctcacaagcc aaatggacct aaagaatcag tccgttatca 4980

gttggcaatc tcacacctgg ccctaattcc tgaaactcct ctaagacaaa gtgaatttcc 5040

aaatggaagg ctcactctcc ttgttactag actgtctggt ggaaatattt atgtgttaaa 5100

ggttcttgcc tgccactctg aggaaatgtg gtgtacagag agtcatcctg tcactgtgga 5160

aatgtttaac acaccagaga aaccttattc cttggttcca gagaacacta gtttgcaatt 5220

taattggaag gctccattga atgttaacct catcagattt tgggttgagc tacagaagtg 5280

gaaatacaat gagttttacc atgttaaaac ttcatgcagc caaggtcctg cttatgtctg 5340

taatatcaca aatctacaac cttatacttc atataatgtc agagtagtgg tggtttataa 5400

-continued

gacgggagaa aatagcacct cacttccaga aagctttaag acaaaagctg gagtcccaaa 5460

taaaccaggc attcccaaat tactagaagg gagtaaaaat tcaatacagt gggagaaagc 5520

tgaagataat ggatgtagaa ttacatacta tatccttgag ataagaaaga gcacttcaaa 5580

taatttacag aaccagaatt taaggtggaa gatgacattt aatggatcct gcagtagtgt 5640

ttgcacatgg aagtccaaaa acctgaaagg aatatttcag ttcagagtag tagctgcaaa 5700

taatctaggg tttggtgaat atagtggaat cagtgagaat attatattag ttggagatga 5760

tttttggata ccagaaacaa gtttcatact tactattata gttggaatat ttctggttgt 5820

tacaatccca ctgacctttg tctggcatag aagattaaag aatcaaaaaa gtgccaagga 5880

aggggtgaca gtgcttataa acgaagacaa agagttggct gagctgcgag gtctggcagc 5940

cggagtaggc ctggctaatg cctgctatgc aatacatact cttccaaccc aagaggagat 6000

tgaaaatctt cctgccttcc ctcgggaaaa actgactctg cgtctcttgc tgggaagtgg 6060

agcctttgga gaagtgtatg aaggaacagc agtggacatc ttaggagttg gaagtggaga 6120

aatcaaagta gcagtgaaga ctttgaagaa gggttccaca gaccaggaga agattgaatt 6180

cctgaaggag gcacatctga tgagcaaatt taatcatccc aacattctga agcagcttgg 6240

agtttgtctg ctgaatgaac cccaatacat tatcctggaa ctgatggagg gaggagacct 6300

tcttacttat ttgcgtaaag cccggatggc aacgttttat ggtcctttac tcaccttggt 6360

tgaccttgta gacctgtgtg tagatatttc aaaaggctgt gtctacttgg aacggatgca 6420

tttcattcac agggatctgg cagctagaaa ttgccttgtt tccgtgaaag actataccag 6480

tccacggata gtgaagattg gagactttgg actcgccaga gacatctata aaaatgatta 6540

ctatagaaag agaggggaag gcctgctccc agttcggtgg atggctccag aaagtttgat 6600

ggatggaatc ttcactactc aatctgatgt atggtctttt ggaattctga tttgggagat 6660

tttaactctt ggtcatcagc cttatccagc tcattccaac cttgatgtgt taaactatgt 6720

gcaaacagga gggagactgg agccaccaag aaattgtcct gatgatctgt ggaatttaat 6780

gacccagtgc tgggctcaag aacccgacca aagacctact tttcatagaa ttcaggacca 6840

acttcagtta ttcagaaatt ttttcttaaa tagcatttat aagtccagag atgaagcaaa 6900

caacagtgga gtcataaatg aaagctttga aggtgaagat ggcgatgtga tttgtttgaa 6960

ttcagatgac attatgccag ttgctttaat ggaaacgaag aaccgagaag ggttaaacta 7020

tatggtactt gctacagaat gtggccaagg tgaagaaaag tctgagggtc ctctaggctc 7080

ccaggaatct gaatcttgtg gtctgaggaa agaagagaag gaaccacatg cagacaaaga 7140

tttctgccaa gaaaaacaag tggcttactg cccttctggc aagcctgaag gcctgaacta 7200

tgcctgtctc actcacagtg gatatggaga tgggtctgat taatagcgtt gtttgggaaa 7260

tagagagttg agataaacac tctcattcag tagttactga aagaaaactc tgctagaatg 7320

ataaatgtca tggtggtcta taactccaaa taaacaatgc aacgttcc 7368

<210> SEQ ID NO 3
<211> LENGTH: 4174
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 3

agtcccgcga ccgaagcagg gcgcgcagca gcgctgagtg ccccggaacg tgcgtcgcgc 60

ccccagtgtc cgtcgcgtcc gccgcgcccc gggcggggat ggggcggcca gactgagcgc 120

cgcacccgcc atccagaccc gccggcccta gccgcagtcc ctccagccgt ggccccagcg 180

-continued

```
cgcacgggcg atggcgaagg cgacgtccgg tgccgcgggg ctgcgtctgc tgttgctgct   240
gctgctgccg ctgctaggca aagtggcatt gggcctctac ttctcgaggg atgcttactg   300
ggagaagctg tatgtggacc aggcggccgg cacgcccttg ctgtacgtcc atgccctgcg   360
ggacgcccct gaggaggtgc ccagcttccg cctgggccag catctctacg gcacgtaccg   420
cacacggctg catgagaaca actggatctg catccaggag gacaccggcc tcctctacct   480
taaccggagc ctggaccata gctcctggga gaagctcagt gtccgcaacc gcggctttcc   540
cctgctcacc gtctacctca aggtcttcct gtcacccaca tcccttcgtg agggcgagtg   600
ccagtggcca ggctgtgccc gcgtatactt ctccttcttc aacacctcct ttccagcctg   660
cagctccctc aagccccggg agctctgctt cccagagaca aggccctcct tccgcattcg   720
ggagaaccga cccccaggca ccttccacca gttccgcctg ctgcctgtgc agttcttgtg   780
ccccaacatc agcgtggcct acaggctcct ggagggtgag ggtctgccct tccgctgcgc   840
cccggacagc ctggaggtga gcacgcgctg ggccctggac cgcgagcagc gggagaagta   900
cgagctggtg gccgtgtgca ccgtgcacgc cggcgcgcgc gaggaggtgg tgatggtgcc   960
cttcccggtg accgtgtacg acgaggacga ctcggcgccc accttccccg cgggcgtcga  1020
caccgccagc gccgtggtgg agttcaagcg gaaggaggac accgtggtgg ccacgctgcg  1080
tgtcttcgat gcagacgtgg tacctgcatc aggggagctg gtgaggcggt acacaagcac  1140
gctgctcccc ggggacacct gggcccagca gaccttccgg gtggaacact ggcccaacga  1200
gacctcggtc caggccaacg gcagcttcgt gcgggcgacc gtacatgact ataggctggt  1260
tctcaaccgg aacctctcca tctcggagaa ccgcaccatg cagctggcgg tgctggtcaa  1320
tgactcagac ttccagggcc caggagcggg cgtcctcttg ctccacttca acgtgtcggt  1380
gctgccggtc agcctgcacc tgcccagtac ctactccctc tccgtgagca ggagggctcg  1440
ccgatttgcc cagatcggga aagtctgtgt ggaaaactgc caggcattca gtggcatcaa  1500
cgtccagtac aagctgcatt cctctggtgc caactgcagc acgctagggg tggtcacctc  1560
agccgaggac acctcgggga tcctgtttgt gaatgacacc aaggccctgc ggcggcccaa  1620
gtgtgccgaa cttcactaca tggtggtggc caccgaccag cagacctcta ggcaggccca  1680
ggcccagctg cttgtaacag tggaggggtc atatgtggcc gaggaggcgg gctgcccccc  1740
gtcctgtgca gtcagcaaga gacggctgga gtgtgaggag tgtggcggcc tgggctcccc  1800
aacaggcagg tgtgagtgga ggcaaggaga tggcaaaggg atcaccagga acttctccac  1860
ctgctctccc agcaccaaga cctgccccga cggccactgc gatgttgtgg agacccaaga  1920
catcaacatt tgccctcagg actgcctccg gggcagcatt gttgggggac acgagcctgg  1980
ggagccccgg gggattaaag ctggctatgg cacctgcaac tgcttccctg aggaggagaa  2040
gtgcttctgc gagcccgaag acatccagga tccactgtgc gacgagctgt gccgcacggt  2100
gatcgcagcc gctgtcctct tctccttcat cgtctcggtg ctgctgtctg ccttctgcat  2160
ccactgctac cacaagtttg cccacaagcc acccatctcc tcagctgaga tgaccttccg  2220
gaggcccgcc caggccttcc cggtcagcta ctcctcttcc ggtgcccgcc ggccctcgct  2280
ggactccatg gagaaccagg tctccgtgga tgccttcaag atcctggagg atccaaagtg  2340
ggaattccct cggaagaact tggttcttgg aaaaactcta ggagaaggcg aatttggaaa  2400
agtggtcaag gcaacggcct tccatctgaa aggcagagca gggtacacca cggtggccgt  2460
gaagatgctg aaagagaacg cctccccgag tgagcttcga gacctgctgt cagagttcaa  2520
cgtcctgaag caggtcaacc acccacatgt catcaaattg tatggggcct gcagccagga  2580
```

-continued

tggcccgctc ctcctcatcg tggagtacgc caaatacggc tccctgcggg gcttcctccg 2640

cgagagccgc aaagtggggc ctggctacct gggcagtgga ggcagccgca actccagctc 2700

cctggaccac ccggatgagc gggccctcac catgggcgac ctcatctcat ttgcctggca 2760

gatctcacag gggatgcagt atctggccga gatgaagctc gttcatcggg acttggcagc 2820

cagaaacatc ctggtagctg aggggcggaa gatgaagatt tcggatttcg gcttgtcccg 2880

agatgtttat gaagaggatt cctacgtgaa gaggagccag ggtcggattc cagttaaatg 2940

gatggcaatt gaatcccttt ttgatcatat ctacaccacg caaagtgatg tatggtcttt 3000

tggtgtcctg ctgtgggaga tcgtgaccct agggggaaac ccctatcctg ggattcctcc 3060

tgagcggctc ttcaaccttc tgaagaccgg ccaccggatg gagaggccag acaactgcag 3120

cgaggagatg taccgcctga tgctgcaatg ctggaagcag gagccggaca aaaggccggt 3180

gtttgcggac atcagcaaag acctggagaa gatgatggtt aagaggagag actacttgga 3240

ccttgcggcg tccactccat ctgactccct gatttatgac gacggcctct cagaggagga 3300

gacaccgctg gtggactgta ataatgcccc cctccctcga gccctccctt ccacatggat 3360

tgaaaacaaa ctctatggta gaatttccca tgcatttact agattctagc accgctgtcc 3420

cctctgcact atccttcctc tctgtgatgc ttttttaaaaa tgtttctggt ctgaacaaaa 3480

ccaaagtctg ctctgaacct ttttatttgt aaatgtctga ctttgcatcc agtttacatt 3540

taggcattat tgcaactatg tttttctaaa aggaagtgaa aataagtgta attaccacat 3600

tgcccagcaa cttaggatgg tagaggaaaa aacagatcag ggcggaactc tcaggggaga 3660

ccaagaacag gttgaataag gcgcttctgg ggtgggaatc aagtcatagt acttctactt 3720

taactaagtg gataaatata caaatctggg gaggtattca gttgagaaag gagccaccag 3780

caccactcag cctgcactgg gagcacagcc aggttccccc agacccctcc tgggcaggca 3840

ggtgcctctc agaggccacc cggcactggc gagcagccac tggccaagcc tcagccccag 3900

tcccagccac atgtcctcca tcaggggtag cgaggttgca ggagctggct ggccctggga 3960

ggacgcaccc ccactgctgt tttcacatcc tttcccttac ccaccttcag gacggttgtc 4020

acttatgaag tcagtgctaa agctggagca gttgcttttt gaaagaacat ggtctgtggt 4080

gctgtggtct tacaatggac agtaaatatg gttcttgcca aaactccttc ttttgtcttt 4140

gattaaatac tagaaattta aaaaaaaaaa aaaa 4174

<210> SEQ ID NO 4
<211> LENGTH: 5629
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 4

agtcccgcga ccgaagcagg gcgcgcagca gcgctgagtg ccccggaacg tgcgtcgcgc 60

ccccagtgtc cgtcgcgtcc gccgcgcccc gggcggggat ggggcggcca gactgagcgc 120

cgcacccgcc atccagaccc gccggcccta gccgcagtcc ctccagccgt ggccccagcg 180

cgcacgggcg atggcgaagg cgacgtccgg tgccgcgggg ctgcgtctgc tgttgctgct 240

gctgctgccg ctgctaggca aagtggcatt gggcctctac ttctcgaggg atgcttactg 300

ggagaagctg tatgtggacc aggcggccgg cacgcccttg ctgtacgtcc atgccctgcg 360

ggacgcccct gaggaggtgc ccagcttccg cctgggccag catctctacg gcacgtaccg 420

cacacggctg catgagaaca actggatctg catccaggag gacaccggcc tcctctacct 480

taaccggagc ctggaccata gctcctggga gaagctcagt gtccgcaacc gcggctttcc 540

-continued

```
cctgctcacc gtctacctca aggtcttcct gtcacccaca tcccttcgtg agggcgagtg   600
ccagtggcca ggctgtgccc gcgtatactt ctccttcttc aacacctcct ttccagcctg   660
cagctccctc aagccccggg agctctgctt cccagagaca aggccctcct tccgcattcg   720
ggagaaccga cccccaggca ccttccacca gttccgcctg ctgcctgtgc agttcttgtg   780
ccccaacatc agcgtggcct acaggctcct ggagggtgag ggtctgccct tccgctgcgc   840
cccggacagc ctggaggtga gcacgcgctg ggccctggac cgcgagcagc gggagaagta   900
cgagctggtg gccgtgtgca ccgtgcacgc cggcgcgcgc gaggaggtgg tgatggtgcc   960
cttcccggtg accgtgtacg acgaggacga ctcggcgccc accttccccg cgggcgtcga  1020
caccgccagc gccgtggtgg agttcaagcg gaaggaggac accgtggtgg ccacgctgcg  1080
tgtcttcgat gcagacgtgg tacctgcatc aggggagctg gtgaggcggt acacaagcac  1140
gctgctcccc ggggacacct gggcccagca gaccttccgg gtggaacact ggcccaacga  1200
gacctcggtc caggccaacg gcagcttcgt gcgggcgacc gtacatgact ataggctggt  1260
tctcaaccgg aacctctcca tctcggagaa ccgcaccatg cagctggcgg tgctggtcaa  1320
tgactcagac ttccagggcc caggagcggg cgtcctcttg ctccacttca acgtgtcggt  1380
gctgccggtc agcctgcacc tgcccagtac ctactccctc tccgtgagca ggagggctcg  1440
ccgatttgcc cagatcggga aagtctgtgt ggaaaactgc caggcattca gtggcatcaa  1500
cgtccagtac aagctgcatt cctctggtgc caactgcagc acgctagggg tggtcacctc  1560
agccgaggac acctcgggga tcctgtttgt gaatgacacc aaggccctgc ggcggcccaa  1620
gtgtgccgaa cttcactaca tggtggtggc caccgaccag cagacctcta ggcaggccca  1680
ggcccagctg cttgtaacag tggaggggtc atatgtggcc gaggaggcgg gctgcccctt  1740
gtcctgtgca gtcagcaaga gacggctgga gtgtgaggag tgtggcggcc tgggctcccc  1800
aacaggcagg tgtgagtgga ggcaaggaga tggcaaaggg atcaccagga acttctccac  1860
ctgctctccc agcaccaaga cctgccccga cggccactgc gatgttgtgg agacccaaga  1920
catcaacatt tgccctcagg actgcctccg gggcagcatt gttgggggac acgagcctgg  1980
ggagccccgg gggattaaag ctggctatgg cacctgcaac tgcttccctg aggaggagaa  2040
gtgcttctgc gagcccgaag acatccagga tccactgtgc gacgagctgt gccgcacggt  2100
gatcgcagcc gctgtcctct tctccttcat cgtctcggtg ctgctgtctg ccttctgcat  2160
ccactgctac cacaagtttg cccacaagcc acccatctcc tcagctgaga tgaccttccg  2220
gaggcccgcc caggccttcc cggtcagcta ctcctcttcc ggtgcccgcc ggccctcgct  2280
ggactccatg gagaaccagg tctccgtgga tgccttcaag atcctggagg atccaaagtg  2340
ggaattccct cggaagaact tggttcttgg aaaaactcta ggagaaggcg aatttggaaa  2400
agtggtcaag gcaacggcct tccatctgaa aggcagagca gggtacacca cggtggccgt  2460
gaagatgctg aaagagaacg cctccccgag tgagcttcga gacctgctgt cagagttcaa  2520
cgtcctgaag caggtcaacc acccacatgt catcaaattg tatggggcct gcagccagga  2580
tggcccgctc ctcctcatcg tggagtacgc caaatacggc tccctgcggg gcttcctccg  2640
cgagagccgc aaagtggggc ctggctacct gggcagtgga ggcagccgca actccagctc  2700
cctggaccac ccggatgagc gggccctcac catgggcgac ctcatctcat ttgcctggca  2760
gatctcacag gggatgcagt atctggccga gatgaagctc gttcatcggg acttggcagc  2820
cagaaacatc ctggtagctg aggggcggaa gatgaagatt tcggatttcg gcttgtcccg  2880
agatgtttat gaagaggatt cctacgtgaa gaggagccag ggtcggattc cagttaaatg  2940
```

-continued

gatggcaatt gaatcccttt ttgatcatat ctacaccacg caaagtgatg tatggtcttt 3000

tggtgtcctg ctgtgggaga tcgtgaccct agggggaaac ccctatcctg ggattcctcc 3060

tgagcggctc ttcaaccttc tgaagaccgg ccaccggatg gagaggccag acaactgcag 3120

cgaggagatg taccgcctga tgctgcaatg ctggaagcag gagccggaca aaaggccggt 3180

gtttgcggac atcagcaaag acctggagaa gatgatggtt aagaggagag actacttgga 3240

ccttgcggcg tccactccat ctgactccct gatttatgac gacggcctct cagaggagga 3300

gacaccgctg gtggactgta ataatgcccc cctccctcga gccctccctt ccacatggat 3360

tgaaaacaaa ctctatggca tgtcagaccc gaactggcct ggagagagtc ctgtaccact 3420

cacgagagct gatggcacta acactgggtt tccaagatat ccaaatgata gtgtatatgc 3480

taactggatg ctttcaccct cagcggcaaa attaatggac acgtttgata gttaacattt 3540

ctttgtgaaa ggtaatggac tcacaagggg aagaaacatg ctgagaatgg aaagtctacc 3600

ggccctttct ttgtgaacgt cacattggcc gagccgtgtt cagttcccag gtggcagact 3660

cgtttttggt agtttgtttt aacttccaag gtggttttac ttctgatagc cggtgatttt 3720

ccctcctagc agacatgcca caccgggtaa gagctctgag tcttagtggt taagcattcc 3780

tttctcttca gtgcccagca gcacccagtg ttggtctgtg tccatcagtg accaccaaca 3840

ttctgtgttc acatgtgtgg gtccaacact tactacctgg tgtatgaaat tggacctgaa 3900

ctgttggatt tttctagttg ccgccaaaca aggcaaaaaa atttaaacat gaagcacaca 3960

cacaaaaaag gcagtaggaa aaatgctggc cctgatgacc tgtccttatt cagaatgaga 4020

gactgcgggg ggggcctggg ggtagtgtca atgcccctcc agggctggag gggaagaggg 4080

gccccgagga tgggcctggg ctcagcattc gagatcttga gaatgatttt tttttaatca 4140

tgcaaccttt ccttaggaag acatttggtt ttcatcatga ttaagatgat tcctagattt 4200

agcacaatgg agagattcca tgccatcttt actatgtgga tggtggtatc agggaagagg 4260

gctcacaaga cacatttgtc ccccgggccc accacatcat cctcacgtgt tcggtactga 4320

gcagccacta cccctgatga gaacagtatg aagaaagggg gctgttggag tcccagaatt 4380

gctgacagca gaggctttgc tgctgtgaat cccacctgcc accagcctgc agcacacccc 4440

acagccaagt agaggcgaaa gcagtggctc atcctacctg ttaggagcag gtagggcttg 4500

tactcacttt aatttgaatc ttatcaacct actcataaag ggacaggcta gctagctgtg 4560

ttagaagtag caatgacaat gaccaaggac tgctacacct ctgattacaa ttctgatgtg 4620

aaaaagatgg tgtttggctc ttatagagcc tgtgtgaaag gcccatggat cagctcttcc 4680

tgtgtttgta atttaatgct gctacaagat gtttctgttt cttagattct gaccatgact 4740

cataagcttc ttgtcattct tcattgcttg tttgtggtca cagatgcaca acactcctcc 4800

agtcttgtgg gggcagcttt tgggaagtct cagcagctct tctggctgtg ttgtcagcac 4860

tgtaacttcg cagaaaagag tcggattacc aaaacactgc ctgctcttca gacttaaagc 4920

actgatagga cttaaaatag tctcattcaa atactgtatt ttatataggc atttcacaaa 4980

aacagcaaaa ttgtggcatt ttgtgaggcc aaggcttgga tgcgtgtgta atagagcctt 5040

gtggtgtgtg cgcacacacc cagagggaga gtttgaaaaa tgcttattgg acacgtaacc 5100

tggctctaat ttgggctgtt tttcagatac actgtgataa gttcttttac aaatatctat 5160

agacatggta aacttttggt tttcagatat gcttaatgat agtcttacta aatgcagaaa 5220

taagaataaa ctttctcaaa ttattaaaaa tgcctacaca gtaagtgtga attgctgcaa 5280

caggtttgtt ctcaggaggg taagaactcc aggtctaaac agctgaccca gtgatgggga 5340

-continued

atttatcctt gaccaattta tccttgacca ataacctaat tgtctattcc tgagttataa 5400

aagtccccat ccttattagc tctactggaa ttttcataca cgtaaatgca gaagttacta 5460

agtattaagt attactgagt attaagtagt aatctgtcag ttattaaaat ttgtaaaatc 5520

tatttatgaa aggtcattaa accagatcat gttccttttt ttgtaatcaa ggtgactaag 5580

aaaatcagtt gtgtaaataa aatcatgtat cataaaaaaa aaaaaaaaa 5629

<210> SEQ ID NO 5
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 5

cctctctatg ggcagtcggt gatgcgagag tggcaggtgt gg 42

<210> SEQ ID NO 6
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 6

ggatctcgac gctctccctc agaggtcacc acagagagga tcag 44

<210> SEQ ID NO 7
<211> LENGTH: 43
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 7

cctctctatg ggcagtcggt gatcatggct tgcagctcct ggt 43

<210> SEQ ID NO 8
<211> LENGTH: 41
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 8

ggatctcgac gctctccctg cagctccatc tgcatggctt g 41

<210> SEQ ID NO 9
<211> LENGTH: 45
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 9

cctctctatg ggcagtcggt gatggccttc atacacctcc ccaaa 45

-continued

<210> SEQ ID NO 10
<211> LENGTH: 37
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 10

ggatctcgac gctctccctt tgggcattcc ggacacc 37

<210> SEQ ID NO 11
<211> LENGTH: 49
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 11

cctctctatg ggcagtcggt gataggaaat ccagttcgtc ctgttcaga 49

<210> SEQ ID NO 12
<211> LENGTH: 43
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 12

ggatctcgac gctctccctg atcagggctt ccatgaggaa atc 43

<210> SEQ ID NO 13
<211> LENGTH: 41
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 13

cctctctatg ggcagtcggt gatggctccc caggctcgtg t 41

<210> SEQ ID NO 14
<211> LENGTH: 43
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 14

ggatctcgac gctctcccta ggtgccatag ccagctttaa tcc 43

<210> SEQ ID NO 15
<211> LENGTH: 43
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 15

cctctctatg ggcagtcggt gatatcaccg tgcggcacag ctc 43

-continued

<210> SEQ ID NO 16
<211> LENGTH: 39
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 16

ggatctcgac gctctccctg aggacagcgg ctgcgatca 39

<210> SEQ ID NO 17
<211> LENGTH: 47
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 17

cctctctatg ggcagtcggt gatagaacca agttcttccg agggaat 47

<210> SEQ ID NO 18
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 18

ggatctcgac gctctccctt ccaaattcgc cttctcctag agtt 44

<210> SEQ ID NO 19
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 19

cctctctatg ggcagtcggt gatacagcag gtctcgcagc tcac 44

<210> SEQ ID NO 20
<211> LENGTH: 41
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 20

ggatctcgac gctctccctt gacctgcttc aggacgttga a 41

<210> SEQ ID NO 21
<211> LENGTH: 45
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 21

cctctctatg ggcagtcggt gatcttgctg actgcacagg acagg 45

-continued

<210> SEQ ID NO 22
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 22

ggatctcgac gctctccctt cctcacactc cagccgtctc 40

<210> SEQ ID NO 23
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 23

cctctctatg ggcagtcggt gattggtgct gggagagcag gt 42

<210> SEQ ID NO 24
<211> LENGTH: 37
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 24

ggatctcgac gctctccctc cgtcggggca ggtcttg 37

<210> SEQ ID NO 25
<211> LENGTH: 46
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 25

cctctctatg ggcagtcggt gatggctgca tgaagtttta acatgg 46

<210> SEQ ID NO 26
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 26

ggatctcgac gctctccctt gatattacag acataagcag gaccttgg 48

<210> SEQ ID NO 27
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 27

cctctctatg ggcagtcggt gatctagtaa tttgggaatg cctggttt 48

-continued

<210> SEQ ID NO 28
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 28

ggatctcgac gctctccctt tcagctttct cccactgtat tgaa 44

<210> SEQ ID NO 29
<211> LENGTH: 52
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 29

cctctctatg ggcagtcggt gatcatcttc caccttaaat tctggttctg ta 52

<210> SEQ ID NO 30
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 30

ggatctcgac gctctccctc aggatccatt aaatgtcatc ttcc 44

<210> SEQ ID NO 31
<211> LENGTH: 52
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 31

cctctctatg ggcagtcggt gatagtaagt atgaaacttg tttctggtat cc 52

<210> SEQ ID NO 32
<211> LENGTH: 43
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 32

ggatctcgac gctctccctg gtcagtggga ttgtaacaac cag 43

<210> SEQ ID NO 33
<211> LENGTH: 43
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 33

cctctctatg ggcagtcggt gatcacccct tccttggcac ttt 43

-continued

<210> SEQ ID NO 34
<211> LENGTH: 46
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 34

ggatctcgac gctctccctt ctttgtcttc gtttataagc actgtc 46

<210> SEQ ID NO 35
<211> LENGTH: 45
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 35

cctctctatg ggcagtcggt gatttcaatc tcctcttggg ttgga 45

<210> SEQ ID NO 36
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 36

ggatctcgac gctctccctc cgagggaagg caggaagatt 40

<210> SEQ ID NO 37
<211> LENGTH: 47
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 37

cctctctatg ggcagtcggt gatcaggaat tcaatcttct cctggtc 47

<210> SEQ ID NO 38
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 38

ggatctcgac gctctccctc tcatcagatg tgcctccttc ag 42

<210> SEQ ID NO 39
<211> LENGTH: 103
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 39

aggatgtccc tgtcaaggct cctgagacct ttgataacat aaccattagc agagaggctc 60

aggctggagt cccaaataaa ccaggcattc ccaaattact aga 103

-continued

<210> SEQ ID NO 40
<211> LENGTH: 88
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 40

tcaaggctct gagacctttg ataacataac cattagcaga gaggctcagg ctggagtcca 60

aatcaaacag gcattcccaa atactaga 88

<210> SEQ ID NO 41
<211> LENGTH: 81
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 41

gagacctttg ataacataac cattagcaga gaggctcagg ctggagtccc aaataaacca 60

ggcattccca aaattactag a 81

<210> SEQ ID NO 42
<211> LENGTH: 96
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 42

gtcctgtcaa ggctctgaga cctttgataa cataaccatt agcagagagg ctcaggctgg 60

agtcccaaat aaaccaggca ttcccaaatt actaga 96

<210> SEQ ID NO 43
<211> LENGTH: 103
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 43

gcaggatgtc cctgtcaagg ctctgagacc tttgataaca taaccattag cagagaggct 60

caggatgatt tttggatacc agaaacaagt ttcatactta cta 103

<210> SEQ ID NO 44
<211> LENGTH: 115
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 44

ggcgcaggag gggcaggatg tccctgtcaa ggctctgaga cctttgataa cataaccatt 60

agcagagagg ctcaggatga tttttggata ccagaaacaa gtttcatact tacta 115

<210> SEQ ID NO 45
<211> LENGTH: 46
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence

-continued

<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 45

ccatctcatc cctgcgtgtc tccgactcag ctaaggtaac tggcgt 46

<210> SEQ ID NO 46
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 46

cgccaggttc cagtcacgac 20

<210> SEQ ID NO 47
<211> LENGTH: 70
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 47

aatgatacgg cgaccaccga gatctacact agatcgcaca ctctttccct acacgacgct 60

cttccgatct 70

<210> SEQ ID NO 48
<211> LENGTH: 54
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 48

gatcggaaga gccagttacc ttagctgagt cggagacacg cagggatgag atgg 54

What is claimed herein is:

1. A method for preparing a nucleic acid for sequencing, the method comprising:

(i) ligating a universal oligonucleotide tail adaptor that comprises a first ligatable duplex end and a second unpaired end to a nucleic acid comprising a known target nucleotide sequence to produce a ligation product, the universal oligonucleotide tail adaptor comprising an amplification strand and a blocking strand, wherein a 3' duplex portion of the amplification strand and a 5' duplex portion of the blocking strand are substantially complementary and form the first ligatable duplex end;

(ii) amplifying the ligation product using a first target-specific primer that specifically anneals to the known target nucleotide sequence and a first adaptor primer having a nucleotide sequence identical to a first portion of the amplification strand; and

(iii) amplifying an amplification product of (ii) using a second target-specific primer that specifically anneals to the amplification product of (ii) and a second adaptor primer having a nucleotide sequence identical to a second portion of the amplification strand, wherein the second target-specific primer is nested relative to the first target-specific primer.

2. The method of claim **1**, wherein the entirety of the blocking strand is a duplex portion that is substantially complementary to the 3' duplex portion of the amplification strand.

3. The method of claim **1**, wherein the blocking strand further comprises a 3' unpaired portion which is not substantially complementary to the 5' unpaired portion of the amplification strand.

4. The method of claim **1**, wherein the duplex portions of the amplification and blocking strands are at least 95% complementary, at least 98% complementary, at least 99% complementary, or 100% complementary.

5. The method of claim **1**, wherein the nucleic acid is a deoxyribonucleic acid.

6. The method of claim **5**, wherein the method further comprises subjecting a ribonucleic acid to a reverse transcriptase regimen to generate the deoxyribonucleic acid.

7. The method of claim **1** further comprising, prior to step (i), mechanically shearing a nucleic acid preparation to produce the nucleic acid comprising the known target nucleotide sequence.

8. The method of claim **7** further comprising, prior to step (i), end-repairing the nucleic acid to blunt the ends of the nucleic acid produced by mechanically shearing the nucleic acid preparation.

9. The method of claim 7 further comprising, prior to step (i), phosphorylating the nucleic acid.

10. The method of claim 1 further comprising, prior to step (i), adenylating the nucleic acid to produce a 3'-adenosine overhang on the nucleic acid.

11. The method of claim 1, wherein ligating in step (i) comprises performing an overhang ligation reaction.

12. The method of claim 1, wherein ligating in step (i) comprises performing a TA ligation reaction.

13. The method of claim 1, wherein the first target-specific primer can specifically anneal to the known target nucleotide sequence at an annealing temperature in a range of 61° C. to 72° C.

14. The method of claim 1, wherein the known target nucleotide sequence comprises a sequence associated with a gene rearrangement.

15. The method of claim 1, wherein the universal oligonucleotide tail adaptor further comprises a barcode portion.

16. A method for preparing a nucleic acid for sequencing, the method comprising:

(i) ligating a universal oligonucleotide tail adaptor that comprises a first ligatable duplex end and a second unpaired end to a nucleic acid comprising a known target nucleotide sequence to produce a ligation product, the universal oligonucleotide tail adaptor comprising an amplification strand and a blocking strand, wherein a 3' duplex portion of the amplification strand and a 5' duplex portion of the blocking strand are substantially complementary and form the first ligatable duplex end;

(ii) amplifying the ligation product using a first target-specific primer that specifically anneals to the known target nucleotide sequence and a first adaptor primer having a nucleotide sequence identical to a first portion of the amplification strand; and

(iii) amplifying an amplification product of (ii) using a second target-specific primer that specifically anneals to the amplification product of (ii) and a second adaptor primer having a nucleotide sequence identical to a second portion of the amplification strand, wherein ligating in step (i) comprises performing an overhang ligation reaction, and wherein the universal oligonucleotide tail adaptor further comprises a barcode portion.

17. The method of claim 16 further comprising, prior to step (i), phosphorylating the nucleic acid.

18. The method of claim 16 further comprising, prior to step (i), adenylating the nucleic acid to produce a 3'-adenosine overhang on the nucleic acid.

19. The method of claim 16, wherein ligating in step (i) comprises performing a TA ligation reaction.

20. The method of claim 16 further comprising sequencing an amplification product of (iii) using a first sequencing primer and a second sequencing primer.

* * * * *

US010450597B2

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

(10) **Patent No.: US 10,450,597 B2**
(45) **Date of Patent: Oct. 22, 2019**

(54) **METHODS OF PREPARING NUCLEIC ACIDS FOR SEQUENCING**

(71) Applicants: **The General Hospital Corporation**, Boston, MA (US); **ArcherDX, Inc.**, Boulder, CO (US)

(72) Inventors: **Anthony John Iafrate**, Newton, MA (US); **Long Phi Le**, Boston, MA (US); **Zongli Zheng**, Boston, MA (US); **Jason Myers**, Golden, CO (US); **Joshua Stahl**, Boulder, CO (US)

(73) Assignees: **THE GENERAL HOSPITAL CORPORATION**, Boston, MA (US); **ARCHERDX, INC.**, Boulder, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 274 days.

(21) Appl. No.: **14/605,363**

(22) Filed: **Jan. 26, 2015**

(65) **Prior Publication Data**

US 2015/0211050 A1 Jul. 30, 2015

Related U.S. Application Data

(60) Provisional application No. 61/931,959, filed on Jan. 27, 2014.

(51) **Int. Cl.**

C12Q 1/68	(2018.01)
C12Q 1/6806	(2018.01)
C12Q 1/6809	(2018.01)

(52) **U.S. Cl.**

CPC ***C12Q 1/6806*** (2013.01); ***C12Q 1/6809*** (2013.01)

(58) **Field of Classification Search**

None

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,868,104	A	9/1989	Kurn et al.	
5,827,658	A	10/1998	Liang	
6,087,101	A	6/2000	Gruelich et al.	
6,172,214	B1	1/2001	Brenner	
6,576,448	B2	6/2003	Weissman et al.	
6,632,611	B2	10/2003	Su et al.	
6,833,246	B2	12/2004	Balasubramanian	
7,214,490	B2	5/2007	Su et al.	
7,244,559	B2	7/2007	Rothberg et al.	
7,273,730	B2	9/2007	Du Breuil	
7,824,856	B2	11/2010	Monforte	
9,487,828	B2 *	11/2016	Iafrate	C12C 1/6874
10,017,810	B2	7/2018	Iafrate et al.	
2002/0086317	A1	7/2002	Nagayama et al.	
2003/0104432	A1	6/2003	Xu et al.	
2003/0143553	A1	7/2003	Sommer	
2007/0172824	A1 *	7/2007	Chun	C12Q 1/686 435/6.11
2009/0203085	A1 *	8/2009	Kurn	C12C 1/686 435/91.2
2010/0286143	A1	11/2010	Dias-Santagata et al.	
2012/0122701	A1	5/2012	Ryan et al.	
2013/0005585	A1 *	1/2013	Anderson	C12N 15/10 506/2
2013/0231253	A1	9/2013	Amorese et al.	
2013/0303461	A1	11/2013	Iafrate et al.	
2015/0011396	A1	1/2015	Schroeder et al.	
2015/0140553	A1	5/2015	Cushing et al.	
2015/0252361	A1	9/2015	Hayden et al.	
2018/0127806	A1	5/2018	Stahl et al.	
2018/0127807	A1	5/2018	Stahl et al.	
2018/0155767	A1	6/2018	Myers et al.	

FOREIGN PATENT DOCUMENTS

AU	2005203617	A1	9/2005
WO	97/23646	A1	7/1997
WO	97/23647	A1	7/1997
WO	98/28443	A1	7/1998
WO	99/42618	A1	8/1999
WO	00/43543	A1	7/2000
WO	00/70095	A2	11/2000
WO	01/12859	A2	2/2001
WO	01/20035	A2	3/2001
WO	2001/083696	A2	11/2001
WO	02/00938	A2	1/2002
WO	02/29117	A2	4/2002
WO	02/48402	A2	6/2002
WO	02/072772	A2	9/2002
WO	03/078645	A2	9/2003
WO	03/083435	A2	10/2003
WO	2004/011665	A2	2/2004
WO	2004/092418	A2	10/2004
WO	2005/065321	A2	7/2005
WO	2007/030759	A2	3/2007

(Continued)

OTHER PUBLICATIONS

Callaway, AS. et al. A sobemovirus coat protein gene complements long-distance movement of a coat protein-null dianthovirus. Virology, vol. 330, p. 186-195, 2004.*

Ali et al. "Sequence Analysis of TnphoA Insertion Sites in Vibrio cholera Mutants Defective in Rugose Polysaccharide Production", Infection and Immunity 68(12):6857-6864 (2000).

Baetens et al. "Applying Massive Parallel Sequencing to Molecular Diagnosis of Marfan and Loeys-Dietz Syndromes", Human Mutation 32:1053-1062 (2011).

Bohmer, "Novel application for isothermal nucleic acid sequence-based amplification (NASBA)", Journal of Virological Methods 158(1-2):199-201 (2009).

Chang et al., "Identification of a biomarker panel using a multiplex proximity ligation assay improves accuracy of pancreatic cancer diagnosis", Journal of Translational Medicine 7:105 (2009).

(Continued)

Primary Examiner — Suryaprabha Chunduru

(74) *Attorney, Agent, or Firm* — Nixon Peabody LLP; David S. Resnick; Nicole D. Kling

(57) **ABSTRACT**

Aspects of the technology disclosed herein relate to methods for preparing and analyzing nucleic acids. In some embodiments, methods for preparing nucleic acids for sequence analysis (e.g., using next-generation sequencing) are provided herein.

29 Claims, 5 Drawing Sheets

Specification includes a Sequence Listing.

ArcherDX v. Qiagen 18-1019-MN
Joint Trial Exhibit
JX-003

JX-003.1

Appx129

Tecan Ex. 2207, Page 211 of 258
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00029

(56) **References Cited**

FOREIGN PATENT DOCUMENTS

WO	2007/057652	A1	5/2007
WO	2007/136717	A1	11/2007
WO	2008/005459	A2	1/2008
WO	2008/093098	A2	8/2008
WO	WO2001/083696	A2 *	11/2008
WO	2009/102878	A2	8/2009
WO	2009/102896	A2	8/2009
WO	2009/117698	A2	9/2009
WO	2009/133466	A2	11/2009
WO	2009/148617	A2	12/2009
WO	2010/077288	A2	7/2010
WO	2010/083046	A2	7/2010
WO	2011/019964	A1	2/2011
WO	2011/032053	A1	3/2011
WO	2011/053987	A1	5/2011
WO	2011/156529	A2	12/2011
WO	2012/003374	A2	1/2012
WO	2012/040387	A1	3/2012
WO	2012/044956	A1	4/2012
WO	2012/064739	A2	5/2012
WO	2012/103154	A1	8/2012
WO	2013/059746	A1	4/2013
WO	2013/074833	A1	5/2013
WO	2013/112923	A1	8/2013
WO	2013/169339	A1	11/2013
WO	2013/191775	A2	12/2013
WO	2014/144092	A1	9/2014
WO	2014/150931	A1	9/2014
WO	2015/073711	A1	5/2015

OTHER PUBLICATIONS

Chen et al., "Allele-specific copy number profiling by next-generation DNA sequencing", Nucleic Acids Research 43(4):e23 (2015).

Chenchik, et al., "Full-Length cDNA Cloning and Determination of mRNA 5' and 3' Ends by Amplification of Adaptor-Ligated cDNA", BioTechniques 21(3):526-534 (1996).

Compton, "Nucleic acid sequence-based amplification", Nature 350(6313): 91-92 (1991).

Cushing et al., "RVD: a command-line program for ultrasensitive rare single nucleotide variant detection using targeted next-generation DNA resequencing", BMC Research Notes 6:206 (2013).

Dafforn et al., "Linear mRNA amplification from as little as 5 ng total RNA for global gene expression analysis", BioRechniques 37(5):854-857 (2004).

Dahl et al., "Multigene amplification and massively parallel sequencing for cancer mutation discovery", PNAS 104(22):9387-9392 (2007).

Zhang et al., "Detecting simultaneous changepoints in multiple sequences", Biometrika 97(3):631-645 (2010).

Zhang et al., "Titration-free 454 sequencing using Y adapters", Nature Protocols 6(9):1367-1376 (2011).

Lamant et al., "A New Fusion Gene TPM3-ALK in Anaplastic Large Cell Lymphoma Created by a (1;2)(q25;p23) Translocation", Blood 93(9):3088-3095 (1999).

Yuanxin et al., "T-linker-specific ligation PCT (T-linker PCR): an advanced PCR technique for chromosome walking or for isolation of tagged DNA ends", Nucleic Acids Research 31(12):e68 (2003).

Flaherty et al., "Ultrasensitive detection of rare mutations using next-generation targeted resequencing", Nucleic Acids Research 40(1):e2 (2012).

Fredriksson et al., "Multiplex amplification of all coding sequences within 10 cancer genes by Gene-Collector", Nucleic Acids Research 35(7):e47 (2007).

Fredriksson et al., "Multiplexed protein detection by proximity ligation for cancer biomarker validation". Nature Methods 4(4):327-329 (2007).

Grace et al., "Degradable dUMP Outer Primers in Merged Tandem (M/T)-Nested PCR: Low- and Single-Copy DNA Target Amplification", Analytical Biochemistry 263(1):85-92 (1998).

Green et al., "Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma", Blood 121(9):1604-1611 (2013).

Grimes et al., "MendeLIMS: a web-based laboratory information management system for clinical genome sequencing", BMC Bioinformatics 15:290 (2014).

Guled et al., "Array comparative genomic hybridization analysis of olfactory neuroblastoma", Modern Pathology 21(6):770-778 (2008).

Head et al., "Method for improved Illumina sequencing library preparation using NuGEN Ovation RNA-Seq System", Biotechniques 50(3):177-181 (2011).

Hoeijmakers et al., "Linear amplification for deep sequencing", Nature Protocols 6(7):1026-1036 (2011).

Hoon et al., "Aptamer selection by high-throughput sequencing and informatic analysis", Biotechniques 51(6):413-416 (2011).

Hopmans et al., "A programmable method for massively parallel targeted sequencing", Nucleic Acids Research 42(10):e88 (2014).

Ji et al., "Data quality in genomics and microarrays", Nat Biotechnol. 24(9):1112-1113 (2006).

Ji et al., "Identification of a novel deletion mutant strain in *Saccharomyces cerevisiae* that results in a microsatellite instability phenotype", Biodiscovery 1:4 (2012).

Ji, "Improving bioinformatic pipelines for exome variant calling", Genome Medicine 4(1):7 (2012).

Ji et al., "Molecular Inversion Probe Analysis of Gene Copy Alterations Reveals Distinct Categories of Colorectal Carcinoma", Cancer Res. 66(16):7910-7919 (2006).

Ji et al., "Molecular inversion probe assay for allelic quantitation", Methods Mol Biol. 556:67-87 (2009).

Kim et al., "Genetic-based biomarkers and next-generation sequencing: the future of personalized care in colorectal cancer", Per Med. 8(3):331-345 (2011).

Kinde et al., "Detection and quantification of rare mutations with massively parallel sequencing", PNAS 108 (23):9530-95355 (2011).

Kurn et al., "Novel Isothermal, Linear Nucleic Acid Amplification Systems for Highly Multiplexed Applications", Clinical Chemistry 51(10):1973-1981 (2005).

Lam et al., "Performance comparison of whole-genome sequencing platforms", Nat Biotechnol. 30(1):78-82 (2011).

Lee et al., "Systematic genomic identification of colorectal cancer genes delineating advanced from early clinical stage and metastasis", BMC Medical Genomics 6:54 (2013).

Lin et al., "Reproducibility Probability Score—incorporating measurement variability across laboratories for gene selection", Nature Biotechnology 24(12):1476-1477 (2006).

Lishanski et al., "Branch migration inhibition in PCR-amplified DNA: homogeneous mutation detection", Nucleic Acids Research 28(9):e42 (2000).

MAQC Consortium, "The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements", Nat Biotechnol. 24(9):1151-1161 (2006).

Miotke et al., "Correction to High Sensitivity Detection and Quantitation of DNA Copy Number and Single Nucleotide Variants with Single Color Droplet Digital PCR", Anal. Chem. 87:3114 (2015).

Miotke et al., "High Sensitivity Detection and Quantitation of DNA Copy Number and Single Nucleotide Variants with Single Color Droplet Digital PCR", Anal. Chem. 86(5):2618-2624 (2014).

Mugasa et al., "Nucleic Acid Sequence-Based Amplification with Oligochromatography for Detection of Trypanosoma brucei in Clinical Samples", Journal of Clinical Microbiology 47(3):630-635 (2009).

Muralidharan et al., "A cross-sample statistical model for SNP detection in short-read sequencing data", Nucleic Acids Research 40(1):e5 (2012).

Myllykangas et al., "Classification of human cancers based on DNA copy number amplification modeling", BMC Medical Genomics 1:15 (2008).

Myllykangas et al., "DNA copy number amplification profiling of human neoplasms", Oncogene 25(55):7324-7332 (2006).

Myllykangas et al., "Efficient targeted resequencing of human germline and cancer genomes by oligonucleotide-selective sequencing", Nat Biotechnol. 29(11):1024-1027 (2011).

(56) **References Cited**

OTHER PUBLICATIONS

Myllykangas et al., "Integrated gene copy number and expression microarray analysis of gastric cancer highlights potential target genes", Int. J. Cancer 123(4):817-825 (2008).
Myllykangas et al., "Manifestation. mechanisms and mysteries of gene amplifications", Cancer Letters 232(1):79-89 (2006).
Myllykangas et al., "Novel high-throughput sequencing strategies in genetic diagnostics", Duodecim. 129(2):141-148 (2013). English Abstract [Review Finnish].
Myllykangas et al., "Specificity, selection and significance of gene amplifications in cancer", Seminars in Cancer Biology 17(1):42-55 (2007).
Myllykangas et al., "Targeted deep resequencing of the human cancer genome using next-generation technologies", Biotechnol Genet Eng Rev. 27:135-158 (2010).
Myllykangas et al., "Targeted sequencing library preparation by genomic DNA circularization", BMC Biotechnology 11:122 (2011).
Nadauld et al., "Quantitative and Sensitive Detection of Cancer Genome Amplifications from Formalin Fixed Paraffin Embedded Tumors with Droplet Digital PCR", Transl Med (Sunnyvale) 2(2) (2012).
Natsoulis et al., "A Flexible Approach for Highly Multiplexed Candidate Gene Targeted Resequencing", PLoS One. 6(6):e21088 (2011).
Natsoulis et al., "Identification of Insertion Deletion Mutations from Deep Targeted Resequencing", J Data Mining Genomics Proteomics 4(3) (2013).
Newburger et al., "The Human OligoGenome Resource: a database of oligonucleotide capture probes for resequencing target regions across the human genome", Nucleic Acids Research 40(Database issue):D1137-D1143 (2012).
Scheinin et al., "CanGEM: mining gene copy number changes in cancer", Nucleic Acids Research 36(Database issue):D830-835 (2008).
Schiffman et al., "Molecular inversion probes reveal patterns of 9p21 deletion and copy number aberrations in childhood leukemia", Cancer Genet Cytogenet. 193(1):9-18 (2009).
Schiffman et al., "Oncogenic BRAF Mutation with CDKN2A Inactivation is Characteristic of a Subset of Pediatric Malignant Astrocytomas", Cancer Res. 70(2):512-519 (2010).
Shendure et al., "Next-generation DNA sequencing", Nature Biotechnology 26(10):1135-1145 (2008).
Singh et al., "Microarray-based comparison of three amplification methods for nanogram amounts of total RNA", Am J Physiol Cell Physiol. 288(5):C1179-C1189 (2005).
Sinkkonen et al., "Serial Analysis of Gene Expression in the Chicken Otocyst", JARO 12(6):697-710 (2011).
Tong et al., "Genome-scale identification of conditionally essential genes in *E. coli* by DNA microarrays", Biochemical and Biophysical Research Communications 322(1):347-354 (2004).
Turner et al., "Gene Expression Profiling of RNA Extracted from FFPE Tissues: NuGEN Technologies' Whole-Transcriptome Amplification System", Methods Mol Biol. 724:269-280 (2011).
Xiao et al., "Sequential Amplification of Flanking Sequences by Y-shaped Adaptor Dependent Extension Using Multiple Templates", 33(1):85-90 (2007).
Gould et al., "Analysis of sequence variation among smeDEF multi drug efflux pump genes and flanking DNA from defined 16S rRNA subgroups of clinical Stenotrophomonas maltophilia isolates." Journal of Antimicrobial Chemotherapy 54(2):348-353 (2004).
Liu et al., "High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences." Biotechniques Rapid Dispatches 43(5):649-656 (2007).
Wang et al. "Fusion primer and nested integrated PCR (FPNI-PCR): a new high-efficiency strategy for rapid chromosome walking or flanking sequence cloning." BMC Biotechnology 11(1):109 (2011).
Callaway et al., "A Sobemovirus coat protein gene complements long-distance movement of a coat protein-null Dianthovirus." Virology 330(1):186-195 (2004).
Frohman et al., "Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer." PNAS 85(23):8998-9002 (1988).
Kim et al., "FISH-negative cryptic PML-RARA rearrangement detected by long-distance polymerase chain reaction and sequencing analyses: a case study and review of the literature." Cancer Genetics and Cytogenetics 203(2):278-283 (2010).

* cited by examiner

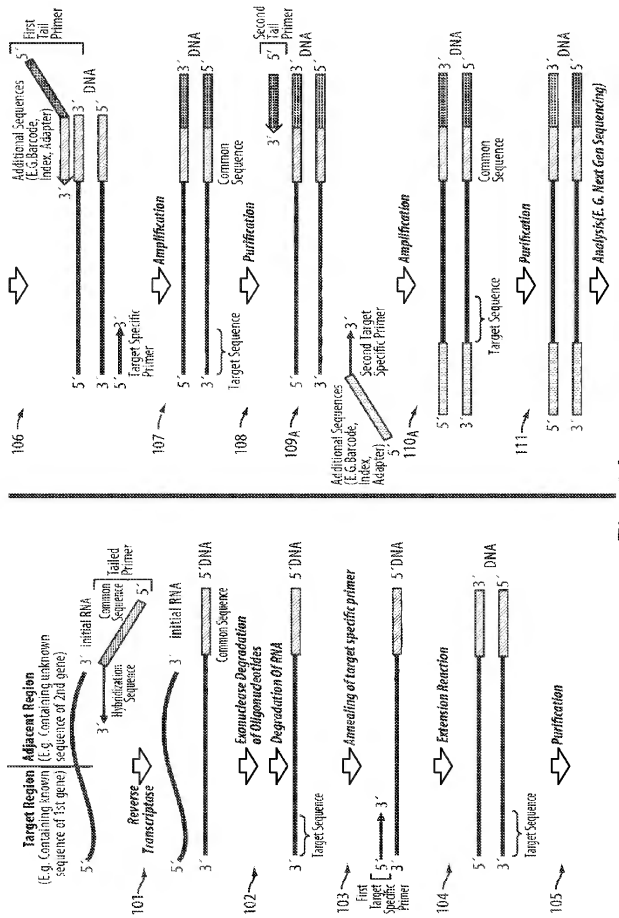


Fig. 1A

Appx132

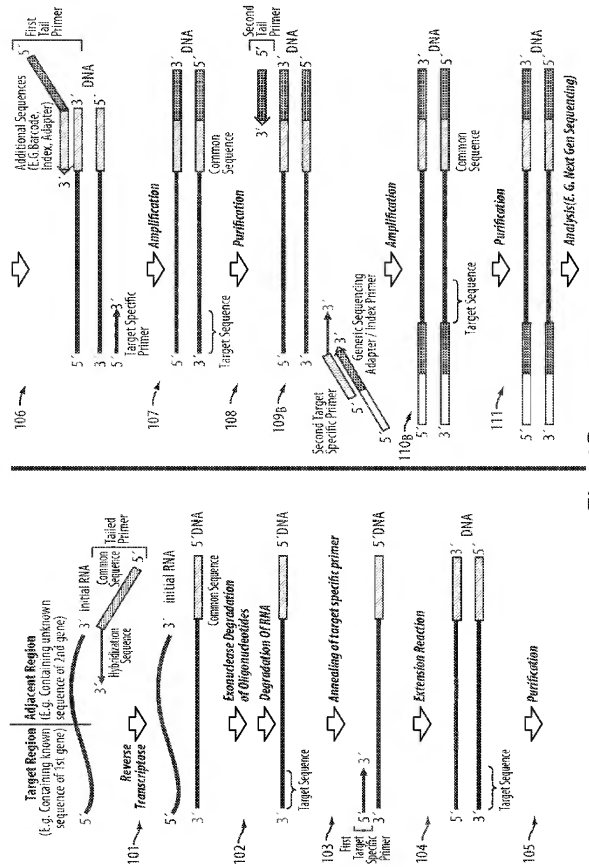


Fig. 1B

Appx133

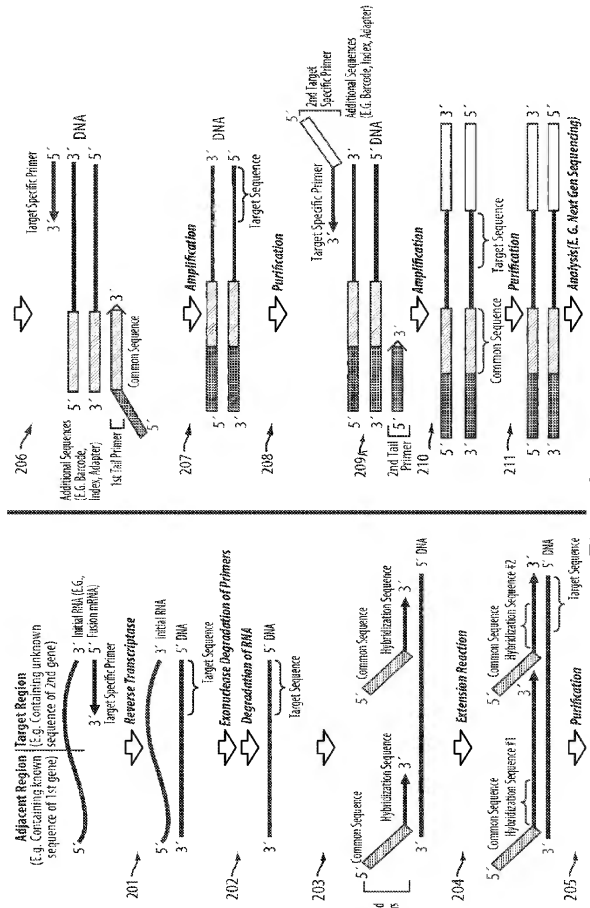

Adjacent Region (E.g. Containing known sequence of 1st gene)
Target Region (E.g. Containing unknown sequence of 2nd gene)
Initial RNA (E.g., Fusion mRNA)
Target Specific Primer
Reverse Transcriptase
Initial RNA
DNA
Target Sequence
Exonuclease Degradation of Primers
Degradation of RNA
Tailed Primers
Common Sequence
Hybridization Sequence
Extension Reaction
Common Sequence #1
Hybridization Sequence #1
Common Sequence #2
Hybridization Sequence #2
Purification
Additional Sequences (E.g. Barcode, Index, Adapter)
1st Tail Primer
Amplification
2nd Tail Primer
2nd Target Specific Primer
Analysis (E. G. Next Gen Sequencing)

Fig. 2A

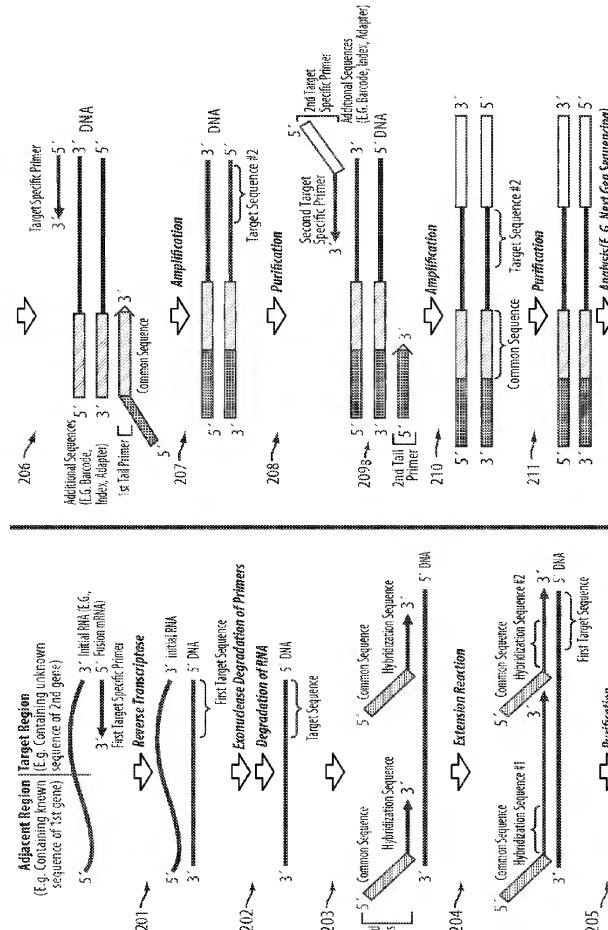


Fig. 2B

Appx135

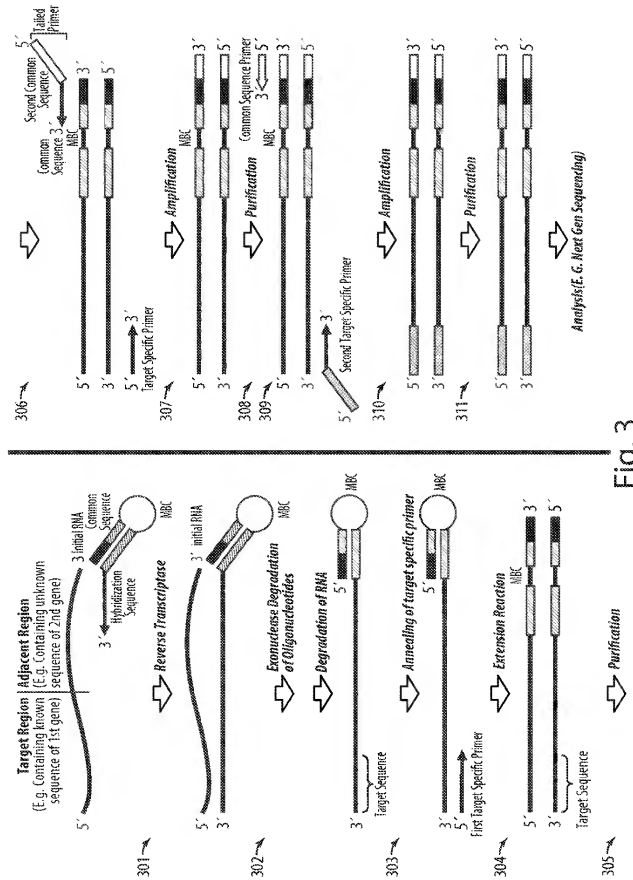

Target Region (Adjacent Region)
(E.g. Containing Known sequence of 1st gene)
(E.g. Containing unknown sequence of 2nd gene)
3' Initial RNA
Common Sequence
Hybridization Sequence
MBC
Reverse Transcriptase
Exonuclease Degradation of Oligonucleotides
Degradation of RNA
Target Sequence
Annealing of target specific primer
First Target Specific Primer
Extension Reaction
Purification
Second Common Sequence
Tailed Primer
Common Sequence
Target Specific Primer
Amplification
Common Sequence Primer
Second Target Specific Primer
Analysis(E. G. Next Gen Sequencing)
301
302
303
304
305
306
307
308
309
310
311

Fig. 3

METHODS OF PREPARING NUCLEIC ACIDS FOR SEQUENCING

RELATED APPLICATIONS

This application claims the benefit under 35 U.S.C. § 119 of U.S. provisional application 61/931,959 filed Jan. 27, 2014, the entire contents of which is incorporated herein by reference.

SEQUENCE LISTING

The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on Mar. 13, 2015, is named 030258-079365-US SL.txt and is 31,392 bytes in size.

TECHNICAL FIELD

The technology described herein relates to methods of preparing and analyzing nucleic acids.

BACKGROUND

Target enrichment prior to next-generation sequencing is more cost-effective than whole genome, whole exome, and whole transcriptome sequencing and therefore more practical for broad implementation; both for research discovery and clinical applications. For example, high coverage depth afforded by target enrichment approaches enables a wider dynamic range for allele counting (in gene expression and copy number assessment) and detection of low frequency mutations, a critical feature for evaluating somatic mutations in cancer. Examples of current enrichment protocols for next generation sequencing include hybridization-based capture assays (TruSeq Capture, Illumina; SureSelect Hybrid Capture, Agilent) and polymerase chain reaction (PCR)-based assays (HaloPlex, Agilent; AmpliSeq, Ion Torrent; TruSeq Amplicon, Illumina; emulsion/digital PCR, Raindance). Hybridization-based approaches capture not only the targeted sequences covered by the capture probes but also near off-target bases that consume sequencing capacity. In addition, these methods are relatively time-consuming, labor-intensive, and suffer from a relatively low level of specificity.

SUMMARY

Aspects of the technology disclosed herein relate to methods for preparing and analyzing nucleic acids. In some embodiments, methods for preparing nucleic acids for sequence analysis (e.g., using next-generating sequencing) are provided herein. In some embodiments, technology described herein is directed to methods of determining nucleotide sequences of nucleic acids. In some embodiments, the methods described herein relate to enriching target nucleic acids prior to sequencing.

Aspects of the technology disclosed herein relate to methods of determining the nucleotide sequence contiguous to a known target nucleotide sequence. In some embodiments, the methods involve (a) contacting a target nucleic acid molecule comprising the known target nucleotide sequence with an initial target-specific primer under hybridization conditions; (b) performing a template-dependent extension reaction that is primed by a hybridized initial target-specific primer and that uses the target nucleic acid molecule as a template; (c) contacting the product of step (b) with a population of tailed random primers under hybridization conditions; (d) performing a template-dependent extension reaction that is primed by a hybridized tailed random primer and that uses the portion of the target nucleic acid molecule downstream of the site of hybridization as a template; (e) amplifying a portion of the target nucleic acid molecule and the tailed random primer sequence with a first tail primer and a first target-specific primer; (f) amplifying a portion of the amplicon resulting from step (e) with a second tail primer and a second target-specific primer; and (g) sequencing the amplified portion from step (f) using a first and second sequencing primer. In some embodiments, the population of tailed random primers comprises single-stranded oligonucleotide molecules having a 5' nucleic acid sequence identical to a first sequencing primer and a 3' nucleic acid sequence comprising from about 6 to about 12 random nucleotides. In some embodiments, the first target-specific primer comprises a nucleic acid sequence that can specifically anneal to the known target nucleotide sequence of the target nucleic acid at the annealing temperature. In some embodiments, the second target-specific primer comprises a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (e), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer and the second target-specific primer is nested with respect to the first target-specific primer. In some embodiments, the first tail primer comprises a nucleic acid sequence identical to the tailed random primer. In some embodiments, the second tail primer comprises a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first tail primer.

In some embodiments. the methods involve (a) contacting a target nucleic acid molecule comprising the known target nucleotide sequence with a population of tailed random primers under hybridization conditions; (b) performing a template-dependent extension reaction that is primed by a hybridized tailed random primer and that uses the portion of the target nucleic acid molecule downstream of the site of hybridization as a template; (c) contacting the product of step (b) with an initial target-specific primer under hybridization conditions; (d) performing a template-dependent extension reaction that is primed by a hybridized initial target-specific primer and that uses the target nucleic acid molecule as a template; (e) amplifying a portion of the target nucleic acid molecule and the tailed random primer sequence with a first tail primer and a first target-specific primer; (f) amplifying a portion of the amplicon resulting from step (e) with a second tail primer and a second target-specific primer; and (g) sequencing the amplified portion from step (f) using a first and second sequencing primer. In some embodiments, the population of tailed random primers comprises single-stranded oligonucleotide molecules having a 5' nucleic acid sequence identical to a first sequencing primer and a 3' nucleic acid sequence comprising from about 6 to about 12 random nucleotides. In some embodiments, the first target-specific primer comprises a nucleic acid sequence that can specifically anneal to the known target nucleotide sequence of the target nucleic acid at the annealing temperature. In some embodiments, the second target-specific primer comprises a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (c), and a 5' portion comprising a nucleic acid sequence that is identical

to a second sequencing primer and the second target-specific primer is nested with respect to the first target-specific primer. In some embodiments, the first tail primer comprises a nucleic acid sequence identical to the tailed random primer. In some embodiments, the second tail primer comprises a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first tail primer.

In some embodiments, the methods further involve a step of contacting the sample and/or products with RNase after extension of the initial target-specific primer. In some embodiments, the tailed random primer can form a hair-pin loop structure. In some embodiments, the initial target-specific primer and the first target-specific primer are identical. In some embodiments, the tailed random primer further comprises a barcode portion comprising 6-12 random nucleotides between the 5' nucleic acid sequence identical to a first sequencing primer and the 3' nucleic acid sequence comprising 6-12 random nucleotides.

In some embodiments, the methods involve (a) contacting a target nucleic acid molecule comprising the known target nucleotide sequence with a population of tailed random primers under hybridization conditions; (b) performing a template-dependent extension reaction that is primed by a hybridized tailed random primer and that uses the portion of the target nucleic acid molecule downstream of the site of hybridization as a template; (c) amplifying a portion of the target nucleic acid molecule and the tailed random primer sequence with a first tail primer and a first target-specific primer; (d) amplifying a portion of the amplicon resulting from step (c) with a second tail primer and a second target-specific primer; and (e) sequencing the amplified portion from step (d) using a first and second sequencing primer. In some embodiments, the population of tailed random primers comprises single-stranded oligonucleotide molecules having a 5' nucleic acid sequence identical to a first sequencing primer; a middle barcode portion comprising; and a 3' nucleic acid sequence comprising from about 6 to about 12 random nucleotides. In some embodiments, the first target-specific primer comprises a nucleic acid sequence that can specifically anneal to the known target nucleotide sequence of the target nucleic acid at the annealing temperature. In some embodiments, the second target-specific primer comprises a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (c), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer and the second target-specific primer is nested with respect to the first target-specific primer. In some embodiments, the first tail primer comprises a nucleic acid sequence identical to the tailed random primer. In some embodiments, the second tail primer comprises a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first tail primer. In some embodiments, the each tailed random primer further comprises a spacer nucleic acid sequence between the 5' nucleic acid sequence identical to a first sequencing primer and the 3' nucleic acid sequence comprising about 6 to about 12 random nucleotides. In certain embodiments, the unhybridized primers are removed from the reaction after an extension step. In some embodiments, the second tail primer is nested with respect to the first tail primer by at least 3 nucleotides. In certain embodiments, the first target-specific primer further comprises a 5' tag sequence portion comprising a nucleic acid sequence of high GC content which is not substantially complementary to or substantially identical to any other portion of any of the primers. In some embodiments, the second tail primer is identical to the full-length first sequencing primer. In certain embodiments, the portions of the target-specific primers that specifically anneal to the known target will anneal specifically at a temperature of about 65° C. in a PCR buffer. In some embodiments, the sample comprises genomic DNA. In some embodiments, the sample comprises RNA and the method further comprises a first step of subjecting the sample to a reverse transcriptase regimen. In certain embodiments, the nucleic acids present in the sample have not been subjected to shearing or digestion. In some embodiments, the sample comprises single-stranded gDNA or cDNA. In certain embodiments, the reverse transcriptase regimen comprises the use of random hexamers. In some embodiments, a gene rearrangement comprises the known target sequence. In certain embodiments, the gene rearrangement is present in a nucleic acid selected from the group consisting of: genomic DNA; RNA; and cDNA. In some embodiments, the gene rearrangement comprises an oncogene. In certain embodiments, the gene rearrangement comprises a fusion oncogene. In some embodiments, the nucleic acid product is sequenced by a next-generation sequencing method. In certain embodiments, the next-generation sequencing method comprises a method selected from the group consisting of: Ion Torrent, Illumina, SOLiD, 454; Massively Parallel Signature Sequencing solid-phase, reversible dye-terminator sequencing; and DNA nanoball sequencing. In certain embodiments, the first and second sequencing primers are compatible with the selected next-generation sequencing method. In some embodiments, the method comprises contacting the sample, or separate portions of the sample, with a plurality of sets of first and second target-specific primers. In certain embodiments, the method comprises contacting a single reaction mixture comprising the sample with a plurality of sets of first and second target-specific primers. In some embodiments, the plurality of sets of first and second target-specific primers specifically anneal to known target nucleotide sequences comprised by separate genes. In certain embodiments, at least two sets of first and second target-specific primers specifically anneal to different portions of a known target nucleotide sequence. In some embodiments, at least two sets of first and second target-specific primers specifically anneal to different portions of a single gene comprising a known target nucleotide sequence. In certain embodiments, at least two sets of first and second target-specific primers specifically anneal to different exons of a gene comprising a known nucleotide target sequence. In some embodiments, the plurality of first target-specific primers comprise identical 5' tag sequence portions. In certain embodiments, each tailed random primer in a population of tailed random primers further comprises an identical sample barcoding portion. In some embodiments, multiple samples are each contacted with a separate population of tailed random primers with a sample barcoding portion; wherein each population of tailed random primers has a distinct sample barcoding portion; and wherein the samples are pooled after step (b). In certain embodiments, each amplification step comprises a set of cycles of a PCR amplification regimen from 5 cycles to 20 cycles in length. In some embodiments, the target-specific primers and the tail primers are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of from about 61 to 72° C. In some embodiments, the target-specific primers and the tail primers are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of about 65° C. In certain embodiments, the target nucleic acid

molecule is from a sample, optionally which is a biological sample obtained from a subject. In some embodiments, the sample is obtained from a subject in need of treatment for a disease associated with a genetic alteration. In certain embodiments, the disease is cancer. In some embodiments, the sample comprises a population of tumor cells. In certain embodiments, the sample is a tumor biopsy. In some embodiments, the cancer is lung cancer. In certain embodiments, a disease-associated gene comprises the known target sequence. In some embodiments, a gene rearrangement product in the sample comprises the known target sequence. In certain embodiments, the gene rearrangement product is an oncogene.

Aspects of the technology disclosed herein relate to methods of preparing nucleic acids for analysis. In some embodiments, the methods involve method (a) contacting a nucleic acid template comprising a first strand of a target nucleic acid with a complementary target-specific primer that comprises a target-specific hybridization sequence, under conditions to promote template-specific hybridization and extension of the target-specific primer; and (b) contacting a nucleic acid template comprising a second strand that is complementary to the first strand of the target nucleic acid with a plurality of different primers that share a common sequence that is 5' to different hybridization sequences, under conditions to promote template-specific hybridization and extension of at least one of the plurality of different primers, in which an extension product is generated to contain both a sequence that is characteristic of the target-specific primer and a sequence that is characteristic of the at least one of the plurality of different primers. In some embodiments, the target nucleic acid is a ribonucleic acid. In certain embodiments, the target nucleic acid is a deoxyribonucleic acid. In some embodiments, steps (a) and (b) are performed sequentially. In certain embodiments, the nucleic acid template in step (a) comprises an extension product resulting from the hybridization and extension of the at least one of the plurality of different primers in step (b). In some embodiments, the nucleic acid template in step (b) comprises an extension product resulting from the hybridization and extension of the target-specific primer in step (a). In certain embodiments, the target nucleic acid is a messenger RNA encoded from a chromosomal segment that comprises a genetic rearrangement. In some embodiments, the target nucleic acid is a chromosomal segment that comprises a portion of a genetic rearrangement. In certain embodiments, the genetic rearrangement is an inversion, deletion, or translocation. In some embodiments, the methods further involve amplifying the extension product. In certain embodiments, the methods further involve contacting the extension product or amplified extension product with an immobilized oligonucleotide under conditions in which hybridization occurs between the extension product and immobilized oligonucleotide. In certain embodiments, the target nucleic acid comprises a target portion having a known sequence and a flanking portion having an unknown sequence. In some embodiments, different hybridization sequences are complementary to the flanking portion. In certain embodiments, the target-specific hybridization sequence is complementary to the target portion. In some embodiments, the target-specific primer further comprises, 5' to the target-specific hybridization sequence, at least one of an index sequence, a barcode sequence and an adaptor sequence. In certain embodiments, the common sequence comprises at least one of an index sequence, barcode sequence and an adaptor sequence. In some embodiments, the adaptor sequence is a cleavable adaptor sequence for immobilizing oligonucleotides in a flow cell.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. **1**A and **1**B depict non-limiting embodiments of a work flow for amplifying and sequencing target nucleic acids that are flanked by a 3' unknown fusion partner, as described herein.

FIGS. **2**A and **2**B depict non-limiting embodiments of a work flow for amplifying sequencing target nucleic acids that are flanked by a 5' unknown fusion partner, as described herein.

FIG. **3** depicts a non-limiting embodiment of a work flow for amplifying sequencing target nucleic acids that are flanked by a 3' unknown fusion pattern using an oligonucleotide comprising a hairpin, as described herein.

DETAILED DESCRIPTION

Aspects of the technology disclosed herein relate to methods for preparing and analyzing nucleic acids. In some embodiments, methods provided herein are useful for determining unknown nucleotide sequences contiguous to (adjacent to) a known target nucleotide sequence. Traditional sequencing methods generate sequence information randomly (e.g. "shotgun" sequencing) or between two known sequences which are used to design primers. In contrast, methods described herein, in some embodiments, allow for determining the nucleotide sequence (e.g. sequencing) upstream or downstream of a single region of known sequence with a high level of specificity and sensitivity. Accordingly, in some embodiments, methods provided herein are useful for determining the sequence of fusions (e.g., fusion mRNAs) that result from gene arrangements (e.g., rearrangements that give rise to cancer or other disorders). In some embodiments, methods provided herein for preparing nucleic acids for analysis (e.g., for sequencing) involve a first round of extension using a target-specific primer that targets a known sequence of a target nucleic acid (e.g., a known sequence of a 1^{st} gene) followed by a second round of extension that involves the use of a heterogenous population of tailed primers (e.g., tailed random primers) that include tailed primers that have hybridization sequences that are complementary with an unknown sequence adjacent to the known sequence in the target nucleic acid. In some embodiments, the tail region of tailed primers includes barcode or index sequences that facilitate multiplex amplification and enrichment of target nucleic acids.

In some aspects of the technology disclosed herein methods are provided of preparing nucleic acids for analysis that involve (a) contacting a nucleic acid template comprising a first strand of a target nucleic acid with a complementary target-specific primer that comprises a target-specific hybridization sequence, under conditions to promote template-specific hybridization and extension of the target-specific primer and (b) contacting a nucleic acid template comprising a second strand that is complementary to the first strand of the target nucleic acid with a plurality of different primers that share a common sequence that is 5' to different hybridization sequences, under conditions to promote template-specific hybridization and extension of at least one of the plurality of different primers, in which an extension product is generated to contain both a sequence that is characteristic of the target-specific primer and a sequence that is characteristic of the at least one of the plurality of

different primers. In some embodiments, steps (a) and (b) above are performed sequentially. In some embodiments, the nucleic acid template in step (a) comprises an extension product resulting from the hybridization and extension of the at least one of the plurality of different primers in step (b). In some embodiments, the nucleic acid template in step (b) comprises an extension product resulting from the hybridization and extension of the target-specific primer in step (a).

In some embodiments, methods are provided for preparing nucleic acids that have a target region 5' to an adjacent region (e.g., an adjacent region of unknown sequence). In some embodiments, methods provided herein can be accomplished using one or more rounds of PCR.

For example, FIGS. **1**A-B present schematics of exemplary methods of amplifying target nucleic acids that have a known target region 5' to an adjacent region (e.g., for purposes of sequencing the adjacent region). At step **101**, initial RNA is obtained or provided in a sample and is used as a template. RNA template is exposed to a plurality of tailed primers (e.g., tailed random primers) that comprise a common sequence that is 5' to different hybridization sequences and shared between all of the tailed primers of the population. In some embodiments, at least one primer hybridizes to an RNA molecule and primes a reverse transcriptase reaction to produce a complementary DNA strand. In step **102**, unhybridized oligonucleotides are degraded (e.g., enzymatically, e.g., by an exonuclease). In step **102**, RNA template is degraded from the complementary DNA strand.

In some embodiments, a tailed primer is provided that hybridizes to the poly-A tail of an RNA molecule. In some embodiments, the sequence of a primer is provided that hybridizes to the poly-A tail comprises a poly-dT (e.g., a 3' positioned stretch of 2 dTs, 3 dTs, 4 dTs, 5 dTs, 6 dTs, 7 dTs, 8 dTs, 9 dTs, 10 dTs, or more.). In some embodiments. a plurality of tailed primers are provided each of which comprises a common sequence. In some embodiments, a plurality of tailed primers is provided each of which further comprises a barcode or index sequence.

It should be appreciated that in methods disclosed herein an RNA template may be degraded by any appropriate method, including, for example, by enzymatic degradation (e.g., using RNaseH, Uracyl glycosylase, etc.), by hydrolysis (e.g., by exposing the RNA to relatively high pH conditions (e.g., pH 10, pH 11, pH 12), etc. In some embodiments, hydrolyzing RNA by exposure to relatively high pH is advantageous because it is relatively inexpensive (compared with certain other methods, e.g., certain enzymatic methods) and because it destroys both RNA and DNA:RNA hybrids. In some embodiments, RNA is degraded by hydrolysis caused by exposure to relatively high pH conditions at a relatively high temperature (e.g., temperature greater than 60° C. (e.g., 60° C. to 95° C.)). In some embodiments, the use of relatively high temperatures is advantageous because it heat inactivates enzymes used in prior preparative steps (e.g., RT enzymes). In some embodiments, an initial nucleic acid may be DNA that is obtained or provided in a sample and is used as a template. In such embodiments, steps **101** and **102** may be omitted.

In step **103**, DNA molecules produced by reverse transcription are contacted by one or more initial target-specific primers which may or may not be the same as the first target-specific primer. In step **104**, hybridization of the initial target-specific primer to a portion of the target nucleic acid (the "target sequence") primes an extension reaction using a DNA molecule as a template to produce a complementary DNA strand. Extension products are purified in step **105**. However, in some embodiments, DNA produced in step **104** may be amplified directly, e.g., by PCR, without purification.

In step **106**, DNA molecules are contacted by a first target-specific primer and a first tail primer. The first target-specific primer hybridizes to a portion of the target nucleic acid. In some embodiments, pools of different first target-specific primers can be used that hybridize to different portions of a target nucleic acid. In some embodiments, use of different target specific primers can be advantegous because it allows for generation of different extension products having overlapping but staggered sequences relative to a target nucleic acid. In some embodiments, different extension products can be sequenced to produce overlapping sequence reads. In some embodiments, overlapping sequence reads can be evaluated to assess accuracy of sequence information, fidelity of nucleic acid amplification, and/or to increase confidence in detecting mutations, such as detecting locations of chromosomal rearrangements (e.g., fusion breakpoints). In some embodiments, pools of different first target-specific primers can be used that hybridize to different portions of different target nucleic acids present in sample. In some embodiments, use of pools of different target-specific primers is advantageous because it facilitates processing (e.g., amplification) and analysis of different target nucleic acids in parallel. In some embodiments, up to 2, up to 3, up to 4, up to 5, up to 6, up to 7, up to 8, up to 9, up to 10, up to 15, up to 20, up to 100 or more pools of different first target-specific primers are used. In some embodiments, 2 to 5, 2 to 10, 5 to 10, 5 to 15, 10 to 15, 10 to 20, 10 to 100, 50 to 100, or more pools of different first target-specific primers are used.

In FIGS. **1**A and **1**B, a first tail primer hybridizes to at least a portion of a DNA molecule provided by the tail portion of the tailed primer of step **101**. In some embodiments, the first tail primer hybridizes to the common sequence provided by the tail of the one or more primers of step **101**. In some embodiments, a nested target specific primer (nested with respect to the target specific primer of step **103**) is used in step **106**. In some embodiments, a first tail primer may comprise an additional sequence 5' to the hybridization sequence that may include barcode, index, adapter sequences, or sequencing primer sites, for example. In step **107**, hybridization of the first target-specific primer and the first tail nucleic acid molecule allows for amplification of a product in a polymerase chain reaction (PCR). In some embodiments, amplified products are purified in step **108**.

In some embodiments, the ssDNA product of step **103** is amplified directly (e.g., by PCR) rather than performing the extension reaction of step **104** and purification of step **105**. Similarly, in some embodiments of any of the methods disclosed herein, ssDNA products may be amplified directly (e.g., by PCR) rather than performing an extension reaction to produce dsDNA prior to purification and/or PCR. In some embodiments, first and second tail primers can be incorporated during a PCR. In some embodiments, primers (e.g., first tail primers, target specific primers) are used in a PCR or extension reaction and then excess primer is removed using a single stranded nuclease. Subsequent rounds of PCR or extension may be performed using different primers (e.g., second tail primers or nested primers or a second target specific primer) to incorporate different sequences into the resulting products.

In some embodiments, an exonuclease (e.g., Exol) may be used to degrade single-stranded DNA. In some embodiments, an exonuclease is used to degrade ssDNA and the

amplified product is processed directly according to steps $\mathbf{109}_{A\text{-}B}$ and $\mathbf{110}_{A\text{-}B}$, without purification at step **108**.

In FIG. **1**A, at step $\mathbf{109}_{A}$, amplified DNA products of step **107** (e.g., as purified in step **108**) are contacted with a second target-specific primer and a second tail primer. In some embodiments, the second target-specific primer hybridizes to a sequence that is present within the template DNA molecule 3' of the sequence of the first target-specific primer such that the reactions are nested. In some embodiments, nesting of the second target-specific primer relative to the first target-specific primer may improve specificity of the hybridization reaction. In some embodiments, the second target-specific primer may comprise an additional sequence 5' to the hybridization sequence that may include barcode, index, adapter sequences, or sequencing primer sites, for example. In step $\mathbf{110}_{A}$, the amplified DNA products of step **107** (e.g., as purified in step **108**) are amplified by PCR in which the extensions are primed by the second target-specific primer and a second tail primer. In some embodiments, a portion of the amplified product from step **107** is further amplified. In some embodiments, a third primer is used that hybridizes to the common tail in the second target specific primer and adds additional sequences such as a barcodes, adapters, etc.

In some embodiments, the second target-specific primer comprises a nucleotide sequence 5' to the target-specific sequence that comprises a barcode, index, or adapter sequences. In some embodiments, the second tail primer hybridizes to a sequence that is present within the template DNA molecule 3' of the sequence of the first tail primer such that the reactions are nested. In such embodiments, a portion of the product from step **106** is amplified. In some embodiments, the second tail primer may comprise additional sequences 5' to the hybridization sequence that may include barcode, index, adapter sequences or sequencing primer sites. Hybridization of the second target-specific primer and the second tail primer allows for exponential amplification of a portion of the target nucleic acid molecule in a PCR reaction.

In some embodiment, the first target-specific primer of step **106** may be used with a second target-specific primer. In such embodiments, hybridization of the first target-specific primer and the second target-specific primer allows for amplification of product in a polymerase chain reaction (PCR). In some embodiments, steps **108-110** may be omitted. The amplification products are purified in reaction **111** and ready for analysis. For example, products purified in step **111** can be sequenced (e.g., using a next generation sequencing platform.) In some embodiments the first target-specific primer or second target-specific primer may comprise additional sequences 5' to a hybridization sequence that may include barcode, index, adapter sequences or sequencing primer sites.

In some embodiments, as depicted in FIG. **1**B, in step $\mathbf{109}_{B}$, DNA products of step **107** (e.g., as purified in step **108**) are contacted with a second target-specific primer and a second tail primer. The second target-specific primer is further contacted by an additional primer (e.g., a primer having 3' sequencing adapter/index sequences) that hybridizes with the common sequence of the second target-specific primer. In some embodiments the additional primer may comprise additional sequences 5' to the hybridization sequence that may include barcode, index, adapter sequences or sequencing primer sites. In some embodiments, the additional primer is a generic sequencing adapter/index primer. In some embodiments, the second target-specific primer may be nested relative to the target-specific primer used in step **107**. In step $\mathbf{110}_{B}$, the DNA products of step **107** (e.g., as purified in step **108**) are amplified by PCR in which the extensions are primed by the second target-specific primer and a second tail primer. Hybridization of the second target-specific primer, the additional primer, and the second tail primer allows for exponential amplification of a portion of the target nucleic acid molecule in a PCR reaction. In such embodiments, a portion of the amplified product from step **108** is amplified.

In some embodiment, the first tail primer of step **106** may be used with a second target-specific primer. In such embodiments, hybridization of the first tail primer and the second target-specific primer allows for amplification of product in a polymerase chain reaction (PCR), optionally with an additional primer (e.g., a primer having 3' sequencing adapter/index sequences) that hybridizes with the common sequence of the second target-specific primer. In some embodiments, steps **108-110** may be omitted.

The products are purified in reaction **111** and ready for analysis. For example, products purified in step **111** can be sequenced (e.g., using a next generation sequencing platform).

In some embodiments, steps **101-104**, **106-107**, and **109-110** are performed consecutively in a single reaction tube without any intervening purification steps. In some embodiments, all of the components involved in steps **101-104**, **106-107**, **109-110** are present at the outset and throughout the reaction. In some embodiments, steps **101-104** are performed consecutively in a single reaction tube. In some embodiments, all of the components involved in steps **101-104** are present at the outset and throughout the reaction. In some embodiments, steps **106-107** are performed consecutively in a single reaction tube. In some embodiments, all of the components involved in steps **106-107** are present at the outset and throughout the reaction. In some embodiments, steps $\mathbf{109}_{A}$-$\mathbf{110}_{A}$ or $\mathbf{109}_{B}$-$\mathbf{110}_{B}$ are performed consecutively in a single reaction tube. In some embodiments, all of the components involved in steps $\mathbf{109}_{A}$-$\mathbf{110}_{A}$ or $\mathbf{109}_{B}$-$\mathbf{110}_{B}$ are present at the outset and throughout the reaction.

In some embodiments, methods are provided for preparing nucleic acids that have a target region 3' to an adjacent region (e.g., an adjacent region of unknown sequence content). For example, FIG. **2** presents a schematic of an exemplary method of amplifying and sequencing target nucleic acids that have a known target region 3' to an adjacent region. In FIGS. **2**A-B, an initial RNA (e.g., a fusion mRNA) is obtained or provided in a sample and is used as a template for the proceeding method. At step **201**, the RNA template is exposed to one or more initial target-specific primers that hybridize to one or more target nucleotide sequences and function to prime a reverse transcription reaction such that a complementary DNA molecule is produced using the initial RNA as a template. In some embodiments, an initial target-specific primer hybridizes to the poly-A tail of an RNA template. In some embodiments, the sequence of the primer that hybridizes to the poly-A tail comprises a poly-dT (e.g., a 3' positioned stretch of 2 dTs, 3 dTs, 4 dTs, 5 dTs, 6 dTs, 7 dTs, 8 dTs, 9 dTs, 10 dTs, or more). In step **202** the unhybridized primers are degraded (e.g., enzymatically, e.g., by an exonuclease). In step **202**, RNA template is degraded from the complementary DNA strand (e.g., enzymatically, e.g., by RNaseH).

In step **203**, DNA molecules that were generated by reverse transcription are contacted by a heterogenous population of tailed primers (e.g., tailed random primers). In some embodiments, the tail portion of each of the tailed primers is a shared or common sequence, identical between

all primers of the population of tailed primers. In some embodiments, at least one primer comprises a hybridization sequence that is complementary to and hybridizes to the target acid template. In step **204**, tailed primers that are hybridized with the template nucleic acid are extended in a template-dependent extension reaction to produce complementary DNA strands that incorporate the tailed primer sequence and the template sequence. The resulting double stranded DNA product is purified in step **205**.

In step **206**, DNA products purified in step **205** are contacted with a first target-specific primer and a first tail primer. The first target-specific primer hybridizes to a target sequence (region) of the DNA. The first tail primer hybridizes to a portion of the DNA molecule characteristic of the tail of the tailed primers of step **203**. In some embodiments, the first tail primer hybridizes to the common sequence provided by the tail of the one or more primers of step **203**. In some embodiments, a first tail primer comprises a barcode or index sequence. In step **207** hybridization of the first target-specific primer and the first tail primer facilitates exponential amplification of a portion of the target nucleic acid molecule in an amplification reaction (e.g., a PCR reaction). In some embodiments, the first tail primer may comprise additional sequences 5' to the hybridization sequence that may include barcode, index, adapter sequences or sequencing primer sites. Amplified products are purified in step **208**. In some embodiments, the purification in step **208** is skipped. For example, in some embodiments, primers (e.g., second target specific primers and 2^{nd} tail primers) are used in a PCR or extension reaction and then excess primer is removed using a single stranded DNA nuclease. Subsequent rounds of PCR or extension may be performed using different primers to incorporate different sequences into the resulting products.

In some embodiments, as depicted in FIG. **2**A at step $\mathbf{209}_A$, DNA molecules produced in step **207** (e.g., as purified in step **208**) are contacted with a second target-specific primer and a second tail primer. In some embodiments, the second target-specific primer is nested relative to the target specific primer used in step **207**. In some embodiments, at least a portion of the product from step $\mathbf{209}_A$ is amplified.

In some embodiments, the second target-specific primer comprises a nucleotide sequence 5' to the target-specific sequence that comprises a barcode, index, or adapter sequences. In some embodiments, the second tail primer hybridizes to a sequence that is present within the template DNA molecule 3' of the sequence of the first tail primer such that the reactions are nested. In some embodiments, at least a portion of the product from step $\mathbf{209}_A$ is amplified.

In some embodiments, first tail primer of step **206** may be used with a second target-specific primer. In such embodiments, hybridization of the first tail primer and the second target-specific primer allows for amplification of product in a polymerase chain reaction (PCR). In such embodiments, steps **208**-**210** may be omitted.

In some embodiments, as depicted in FIG. **2**B at step $\mathbf{209}_B$, DNA molecules produced in step **207** (e.g., as purified in step **208**) are contacted with a second target-specific primer and a second tail primer wherein the second target-specific primer hybridizes to a target sequence that is present within the template DNA molecule 3' of the sequence of the first target-specific primer such that the reactions are nested. In some embodiments, at least a portion of the product from step $\mathbf{209}_B$ is amplified.

In some embodiments, the second target-specific primer is further contacted by an additional primer (e.g., a primer having 3' sequencing adapter/index sequences) that hybridizes with the common sequence of the second target-specific primer. In some embodiments the additional primer may comprise additional sequences 5' to the hybridization sequence that may include barcode, index, adapter sequences or sequencing primer sites. In some embodiments, the additional primer is a generic sequencing adapter/index primer. In such embodiments, hybridization of the second target-specific primer, the additional primer, and a second tail primer allows for exponential amplification of a portion of the target nucleic acid molecule in a PCR reaction.

In some embodiments, the second target-specific primer comprises a nucleotide sequence 5' to the target-specific sequence that comprises a barcode, index, or adapter sequences. In some embodiments, the second tail primer hybridizes to a sequence that is present within the template DNA molecule 3' of the sequence of the first tail primer such that the reactions are nested. In some embodiments, at least a portion of the product from step $\mathbf{209}_B$ is amplified.

In some embodiments, the second tail primer may comprise additional sequence 5' to the hybridization sequence that may include barcode, index, adapter sequences or sequencing primer sites. In step **210**, hybridization of the second target-specific primer and the second tail primer facilitates exponential amplification of a portion of the target nucleic acid molecule in an amplification reaction (e.g., a PCR reaction). The amplification product of step **210** is purified in reaction **211** and ready for analysis. For example, products purified in step **211** can be sequenced (e.g., using a next generation sequencing platform.)

In some embodiments, steps **201**-**204** are performed consecutively in a single reaction tube. In some embodiments, all of the components involve in steps **201**-**204** are present at the outset and throughout the reaction. In some embodiments, steps **206**-**207** are performed consecutively in a single reaction tube. In some embodiments, all of the components involve in steps **206**-**207** are present at the outset and throughout the reaction. In some embodiments, steps $\mathbf{209}_A$-**210** or $\mathbf{209}_B$-**210** are performed consecutively in a single reaction tube. In some embodiments, all of the components involve in steps $\mathbf{209}_A$-**210** or $\mathbf{209}_B$-**210** are present at the outset and throughout the reaction.

In some embodiments, methods provided herein involve use of random sequences as molecular barcodes. In some embodiments, molecular barcodes are built into primers (e.g., RT primers, target-specific primers, extension sequence primers) such that each individual molecule produced by a primer obtains a unique barcode tag. Thus, in some embodiments, the molecular barcode tag permits a determination of whether a sequenced molecular is unique. In some embodiments, molecular barcodes may be used to silence sequencing errors, improve confidence calling of fusions or other mutations, and improved detection limits.

In some embodiments, methods are provided for preparing nucleic acids that have a target region 5' to an adjacent region (e.g., an adjacent region of unknown sequence) using an oligonucleotide comprising a hairpin. In some embodiments, the oligonucleotide may have a structure that is not a hairpin, for example, the oligonucleotide may be linear. For example, FIG. **3** presents a schematic of an exemplary method for amplifying target nucleic acids that have a known target region 5' to an adjacent region (for purposes of sequencing the adjacent region). At step **301**, an initial RNA is obtained or provided in a sample and is used as a template for the proceeding method. The RNA template is exposed to a plurality of hairpin primers (e.g., random primers with a hairpin tail) that comprise a hairpin sequence that is 5' to different hybridization sequences and shared between all of

the primers of the population. The hairpin sequence comprises two complementary common sequences that flank a molecular barcode sequence (MBC). The complementary common sequences base pair to form the stem-loop hairpin structure and protect the MBC sequence. In some embodiments, the plurality of primers is in another structure (not a hairpin) and comprises two complementary common sequences that flank a molecular barcode sequence (MBC). In some embodiments, at least one primer hybridizes to the RNA molecule and primes a reverse transcriptase reaction to produce a complementary DNA strand. In some embodiments, a primer hybridizes to the poly-A tail of the RNA molecule. In some embodiments, the sequence of the primer that hybridizes to the poly-A tail comprises a poly-dT (e.g., a 3' positioned stretch of 2 dTs, 3 dTs, 4 dTs, 5 dTs, 6 dTs, 7 dTs, 8 dTs, 9 dTs, 10 dTs, or more.). In step **302**, any unhybridized oligonucleotides are enzymatically degraded (e.g., by an exonuclease). Also in step **302**, the RNA template is enzymatically degraded from the complementary DNA strand (e.g., by RNaseH).

In step **303**, DNA molecules produced by reverse transcription are contacted by one or more initial target-specific primers which may or may not be the same as the first target-specific primer. In step **304**, hybridization of the first target-specific primer to a portion of the target nucleic acid primes an extension reaction using the DNA molecule as a template to produce a complementary DNA strand. In some embodiments, synthesis of the complementary DNA strand may reduce or eliminate hairpin formation of the complementary common sequences. Extension products are purified in step **305**.

In step **306**, DNA molecules are contacted by a first target-specific primer and a tailed primer. The first target-specific primer hybridizes to a portion of the target nucleic acid. The first tail primer hybridizes to a portion of the DNA molecule provided by the common sequence involve in hairpin formation step **301**. In some embodiments, a nested target-specific primer (e.g., nested with respect to the target-specific primer of step **303**) is used in step **306**. In some embodiments, the first tailed primer may comprise an additional sequence 5' to the hybridization sequence that may include barcode, index, adapter sequences, or sequencing primer sites, for example. In step **307**, hybridization of each of the first target-specific primer and the tailed primer allows for amplification of a portion of the target nucleic acid molecule in a polymerase chain reaction (PCR). In some embodiments, amplified products are purified in step **308**.

In step **309**, amplified DNA products (e.g., those purified in step **308**) are contacted with a second target-specific primer and a common sequence primer. In some embodiments, the second target-specific primer hybridizes to a sequence that is present within the template DNA molecule 3' of the sequence of the first target-specific primer such that the reactions are nested. In some embodiments, the common sequence primer hybridizes with a sequence provided by the first tail primer in step **306**. In some embodiments, in step **310**, DNA products purified in step **308** are amplified by PCR in which the extensions are primed by a second target-specific primer and a common sequence primer. In some embodiments, amplified product from step **308** may be amplified.

In some embodiments, the second target-specific primer comprises a nucleotide sequence 5' to the target-specific sequence that comprises a barcode, index, or adapter sequences. In some embodiments, the second tail primer hybridizes to a sequence that is present within the template DNA molecule 3' of the sequence of the first tail primer such that the reactions are nested. In such embodiments, a portion of the product from step **308** is amplified. In some embodiments, the common sequence primer may comprise additional sequences 5' to the hybridization sequence that may include barcode, index, adapter sequences or sequencing primer sites. Hybridization of the second target-specific primer and the common sequence primer allows for exponential amplification of a portion of the target nucleic acid molecule in a PCR reaction. In some embodiments, products are purified in reaction **311** useful for analysis. For example, products purified in step **311** can be sequenced (e.g., using a next generation sequencing platform.)

In some embodiments, all of the components involved in steps **301-311** are present at the outset and throughout the reaction. In some embodiments, steps **301-304**, **306-307**, and **309-310** are performed consecutively in a single reaction tube without any intervening purification steps. In some embodiments, steps **301-304** are performed consecutively in a single reaction tube. In some embodiments, all of the components involve in steps **301-304** are present at the outset and throughout the reaction. In some embodiments, steps **306-307** are performed consecutively in a single reaction tube. In some embodiments, all of the components involve in steps **306-307** are present at the outset and throughout the reaction. In some embodiments, steps **309-310** are performed consecutively in a single reaction tube. In some embodiments, all of the components involve in steps **309-310** are present at the outset and throughout the reaction.

In some embodiments, methods are provided herein that involve determining the nucleotide sequence contiguous to (adjacent to) a known target nucleotide sequence. In some embodiments, methods comprise contacting a target nucleic acid molecule comprising the known target nucleotide sequence with an initial target-specific primer under suitable hybridization conditions. In some embodiments, the methods further comprise maintaining the target nucleic acid molecule under conditions that promote extension of the hybridized initial target-specific primer (e.g., using the target nucleic acid molecule as a template), thereby producing a first extension product. In some embodiments, the methods further comprise contacting the extension product with a population of tailed random primers under suitable hybridization conditions. In some embodiments, the methods further comprise maintaining the extension product under conditions that promote extension of a hybridized tailed random primer using the portion of the target nucleic acid molecule downstream of the site of hybridization as a template, thereby producing a second extension product. In some embodiments, the methods further comprise amplifying a portion of the target nucleic acid molecule and the tailed random primer sequence with a first tail primer and a first target-specific primer. thereby producing a first amplicon. In some embodiments, the methods further comprise amplifying a portion of the amplicon with a second tail primer and a second target-specific primer, thereby producing a second amplicon.

In some embodiments, one or more target-specific primers used in the methods may be nested with respect to one or more other target-specific primers. For example, in some embodiments, a second target-specific primer is internal to a first target-specific primer. In some embodiments, target-specific primers are the same. In some embodiments, target-specific primers are nested but overlapping with respect to target complementarity. In some embodiments, target-specific primers are nested and non-overlapping. In some embodiments, combinations of identical and nested target

specific primers are used in the same or different amplification steps. In some embodiments, nesting of primers increases target specificity. In some embodiments, the methods further comprise sequencing the second amplicon using a first and second sequencing primer. In some embodiments, the population of tailed random primers comprises single-stranded oligonucleotide molecules having a 5' nucleotide sequence identical to a first sequencing primer and a 3' nucleotide comprising from random nucleotides (e.g., about 6 to about 12 random nucleotides). In some embodiments, the first target-specific primer comprises a nucleic acid sequence that can specifically anneal to the known nucleotide sequence of the target nucleic acid at an appropriate annealing temperature. In some embodiments, the second target-specific primer comprises a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the first amplicon, and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer and the second target-specific primer is nested with respect to the first target-specific primer. In some embodiments, the first tail primer comprises a nucleic acid sequence identical to the common sequence of the tail of the tailed random primer. In some embodiments, the common sequence on the tailed random primer is the exact match of the common sequence on the first tail primer. In some embodiments, the second tail primer comprises a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first tail primer.

As used herein, the term "target nucleic acid" refers to a nucleic acid molecule of interest (e.g., an nucleic acid to be analyzed). In some embodiments, a target nucleic acid comprises both a target nucleotide sequence (e.g., a known or predetermined nucleotide sequence) and an adjacent nucleotide sequence which is to be determined (which may be referred to as an unknown sequence). A target nucleic acid can be of any appropriate length. In some embodiments, a target nucleic acid is double-stranded. In some embodiments, the target nucleic acid is DNA. In some embodiments, the target nucleic acid is genomic or chromosomal DNA (gDNA). In some embodiments, the target nucleic acid can be complementary DNA (cDNA). In some embodiments, the target nucleic acid is single-stranded. In some embodiments, the target nucleic acid can be RNA, e.g., mRNA, rRNA, tRNA, long non-coding RNA, microRNA.

As used herein, the term "known target nucleotide sequence" refers to a portion of a target nucleic acid for which the sequence (e.g. the identity and order of the nucleotide bases of the nucleic acid) is known. For example, in some embodiments, a known target nucleotide sequence is a nucleotide sequence of a nucleic acid that is known or that has been determined in advance of an interrogation of an adjacent unknown sequence of the nucleic acid. A known target nucleotide sequence can be of any appropriate length.

In some embodiments, a target nucleotide sequence (e.g., a known target nucleotide sequence) has a length of 10 or more nucleotides, 30 or more nucleotides, 40 or more nucleotides, 50 or more nucleotides, 100 or more nucleotides, 200 or more nucleotides, 300 or more nucleotides, 400 or more nucleotides, 500 or more nucleotides. In some embodiments, a target nucleotide sequence (e.g., a known target nucleotide sequence) has a length in range of 10 to 100 nucleotides, 10 to 500 nucleotides, 10 to 1000 nucleotides, 100 to 500 nucleotides, 100 to 1000 nucleotides, 500 to 1000 nucleotides, 500 to 5000 nucleotides.

In some embodiments, methods are provided herein for determining sequences of contiguous (or adjacent) portions of a nucleic acid. As used herein, the term "nucleotide sequence contiguous to" refers to a nucleotide sequence of a nucleic acid molecule (e.g., a target nucleic acid) that is immediately upstream or downstream of another nucleotide sequence (e.g., a known nucleotide sequence). In some embodiments, a nucleotide sequence contiguous to a known target nucleotide sequence may be of any appropriate length. In some embodiments, a nucleotide sequence contiguous to a known target nucleotide sequence comprises 1 kb or less of nucleotide sequence, e.g. 1 kb or less of nucleotide sequence, 750 bp or less of nucleotide sequence, 500 bp or less of nucleotide sequence, 400 bp or less of nucleotide sequence, 300 bp or less of nucleotide sequence, 200 bp or less of nucleotide sequence, 100 bp or less of nucleotide sequence. In some embodiments, in which a sample comprises different target nucleic acids comprising a known target nucleotide sequence (e.g. a cell in which a known target nucleotide sequence occurs multiple times in its genome, or on separate, non-identical chromosomes), there may be multiple sequences which comprise "a nucleotide sequence contiguous to" the known target nucleotide sequence. As used herein, the term "determining a (or the) nucleotide sequence," refers to determining the identity and relative positions of the nucleotide bases of a nucleic acid.

In some embodiments of methods disclosed herein one or more tailed random primers are hybridized to a nucleic acid template (e.g., a template comprising a strand of a target nucleic acid). In some embodiments, a target nucleic acid is present in or obtained from a sample comprising a plurality of nucleic acids, one or more of which plurality do not comprise the target nucleic acid. In some embodiments, one or more primers (e.g., one or more tailed random primers) hybridize to substantially all of the nucleic acids in a sample. In some embodiments, one or more primers (e.g., one or more tailed random primers) hybridize to nucleic acids that comprise a target nucleic acid and to nucleic acids that do not comprise the target nucleotide sequence.

Aspects of certain methods disclosed herein relate to contacting a nucleic acid template with a plurality of different primers that share a common sequence that is 5' (or upstream) to different hybridization sequences. In some embodiments the plurality of different primers may be referred to as a population of different primers. In some embodiments, the common sequence may be referred to as a tail, as such the primers are referred to as "tailed primers." In some embodiments, different hybridization sequences of a population comprise nucleotide sequences that occur randomly or pseudorandomly within the population. In some embodiments, nucleotide sequences that occur randomly within a population contain no recognizable regularities, such that, for each nucleotide of each sequence in the population, there is an equal likelihood that the nucleotide comprises a base that is complementary with A, T, G, or C. In such embodiments, it should be appreciated that each nucleotide comprising a base that is complementary with A, T, G, or C may be a naturally occurring nucleotide, a non-naturally occurring nucleotide or a modified nucleotide.

As used herein, a "common sequence" or "shared sequence" refers to a nucleotide sequence that is present in each nucleic acid of a population of nucleic acids. In some embodiments, the common sequence is in a range of about 4 to 75, 4 to 50, 4 to 30, or 4 to 20 nucleotides in length. In some embodiments, the common sequence is 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57,

58, 59, 60, 61. 62. 63. 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, or 75 nucleotides in length.

As used herein, the term "tailed random primer" refers to a single-stranded nucleic acid molecule having a 5' nucleotide sequence (e.g., a 5' nucleotide sequence identical or complementary to a first sequencing primer) and a 3' nucleic acid sequence, in which the 3' nucleotide comprises random nucleotides (e.g., from about 3 to about 15 random nucleotides, about 6 to about 12 random nucleotides). In some embodiments, the 3' nucleotide sequence comprising random nucleotides is at least 6 nucleotides in length, e.g. 6 nucleotides or more, 7 nucleotides or more, 8 nucleotides or more, 9 nucleotides or more, 10 nucleotides or more, 11 nucleotides or more, 12 nucleotides or more, 13 nucleotides or more, 14 nucleotides or more, 15 nucleotides or more, 20 nucleotides or more, 25 nucleotides or more in length. In some embodiments, the 3' nucleotide sequence comprising random nucleotides is 3 to 6 nucleotides in length, 3 to 9 nucleotides in length, 3 to 12 nucleotides in length, 5 to 9 nucleotides in length, 6 to 12 nucleotides in length, 3 to 25 nucleotides in length, 6 to 15 nucleotides in length, or 6 to 25 nucleotides in length. In some embodiments, a tailed random primer can further comprise a spacer between the 5' nucleotide sequence and the 3' nucleotide sequence comprising about 6 to about 12 random nucleotides. In some embodiments, the spacer is a molecular barcode, e.g., that independently tags a template nucleic acid (e.g., a template RNA). In some embodiments, the spacer may be 3 to 6 nucleotides in length, 3 to 12 nucleotides in length, 3 to 25 nucleotides in length, 3 to 45 nucleotides in length, 6 to 12 nucleotides in length, 8 to 16 nucleotides in length, 6 to 25 nucleotides in length, or 6 to 45 nucleotides in length. In some embodiments, for a populations of primers, the spacer is composed of random nucleotides (e.g., in which each of N is independently selected from A, G, C, and T). In some embodiments, the spacer (e.g., a molecular barcode(MBC)) is flanked by two common regions that are complementary. In some embodiments, the complementary common regions base pair to form the stem of a hairpin having a loop portion that comprises the MBC (e.g., as depicted in FIG. **3**). In some embodiments, this hairpin configuration protects the MBC from annealing to other targets inhibiting the RT reaction of the extension reaction in the case of 5' fusions. In some embodiments, a population of tailed random primers can comprise individual primers with varying 3' sequences. In some embodiments, a population of tailed random primers can comprise individual primers with identical 5' nucleotide sequences, e.g., they are all compatible with the same sequencing primer. In some embodiments, a population of tailed random primers can comprise individual primers with varying 5' nucleotide sequences, e.g. an first individual primer is compatible with a first sequencing primer and a second individual primer is compatible with a second sequencing primer.

As used herein, a "hybridization sequence" refers to a sequence of a nucleic acid, such as a portion of a primer, that that has sufficient complementary with a sequence of another nucleic acid (e.g., a template molecule, a target sequence) to enable hybridization between nucleic acid. In some embodiments, the hybridization sequence is about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more nucleotides in length. In some embodiments, the hybridization sequence is in a range of 5 to 50 nucleotides in length, 5 to 40 nucleotides in length, 5 to 35 nucleotides in length, 5 to 30 nucleotides in length, 5 to 25 nucleotides in length, 5 to 20 nucleotides in length, 5 to 15 nucleotides in length, 5 to 10 nucleotides in length, 10 to 40 nucleotides in length, 10 to 30 nucleotides in length, or 10 to 20 nucleotides in length.

In some embodiments, methods described herein comprises an extension regimen or step. In such embodiments, extension may proceed from one or more hybridized tailed random primers, using the nucleic acid molecules which the primers are hybridized to as templates. Extension steps are described herein. In some embodiments, one or more tailed random primers can hybridize to substantially all of the nucleic acids in a sample, many of which may not comprise a known target nucleotide sequence. Accordingly, in some embodiments, extension of random primers may occur due to hybridization with templates that do not comprise a known target nucleotide sequence.

In some embodiments, methods described herein may involve a polymerase chain reaction (PCR) amplification regimen, involving one or more amplification cycles. As used herein, the term "amplification regimen" refers to a process of specifically amplifying (increasing the abundance of) a nucleic acid of interest. In some embodiments, exponential amplification occur when products of a previous polymerase extension serve as templates for successive rounds of extension. In some embodiments, a PCR amplification regimen according to methods disclosed herein may comprise at least one, and in some cases at least 5 or more iterative cycles. In some embodiments each iterative cycle comprises steps of: 1) strand separation (e.g., thermal denaturation); 2) oligonucleotide primer annealing to template molecules; and 3) nucleic acid polymerase extension of the annealed primers. It should be appreciated that any suitable conditions and times involved in each of these steps may be used. In some embodiments, conditions and times selected may depend on the length, sequence content, melting temperature, secondary structural features, or other factors relating to the nucleic acid template and/or primers used in the reaction. In some embodiments, an amplification regimen according to methods described herein is performed in a thermal cycler, many of which are commercially available.

In some embodiments, a nucleic acid extension reaction involves the use of a nucleic acid polymerase. As used herein, the phrase "nucleic acid polymerase" refers an enzyme that catalyzes the template-dependent polymerization of nucleoside triphosphates to form primer extension products that are complementary to the template nucleic acid sequence. A nucleic acid polymerase enzyme initiates synthesis at the 3' end of an annealed primer and proceeds in the direction toward the 5' end of the template. Numerous nucleic acid polymerases are known in the art and commercially available. One group of nucleic acid polymerases are thermostable, i.e., they retain function after being subjected to temperatures sufficient to denature annealed strands of complementary nucleic acids, e.g. 94° C., or sometimes higher. A non-limiting example of a protocol for amplification involves using a polymerase (e.g., Phoenix Taq, VeraSeq) under the following conditions: 98° C. for 30 s, following by 14-22 cycles comprising melting at 98° C. for 10 s, followed by annealing at 68° C. for 30 s, followed by extension at 72° C. 3 min, followed by holding of the reaction at 4° C. However, other appropriate reaction conditions may be used. In some embodiments, annealing/extension temperatures may be adjusted to account for differences in salt concentration (e.g., 3° C. higher to higher salt concentrations). In some embodiments, slowing the ramp rate (e.g., 1° C./s, 0.5° C./s, 0.28° C./s, 0.1° C./s or

slower), for example, from 98° C. to 65° C., improves primer performance and coverage uniformity in highly multiplexed samples.

In some embodiments, a nucleic acid polymerase is used under conditions in which the enzyme performs a template-dependent extension. In some embodiments, the nucleic acid polymerase is DNA polymerase I, Taq polymerase, Pheonix Taq polymerase, Phusion polymerase, T4 polymerase, T7 polymerase, Klenow fragment, Klenow exo-, phi29 polymerase, AMV reverse transcriptase, M-MuLV reverse transcriptase, HIV-1 reverse transcriptase, VeraSeq ULtra polymerase, VeraSeq HF 2.0 polymerase, EnzScript or another appropriate polymerase. In some embodiments, a nucleic acid polymerase is not a reverse transcriptase. In some embodiments, a nucleic acid polymerase acts on a DNA template. In some embodiments, the nucleic acid polymerase acts on an RNA template. In some embodiments, an extension reaction involves reverse transcription performed on an RNA to produce a complementary DNA molecule (RNA-dependent DNA polymerase activity). In some embodiments, a reverse transcriptase is a mouse molony murine leukemia virus (M-MLV) polymerase, AMV reverse transcriptase, RSV reverse transcriptase, HIV-1 reverse transcriptase, HIV-2 reverse transcriptase or another appropriate reverse transcriptase.

In some embodiments, a nucleic acid amplification reaction involves cycles including a strand separation step generally involving heating of the reaction mixture. As used herein, the term "strand separation" or "separating the strands" means treatment of a nucleic acid sample such that complementary double-stranded molecules are separated into two single strands available for annealing to an oligonucleotide primer. In some embodiments, strand separation according to methods described herein is achieved by heating the nucleic acid sample above its melting temperature (T_m). In some embodiments, for a sample containing nucleic acid molecules in a reaction preparation suitable for a nucleic acid polymerase, heating to 94° C. is sufficient to achieve strand separation. In some embodiments, a suitable reaction preparation contains one or more salts (e.g., 1 to 100 mM KCl, 0.1 to 10 $MgCl_2$), at least one buffering agent (e.g., 1 to 20 mM Tris-HCL), and a carrier (e.g., 0.01 to 0.5% BSA). A non-limiting example of a suitable buffer comprises 50 mM KCl, 10 mM Tris-HCl (pH 8.8@25° C.), 0.5 to 3 mM $MgCl_2$, and 0.1% BSA.

In some embodiments, a nucleic acid amplification involves annealing primers to nucleic acid templates having a strands characteristic of a target nucleic acid. In some embodiments, a strand of a target nucleic acid can serve as a template nucleic acid.

As used herein, the term "anneal" refers to the formation of one or more complementary base pairs between two nucleic acids. In some embodiments, annealing involve two complementary or substantially complementary nucleic acids strands hybridizing together. In some embodiments, in the context of an extension reaction annealing involves the hybridize of primer to a template such that a primer extension substrate for a template-dependent polymerase enzyme is formed. In some embodiments, conditions for annealing (e.g., between a primer and nucleic acid template) may vary based of the length and sequence of a primer. In some embodiments, conditions for annealing are based upon a T_m (e.g., a calculated T_m) of a primer. In some embodiments, an annealing step of an extension regimen involves reducing the temperature following strand separation step to a temperature based on the T_m (e.g., a calculated T_m) for a primer, for a time sufficient to permit such annealing. In some embodiments, a T_m can be determined using any of a number of algorithms (e.g., OLIGO™ (Molecular Biology Insights Inc. Colorado) primer design software and VENTRO NTI™ (Invitrogen, Inc. California) primer design software and programs available on the internet, including Primer3, Oligo Calculator, and NetPrimer (Premier Biosoft; Palo Alto, Calif.; and freely available on the world wide web (e.g., at premierbiosoft.com/netpriiner/netprlaunch/Help/xnetpr-launch.html). In some embodiments, the T_m of a primer can be calculated using following formula, which is used by NetPrimer software and is described in more detail in Frieir et al. PNAS 1986 83:9373-9377 which is incorporated by reference herein in its entirety.

$$T_m=\Delta H/(\Delta S+R*\ln(C/4))+16.6\ \log([K^+]/(1+0.7[K^+]))-273.15$$

wherein, ΔH is enthalpy for helix formation; ΔS is entropy for helix formation; R is molar gas constant (1.987 cal/° C.*mol); C is the nucleic acid concentration; and $[K^+]$ is salt concentration. For most amplification regimens, the annealing temperature is selected to be about 5° C. below the predicted T_m, although temperatures closer to and above the T_m (e.g., between 1° C. and 5° C. below the predicted T_m or between 1° C. and 5° C. above the predicted T_m) can be used, as can, for example, temperatures more than 5° C. below the predicted T_m (e.g., 6° C. below, 8° C. below, 10° C. below or lower). In some embodiments, the closer an annealing temperature is to the T_m, the more specific is the annealing. In some embodiments, the time used for primer annealing during an extension reaction (e.g., within the context of a PCR amplification regimen) is determined based, at least in part, upon the volume of the reaction (e.g., with larger volumes involving longer times). In some embodiments, the time used for primer annealing during an extension reaction (e.g., within the context of a PCR amplification regimen) is determined based, at least in part, upon primer and template concentrations (e.g., with higher relative concentrations of primer to template involving less time than lower relative concentrations). In some embodiments, depending upon volume and relative primer/template concentration, primer annealing steps in an extension reaction (e.g., within the context of an amplification regimen) can be in the range of 1 second to 5 minutes, 10 seconds and 2 minutes, or 30 seconds to 2 minutes. As used herein, "substantially anneal" refers to an extent to which complementary base pairs form between two nucleic acids that, when used in the context of a PCR amplification regimen, is sufficient to produce a detectable level of a specifically amplified product.

As used herein, the term "polymerase extension" refers to template-dependent addition of at least one complementary nucleotide, by a nucleic acid polymerase, to the 3' end of an primer that is anneal to a nucleic acid template. In some embodiments, polymerase extension adds more than one nucleotide, e.g., up to and including nucleotides corresponding to the full length of the template. In some embodiments, conditions for polymerase extension are based, at least in part, onthe identity of the polymerase used. In some embodiments, the temperature used for polymerase extension is based upon the known activity properties of the enzyme. In some embodiments, in which annealing temperatures are below the optimal temperatures for the enzyme, it may be acceptable to use a lower extension temperature. In some embodiments, enzymes may retain at least partial activity below their optimal extension temperatures. In some embodiments, a polymerase extension (e.g., performed thermostable polymerases) (e.g., Taq polymerase and variants

thereof) is performed at 65° C. to 75° C. or 68° C. to 72° C. In some embodiments, methods provided herein involve polymerase extension of primers that are anneal to nucleic acid templates at each cycle of a PCR amplification regimen. In some embodiments, a polymerase extension is performed using a polymerase that has relatively strong strand displacement activity. In some embodiments, polymerases having strong strand displacement are useful for preparing nucleic acids for purposes of detecting fusions (e.g., 5' fusions).

In some embodiments, primer extension is performed under conditions that permit the extension of annealed oligonucleotide primers. As used herein, the term "conditions that permit the extension of an annealed oligonucleotide such that extension products are generated" refers to the set of conditions including, for example temperature, salt and co-factor concentrations. pH. and enzyme concentration under which a nucleic acid polymerase catalyzes primer extension. In some embodiments, such conditions are based, at least in part, on the nucleic acid polymerase being used. In some embodiments, a polymerase may perform a primer extension reaction in a suitable reaction preparation. In some embodiments, a suitable reaction preparation contains one or more salts (e.g., 1 to 100 mM KCl, 0.1 to 10 $MgCl_2$), at least one buffering agent (e.g., 1 to 20 mM Tris-HCL), a carrier (e.g., 0.01 to 0.5% BSA) and one or more NTPs (e.g, 10 to 200 uM of each of dATP, dTTP, dCTP, and dGTP). nonlimiting set of conditions is 50 mM KCl, 10 mM Tris-HCl (pH 8.8@25° C.), 0.5 to 3 mM $MgCl_2$, 200 uM each dNTP, and 0.1% BSA at 72° C., under which a polymerase (e.g., Taq polymerase) catalyzes primer extension. In some embodiments, conditions for initiation and extension may include the presence of one, two, three or four different deoxyribonucleoside triphosphates (e.g., selected from dATP, dTTP, dCTP, and dGTP) and a polymerization-inducing agent such as DNA polymerase or reverse transcriptase, in a suitable buffer. In some embodiments, a "buffer" may include solvents (e.g., aqueous solvents) plus appropriate cofactors and reagents which affect pH, ionic strength, etc.).

In some embodiments, nucleic acid amplification involve up to 5, up to 10, up to 20, up to 30, up to 40 or more rounds (cycles) of amplification. In some embodiments, nucleic acid amplification may comprise a set of cycles of a PCR amplification regimen from 5 cycles to 20 cycles in length. In some embodiments, an amplification step may comprise a set of cycles of a PCR amplification regimen from 10 cycles to 20 cycles in length. In some embodiments, each amplification step can comprise a set of cycles of a PCR amplification regimen from 12 cycles to 16 cycles in length. In some embodiments, an annealing temperature can be less than 70° C. In some embodiments, an annealing temperature can be less than 72° C. In some embodiments, an annealing temperature can be about 65° C. In some embodiments, an annealing temperature can be from about 61 to about 72° C.

In various embodiments, methods and compositions described herein relate to performing a PCR amplification regimen with one or more of the types of primers described herein. As used herein, "primer" refers to an oligonucleotide capable of specifically annealing to a nucleic acid template and providing a 3' end that serves as a substrate for a template-dependent polymerase to produce an extension product which is complementary to the template. In some embodiments, a primer useful in methods described herein is single-stranded, such that the primer and its complement can anneal to form two strands. Primers according to methods and compositions described herein may comprise a hybridization sequence (e.g., a sequence that anneals with a nucleic acid template) that is less than or equal to 300 nucleotides in length, e.g., less than or equal to 300, or 250, or 200, or 150, or 100, or 90, or 80, or 70, or 60. or 50, or 40. or 30 or fewer, or 20 or fewer, or 15 or fewer, but at least 6 nucleotides in length. In some embodiments, a hybridization sequence of a primer may be 6 to 50 nucleotides in length, 6 to 35 nucleotides in length, 6 to 20 nucleotides in length, 10 to 25 nucleotides in length.

Any suitable method may be used for synthesizing oligonucleotides and primers. In some embodiments, commercial sources offer oligonucleotide synthesis services suitable for providing primers for use in methods and compositions described herein, e.g. INVITROGEN™ Custom DNA Oligos; Life Technologies; Grand Island, N.Y. or custom DNA Oligos from IDT; Coralville, Iowa).

In some embodiments, after an extension from a tailed random primer has occurred, the extension product and template can be amplified in a first amplification step. In some embodiments. amplification may involve a set of PCR amplification cycles using a first target-specific primer and a first tail primer. In some embodiments, the amplification may result in at least part of the tailed random primer sequence present in the extension product being amplified. In some embodiments, the amplification may result in all of the tailed random primer sequence present in the extension product being amplified.

As used herein, the term "first target-specific primer" refers to a single-stranded oligonucleotide comprising a nucleic acid sequence that can specifically anneal under suitable annealing conditions to a nucleic acid template that has a strand characteristic of a target nucleic acid.

In some embodiments, a primer (e.g., a target specific primer) can comprise a 5' tag sequence portion. In some embodiments, multiple primers (e.g., all first-target specific primers) present in a reaction can comprise identical 5' tag sequence portions. In some embodiments, in a multiplex PCR reaction, different primer species can interact with each other in an off-target manner, leading to primer extension and subsequently amplification by DNA polymerase. In such embodiments, these primer dimers tend to be short, and their efficient amplification can overtake the reaction and dominate resulting in poor amplification of desired target sequence. Accordingly, in some embodiments, the inclusion of a 5' tag sequence in primers (e.g., on target specific primer(s)) may result in formation of primer dimers that contain the same complementary tails on both ends. In some embodiments, in subsequent amplification cycles, such primer dimers would denature into single-stranded DNA primer dimers, each comprising complementary sequences on their two ends which are introduced by the 5' tag. In some embodiments, instead of primer annealing to these single stranded DNA primer dimers, an intra-molecular hairpin (a panhandle like structure) formation may occur due to the proximate accessibility of the complementary tags on the same primer dimer molecule instead of an inter-molecular interaction with new primers on separate molecules. Accordingly, in some embodiments, these primer dimers may be inefficiently amplified, such that primers are not exponentially consumed by the dimers for amplification; rather the tagged primers can remain in high and sufficient concentration for desired specific amplification of target sequences. In some embodiments, accumulation of primer dimers may be undesirable in the context of multiplex amplification because they compete for and consume other reagents in the reaction.

In some embodiments, a 5' tag sequence can be a GC-rich sequence. In some embodiments. a 5' tag sequence may comprise at least 50% GC content, at least 55% GC content,

at least 60% GC content, at least 65% GC content, at least 70% GC content, at least 75% GC content, at least 80% GC content, or higher GC content. In some embodiments, a tag sequence may comprise at least 60% GC content. In some embodiments, a tag sequence may comprise at least 65% GC content.

In some embodiments, a target-specific primer (e.g., a second target-specific primer) is a single-stranded oligonucleotide comprising a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of a known target nucleotide sequence of an amplicon of an amplification reaction, and a 5' portion comprising a tag sequence (e.g., a nucleotide sequence that is identical to or complementary to a sequencing primer (e.g., a second sequencing primer).

In some embodiments, a second target-specific primer of an amplification regimen is nested with respect to a first target-specific primer of the amplification regimen. In some embodiments, the second target-specific primer is nested with respect to the first target-specific primer by at least 3 nucleotides, e.g. by 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more, or 15 or more nucleotides. In some embodiments, all of the target-specific primers (e.g., second target-specific primers) used in an amplification regimen comprise the same 5' portion. In some embodiments, the 5' portion target-specific primer can be configured to suppress primer dimers as described herein.

In some embodiments, first and second target-specific primers are used in an amplification regimen that are substantially complementary to the same strand of a target nucleic acid. In some embodiments, portions of the first and second target-specific primers that specifically anneal to a target sequence (e.g., a known target sequence) can comprise a total of at least 20 unique bases of the known target nucleotide sequence, e.g. 20 or more unique bases, 25 or more unique bases, 30 or more unique bases, 35 or more unique bases, 40 or more unique bases, or 50 or more unique bases. In some embodiments, portions of first and second target-specific primers that specifically anneal to a target sequence (e.g., a known target sequence) can comprise a total of at least 30 unique bases of the known target nucleotide sequence.

As used herein, the term "first tail primer" refers to a nucleic acid molecule comprising a nucleic acid sequence identical to the tail portion of tailed primer.

As used herein, the term "second tail primer" refers to a nucleic acid molecule comprising a nucleic acid sequence identical to a portion of a first sequencing primer, adapter, index primer, etc. and is optionally nested with respect to a first tailed primer. In some embodiments, the second tail primer sits outside of the first tail primer to facilitate addition of appropriate index tags, adapters (e.g., for use in a sequencing platform), etc. In some embodiments, a second tailed primer is identical to a sequencing primer. In some embodiments, a second tailed primer is complementary to a sequencing primer.

In some embodiments, a second tail primer is nested with respect to a first tail primer. In some embodiments, a second tail primer is not nested with respect to a first tail primer. In some embodiments, tail primers of an amplification regimen are nested with respect to one another by at least 3 nucleotides, e.g. by 3 nucleotides, by 4 nucleotides, by 5 nucleotides, by 6 nucleotides, by 7 nucleotides, by 8 nucleotides, by 9 nucleotides, by 10 nucleotides or more.

In some embodiments, a first tail primer comprises a nucleic acid sequence identical to or complementary to the extension product of step (b) strand which is not comprised by the second tail primer and which is located closer to the 5' end of the tailed random primer than any of the sequence identical to or complementary to the second tail primer. Thus, in some embodiments, a second tail primer sits outside of a region added by a random tail primer (5' end), e.g., within the 5' tail added by the first tail primers.

In some embodiments, a first tail primer can comprise a nucleic acid sequence identical to or complementary to a stretch (e.g., of about 20 nucleotides) of the 5'-most nucleotides of a tailed random primer, and a second tail primer can comprise a nucleic acid sequence identical to or complementary to about 30 bases of a tailed random primer, with a 5' nucleotides that is at least 3 nucleotides 3' of the 5' terminus of the tailed random primer.

In some embodiments, use of nested tail primers minimizes or eliminates the production of final amplicons that are amplifiable (e.g. during bridge PCR or emulsion PCR) but cannot be sequenced, a situation that can arise during hemi-nested methods. In some embodiments, hemi-nested approaches using a primer identical to a sequencing primer can result in the carry-over of undesired amplification products from a first PCR step to a second PCR step and may yield artificial sequencing reads. In some embodiments, the use of two tail primers, as described herein can reduce, and in some embodiments eliminate, these problems.

In some embodiments, in a first PCR amplification cycle of a first amplification step, a first target-specific primer can specifically anneal to a template strand of any nucleic acid comprising the known target nucleotide sequence. In some embodiments, depending upon the orientation with which the first target-specific primer was designed, sequence upstream or downstream of the known target nucleotide sequence, and complementary to the template strand will be synthesized. In some embodiments, in which an extension product is formed that comprises the hybridization sequence with which the first target-specific primer forms complementary base pairs, adouble-stranded amplification product can be formed that comprises the first target-specific primer (and the sequence complementary thereto), the target nucleotide sequence downstream of the first target-specific primer (and the sequence complementary thereto), and the tailed random primer sequence (and the sequence complementary thereto). In such embodiments, in subsequent PCR amplification cycles, both the first target-specific primer and the first tail primer are capable of specifically annealling to appropriate strands of the amplification product and the sequence between the known nucleotide target sequence and the tailed random primer can be amplified.

In some embodiments, of methods described herein, a portion of an amplified product (an amplicon) is amplified in further rounds of amplification. In some embodiments, the further rounds of amplification may involve PCR amplification cycles performed using a second target-specific primer and a first sequencing primer or a second tail primer. In some embodiments, a PCR amplification cycles may involve the use of PCR parameters identical to, or which differ from, those of one or more other (e.g., prior) of PCR amplification cycles. In some embodiments, PCR amplification regimens can have the same or different annealing temperatures or the same or different extension step time lengths.

In some embodiments, methods described herein allow for determining the nucleotide sequence contiguous to a known target nucleotide sequence on either or both flanking regions of the known target nucleotide sequence. Regardless of whether the target nucleic acid normally exists as a single-stranded or double-stranded nucleic acid, sequence

information may be represented in a single-stranded format (Strand A), from 5' to 3'. In some embodiments, if the sequence 5' to a known target nucleotide sequence of Strand A is to be determined, gene-specific primers can be complementary to (anneal to) Strand A. If the sequence 3' to a known target nucleotide sequence of Strand A is to be determined, the gene-specific primers can be identical to Strand A, such that they will anneal to the complementary strand of a double-stranded target nucleic acid.

In some embodiments, methods described herein, relating to the use of a first and second gene-specific primer can result in assays with a superior on-target rate, e.g. 70-90%. In some embodiments, the assays and methods described herein can have a target specificity rate of at least 85%.

In some embodiments, primers disclosed herein (e.g., target-specific primers, tail primers) are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of from about 61 to 72° C., e.g. from about 61 to 69° C., from about 63 to 69° C., from about 63 to 67° C., from about 64 to 66° C. In some embodiments, primers disclosed herein are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of less than 72° C. In some embodiments, primers disclosed herein are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of less than 70° C. In some embodiments, primers disclosed herein are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of less than 68° C. In some embodiments, primers disclosed herein are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of about 65° C.

In some embodiments, portions of the target-specific primers that specifically anneal to the known target nucleotide sequence will anneal specifically at a temperature of about 61 to 72° C., e.g. from about 61 to 69° C.. from about 63 to 69° C., from about 63 to 67° C., from about 64 to 66° C. In some embodiments, portions of the target-specific primers that specifically anneal to the known target nucleotide sequence will anneal specifically at a temperature of about 65° C. in a PCR buffer.

In some embodiments, primers described herein do not comprise modified bases (e.g. the primers can not comprise a blocking 3' amine). However, in some embodiments, primers described herein do comprise modified or non-naturally occurring bases. In some embodiments, primers may be modified with a label capable of providing a detectable signal, either directly or indirectly. Non-limiting examples of such labels include radioisotopes, fluorescent molecules, biotin, and others. In some embodiments, primers disclosed herein may include contain a biotin linker or other suitable linker (e.g., for conjugating the primer to a support). In some embodiments, primer may contain a target sequence of an endonucleases such that cleavage with the appropriate enzyme. In other embodiments, the 5' end of a primer may include a sequence that is complementary with a nucleic acid bound to a bead or other support, e.g., a flow cell substrate. Primers may or may not comprise modified internucleoside linkages.

In some embodiments, of methods described herein, nucleic acids (e.g., amplified nucleic acids, extension products, target nucleic acids) can be sequenced. In some embodiments, sequencing can be performed by a next-generation sequencing method. As used herein "next-generation sequencing" refers to oligonucleotide sequencing technologies that have the capacity to sequence oligonucleotides at speeds above those possible with conventional sequencing methods (e.g. Sanger sequencing), due to performing and reading out thousands to millions of sequencing reactions in parallel. Non-limiting examples of next-generation sequencing methods/platforms include Massively Parallel Signature Sequencing (Lynx Therapeutics); 454 pyrosequencing (454 Life Sciences/Roche Diagnostics); solid-phase, reversible dye-terminator sequencing (Solexa/Illumina): SOLiD technology (Applied Biosystems); Ion semiconductor sequencing (ION Torrent); DNA nanoball sequencing (Complete Genomics); and technologies available from Pacific Biosciences, Intelligen Bio-systems, Oxford Nanopore Technologies, and Helicos Biosciences. In some embodiments, the sequencing primers can comprise portions compatible with the selected next-generation sequencing method. Next-generation sequencing technologies and the constraints and design parameters of associated sequencing primers are well known in the art (see, e.g. Shendure, et al., "Next-generation DNA sequencing." Nature, 2008, vol. 26, No. 10, 1135-1145; Mardis, "The impact of next-generation sequencing technology on genetics," Trends in Genetics, 2007, vol. 24, No. 3, pp. 133-141; Su, et al., "Next-generation sequencing and its applications in molecular diagnostics" Expert Rev Mol Diagn, 2011, 11(3):333-43; Zhang et al., "The impact of next-generation sequencing on genomics", J Genet Genomics, 2011, 38(3): 95-109; (Nyren, P. et al. Anal Biochem 208: 17175 (1993); Bentley, D. R. Curr Opin Genet Dev 16:545-52 (2006); Strausberg, R. L., et al. Drug Disc Today 13:569-77 (2008); U.S. Pat. No. 7,282,337; 7,279,563; 7,226,720; 7,220,549; 7,169,560; 6,818,395; 6,911,345; US Pub. Nos. 2006/0252077; 2007/0070349; and 20070070349; which are incorporated by reference herein in their entireties).

In some embodiments, the sequencing step involve the use of a first and second sequencing primers. In some embodiments, the first and second sequencing primers are selected to be compatible with a next-generation sequencing method as described herein.

Methods of aligning sequencing reads to known sequence databases of genomic and/or cDNA sequences are well known in the art and software is commercially available for this process. In some embodiments, reads (less the sequencing primer nucleotide sequence) which do not map, in their entirety, to wild-type sequence databases can be genomic rearrangements or large indel mutations. In some embodiments, reads (less the sequencing primer nucleotide sequence) comprising sequences which map to multiple locations in the genome can be genomic rearrangements.

In some embodiments, primers may contain additional sequences such as an identifier sequence (e.g., a barcode, an index), sequencing primer hybridization sequences (e.g., Rd1), and adapter sequences. In some embodiments the adapter sequences are sequences used with a next generation sequencing system. In some embodiments, the adapter sequences are P5 and P7 sequences for Illumina-based sequencing technology. In some embodiments, the adapter sequence are P1 and A compatible with Ion Torrent sequencing technology.

In some embodiments, as used herein, a "barcode," "molecular barcode," "molecular barcode tag" and "index" may be used interchangeably. generally referring to a nucleotide sequence of a nucleic acid that is useful as an identifier, such as, for example, a source identifier, location identifier, date or time identifier (e.g., date or time of sampling or processing), or other identifier of the nucleic acid. In some embodiments, such barcode or index sequences are useful for identifying different aspects of a

nucleic acid that is present in a population of nucleic acids. In some embodiments, barcode or index sequences may provide a source or location identifier for a target nucleic acid. For example, a barcode or index sequence may serve to identify a patient from whom a nucleic acid is obtained. In some embodiments, barcode or index sequences enable sequencing of multiple different samples on a single reaction (e.g., performed in a single flow cell). In some embodiments, an index sequence can be used to orientate a sequence imager for purposes of detecting individual sequencing reactions. In some embodiments, a barcode or index sequence may be 2 to 25 nucleotides in length, 2 to 15 nucleotides in length, 2 to 10 nucleotides in length, 2 to 6 nucleotides in length. In some embodiments, a barcode or index comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or at least 25 nucleotides.

In some embodiments, when a population of tailed random primers is used in accordance with methods described herein, multiple distinguishable amplification products can be present after amplification. In some embodiments, because tailed random primers hybridize at various positions throughout nucleic acid molecules of a sample, a set of target-specific primers can hybridize (and amplify) the extension products created by more than 1 hybridization event, e.g. one tailed random primer may hybridize at a first distance (e.g., 100 nucleotides) from a target-specific primer hybridization site, and another tailed random primer can hybridize at a second distance (e.g., 200 nucleotides) from a target-specific primer hybridization site, thereby resulting in two amplification products (e.g., a ~100 bp amplification product and a ~200 bp amplification product). In some embodiments, these multiple amplification products can each be sequenced in. In some embodiments, sequencing of these multiple amplification products is advantageous because it provides multiple overlapping sequence reads that can be compare with one another to detect sequence errors introduced during amplification or sequencing processes. In some embodiments, individual amplification products can be aligned and where they differ in the sequence present at a particular base, an artifact or error of PCR and/or sequencing may be present.

In some embodiments, target nucleic acids and/or amplification products thereof can be isolated from enzymes, primers, or buffer components before and/or after any of appropriate step of a method. Any suitable methods for isolating nucleic acids may be usede. In some embodiments, the isolation can comprise Solid Phase Reversible Immobilization (SPRI) cleanup. Methods for SPRI cleanup are well known in the art and kits are commercially available, e.g. Agencourt AMPure XP-PCR Purification (Cat No. A63880, Beckman Coulter; Brea. Calif.). In some embodiments, enzymes can be inactivated by heat treatment.

In some embodiments, unhybridized primers can be removed from a nucleic acid preparation using appropriate methods (e.g., purification, digestion, etc.). In some embodiments, a nuclease (e.g., exonuclease I) is used to remove primer from a preparation. In some embodiments, such nucleases are heat inactivated subsequent to primer digestion. Once the nucleases are inactivated a further set of primers may be added together with other appropriate components (e.g., enzymes, buffers) to perform a further amplification reaction.

In some embodiments, a target nucleic acid genomic DNA or a portion thereof. In some embodiments, a target nucleic acid can be ribonucleic acid (RNA), e.g. mRNA, or a portion thereof. In some embodiments, a target nucleic acid can be a cDNA or a portion thereof.

Many of the sequencing methods suitable for use in methods described herein provide sequencing runs with optimal read lengths of tens to hundreds of nucleotide bases (e.g. Ion Torrent technology can produce read lengths of 200-400 bp). Target nucleic acids may or may not be substantially longer than this optimal read length. In some embodiments, in order for an amplified nucleic acid portion to be of a suitable length for use in a particular sequencing technology, the average distance between the known target nucleotide sequence and an end of the target nucleic acid to which a tailed random primer is hybridizable should be as close to the optimal read length of the selected technology as possible. In some embodiments, if the optimal read-length of a given sequencing technology is 200 bp, then the nucleic acid molecules amplified in accordance with methods described herein should have an average length of about 800 bp. about 700 bp, about 600 bp. about 500 bp, about 400 bp, about 300 bp, about 200 bp or less.

Nucleic acids used herein (e.g., prior to sequencing) can be sheared, e.g. mechanically or enzymatically sheared, to generate fragments of any desired size. Non-limiting examples of mechanical shearing processes include sonication, nebulization, and AFA™ shearing technology available from Covaris (Woburn, Mass.). In some embodiments, a nucleic acid can be mechanically sheared by sonication.

In some embodiments, a target nucleic acid is not sheared or digested. In some embodiments, nucleic acid products of preparative steps (e.g., extension products, amplification products) are not sheared or enzymatically digested.

In some embodiments, when a target nucleic acid an RNA, the sample can be subjected to a reverse transcriptase regimen to generate DNA template and the DNA template can then be sheared. In some embodiments, target RNA can be sheared before performing a reverse transcriptase regimen. In some embodiments, a sample comprising target RNA can be used in methods described herein using total nucleic acids extracted from either fresh or degraded specimens; without the need of genomic DNA removal for cDNA sequencing; without the need of ribosomal RNA depletion for cDNA sequencing; without the need of mechanical or enzymatic shearing in any of the steps; by subjecting the RNA for double-stranded cDNA synthesis using random hexamers.

In some embodiments, a known target nucleic acid can contain a fusion sequence resulting from a gene rearrangement. In some embodiments, methods described herein are suited for determining the presence and/or identity of a gene rearrangement. In some embodiments, identity of one portion of a gene rearrangement is previously known (e.g., the portion of a gene rearrangement that is to be targeted by the gene-specific primers) and the sequence of the other portion may be determined using methods disclosed herein. In some embodiments, a gene rearrangement can involves an oncogene. In some embodiments, a gene rearrangement can comprise a fusion oncogene.

In some embodiments, a target nucleic acid is present in or obtained from an appropriate sample (e.g., a food sample, environmental sample, biological sample e.g., blood sample, etc.). In some embodiments, the is a biological sample obtained from a subject. In some embodiments a sample can be a diagnostic sample obtained from a subject. In some embodiments, a sample can further comprise proteins, cells, fluids, biological fluids, preservatives, and/or other substances. By way of non-limiting example, a sample can be a cheek swab, blood, serum, plasma, sputum, cerebrospinal

fluid, urine, tears, alveolar isolates, pleural fluid, pericardial fluid, cyst fluid, tumor tissue, tissue, a biopsy, saliva, an aspirate, or combinations thereof. In some embodiments, a sample can be obtained by resection or biopsy.

In some embodiments, the sample can be obtained from a subject in need of treatment for a disease associated with a genetic alteration, e.g. cancer or a hereditary disease. In some embodiments, a known target sequence is present in a disease-associated gene.

In some embodiments, a sample is obtained from a subject in need of treatment for cancer. In some embodiments, the sample comprises a population of tumor cells, e.g. at least one tumor cell. In some embodiments, the sample comprises a tumor biopsy, including but not limited to, untreated biopsy tissue or treated biopsy tissue (e.g. formalin-fixed and/or paraffin-embedded biopsy tissue).

In some embodiments, the sample is freshly collected. In some embodiments, the sample is stored prior to being used in methods and compositions described herein. In some embodiments, the sample is an untreated sample. As used herein, "untreated sample" refers to a biological sample that has not had any prior sample pre-treatment except for dilution and/or suspension in a solution. In some embodiments, a sample is obtained from a subject and preserved or processed prior to being utilized in methods and compositions described herein. By way of non-limiting example, a sample can be embedded in paraffin wax, refrigerated, or frozen. A frozen sample can be thawed before determining the presence of a nucleic acid according to methods and compositions described herein. In some embodiments, the sample can be a processed or treated sample. Exemplary methods for treating or processing a sample include, but are not limited to, centrifugation, filtration, sonication, homogenization, heating, freezing and thawing, contacting with a preservative (e.g. anti-coagulant or nuclease inhibitor) and any combination thereof. In some embodiments, a sample can be treated with a chemical and/or biological reagent. Chemical and/or biological reagents can be employed to protect and/or maintain the stability of the sample or nucleic acid comprised by the sample during processing and/or storage. In addition, or alternatively, chemical and/or biological reagents can be employed to release nucleic acids from other components of the sample. By way of non-limiting example, a blood sample can be treated with an anti-coagulant prior to being utilized in methods and compositions described herein. Suitable methods and processes for processing, preservation, or treatment of samples for nucleic acid analysis may be used in the method disclosed herein. In some embodiments, a sample can be a clarified fluid sample, for example, by centrifugation. In some embodiments, a sample can be clarified by low-speed centrifugation (e.g. 3,000×g or less) and collection of the supernatant comprising the clarified fluid sample.

In some embodiments, a nucleic acid present in a sample can be isolated, enriched, or purified prior to being utilized in methods and compositions described herein. Suitable methods of isolating, enriching, or purifying nucleic acids from a sample may be used. For example, kits for isolation of genomic DNA from various sample types are commercially available (e.g. Catalog Nos. 51104, 51304, 56504, and 56404; Qiagen; Germantown, Md.). In some embodiments, methods described herein relate to methods of enriching for target nucleic acids, e.g., prior to a sequencing of the target nucleic acids. In some embodiments, a sequence of one end of the target nucleic acid to be enriched is not known prior to sequencing. In some embodiments, methods described herein relate to methods of enriching specific nucleotide sequences prior to determining the nucleotide sequence using a next-generation sequencing technology. In some embodiments, methods of enriching specific nucleotide sequences do not comprise hybridization enrichment.

Methods described herein can be employed in a multiplex format. In embodiments of methods described herein, multiplex applications can include determining the nucleotide sequence contiguous to one or more known target nucleotide sequences. As used herein, "multiplex amplification" refers to a process involve simultaneous amplification of more than one target nucleic acid in one reaction vessel. In some embodiments, methods involve subsequent determination of the sequence of the multiplex amplification products using one or more sets of primers. Multiplex can refer to the detection of between about 2-1,000 different target sequences in a single reaction. As used herein, multiplex refers to the detection of any range between 2-1,000, e.g., between 5-500, 25-1000, or 10-100 different target sequences in a single reaction, etc. The term "multiplex" as applied to PCR implies that there are primers specific for at least two different target sequences in the same PCR reaction.

In some embodiments, target nucleic acids in a sample, or separate portions of a sample, can be amplified with a plurality of primers (e.g., a plurality of first and second target-specific primers). In some embodiments, the plurality of primers (e.g., a plurality of first and second target-specific primers) can be present in a single reaction mixture, e.g. multiple amplification products can be produced in the same reaction mixture. In some embodiments, the plurality of primers (e.g., a plurality of sets of first and second target-specific primers) can specifically anneal to known target sequences comprised by separate genes. In some embodiments, at least two sets of primers (e.g., at least two sets of first and second target-specific primers) can specifically anneal to different portions of a known target sequence. In some embodiments, at least two sets of primers (e.g., at least two sets of first and second target-specific primers) can specifically anneal to different portions of a known target sequence comprised by a single gene. In some embodiments, at least two sets of primers (e.g., at least two sets of first and second target-specific primers) can specifically anneal to different exons of a gene comprising a known target sequence. In some embodiments, the plurality of primers (e.g., first target-specific primers) can comprise identical 5' tag sequence portions.

In embodiments of methods described herein, multiplex applications can include determining the nucleotide sequence contiguous to one or more known target nucleotide sequences in multiple samples in one sequencing reaction or sequencing run. In some embodiments, multiple samples can be of different origins, e.g. from different tissues and/or different subjects. In such embodiments, primers (e.g., tailed random primers) can further comprise a barcode portion. In some embodiments, a primer (e.g., a tailed random primer) with a unique barcode portion can be added to each sample and ligated to the nucleic acids therein; the samples can subsequently be pooled. In such embodiments, each resulting sequencing read of an amplification product will comprise a barcode that identifies the sample containing the template nucleic acid from which the amplification product is derived.

In some embodiments of methods described herein, a determination of the sequence contiguous to a known oligonucleotide target sequence can provide information relevant to treatment of disease. Thus, in some embodiments, methods disclosed herein can be used to aid in treating

disease. In some embodiments, a sample can be from a subject in need of treatment for a disease associated with a genetic alteration. In some embodiments, a known target sequence a sequence of a disease-associated gene, e.g. an oncogene. In some embodiments, a sequence contiguous to a known oligonucleotide target sequence and/or the known oligonucleotide target sequence can comprise a mutation or genetic abnormality which is disease-associated, e.g. a SNP, an insertion, a deletion, and/or a gene rearrangement. In some embodiments, a sequence contiguous to a known target sequence and/or a known target sequence present in a sample comprised sequence of a gene rearrangement product. In some embodiments, a gene rearrangement can be an oncogene, e.g. a fusion oncogene.

Certain treatments for cancer are particularly effective against tumors comprising certain oncogenes, e.g. a treatment agent which targets the action or expression of a given fusion oncogene can be effective against tumors comprising that fusion oncogene but not against tumors lacking the fusion oncogene. Methods described herein can facilitate a determination of specific sequences that reveal oncogene status (e.g. mutations, SNPs, and/or rearrangements). In some embodiments, methods described herein can further allow the determination of specific sequences when the sequence of a flanking region is known, e.g. methods described herein can determine the presence and identity of gene rearrangements involving known genes (e.g., oncogenes) in which the precise location and/or rearrangement partner are not known before methods described herein are performed.

In some embodiments, technology described herein relates to a method of treating cancer. Accordingly, in some embodiments, methods provided herein may involve detecting, in a tumor sample obtained from a subject in need of treatment for cancer, the presence of one or more oncogene rearrangements: and administering a cancer treatment which is effective against tumors having any of the detected oncogene rearrangements. In some embodiments, technology described herein relates to a method of determining if a subject in need of treatment for cancer will be responsive to a given treatment. Accordingly, in some embodiments, methods provided herein may involve detecting, in a tumor sample obtained from a subject, the presence of an oncogene rearrangement, in which the subject is determined to be responsive to a treatment targeting an oncogene rearrangement product if the presence of the oncogene rearrangement is detected.

In some embodiments, a subject is in need of treatment for lung cancer. In some embodiments, e.g. when the sample is obtained from a subject in need of treatment for lung cancer, the known target sequence can comprise a sequence from a gene selected from the group of ALK, ROS1, and RET. Accordingly, in some embodiments, gene rearrangements result in fusions involving the ALK, ROS1, or RET. Non-limiting examples of gene arrangements involving ALK, ROS1, or RET are described in, e.g., Soda et al. Nature 2007 448561-6: Rikova et al. Cell 2007 131:1190-1203; Kohno et al. Nature Medicine 2012 18:375-7; Takouchi et al. Nature Medicine 2012 18:378-81; which are incorporated by reference herein in their entireties. However, it should be appreciated that the precise location of a gene rearrangement, and the identity of the second gene involved in the rearrangement may not be known in advance. Accordingly, in methods described herein, the presence and identity of such rearrangements can be detected without having to know the location of the rearrangement or the identity of the second gene involved in the gene rearrangement.

In some embodiments, the known target sequence can comprise sequence from a gene selected from the group of: ALK, ROS1. and RET.

In some embodiments, the prescnce of a gene rearrangement of ALK in a sample obtained from a tumor in a subject can indicate that the tumor is susceptible to treatment with a treatment selected from the group consisting of: an ALK inhibitor; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; diamino and aminopyrimidine inhibitors of ALK kinase activity such as NVP-TAE684 and PF-02341066 (see, e.g. Galkin et al, Proc Natl Acad Sci USA, 2007, 104:270-275; Zou et al. Cancer Res, 2007, 67:4408-4417; Hallberg and Palmer F1000 Med Reports 2011 3:21; and Sakamoto et al. Cancer Cell 2011 19:679-690) and molecules disclosed in WO 04/079326. All of the foregoing references are incorporated by reference herein in their entireties. An ALK inhibitor can include any agent that reduces the expression and/or kinase activity of ALK or a portion thereof, including, e.g. oligonucleotides, small molecules, and/or peptides that reduce the expression and/or activity of ALK or a portion thereof. As used herein "anaplastic lymphoma kinase" or "ALK" refers to a transmembrane tyROS line kinase typically involved in neuronal regulation in the wildtype form. The nucleotide sequence of the ALK gene and mRNA are known for a number of species, including human (e.g. SEQ ID NO: 2 (mRNA), NCBI Gene ID: 238).

In some embodiments, the presence of a gene rearrangement of ROS1 in a sample obtained from a tumor in a subject can indicate that the tumor is susceptible to treatment with a treatment selected from the group consisting of: a ROS1 inhibitor and an ALK inhibitor as described herein above (e.g. crizotinib). A ROS1 inhibitor can include any agent that reduces the expression and/or kinase activity of ROS1 or a portion thereof, including, e.g. oligonucleotides, small molecules, and/or peptides that reduce the expression and/or activity of ROS1 or a portion thereof. As used herein "c-ros oncogene 1" or "ROS1" (also referred to in the art as ros-1) refers to a transmembrane tyrosine kinase of the sevenless subfamily and which interacts with PTPN6. Nucleotide sequences of the ROS1 gene and mRNA are known for a number of species, including human (e.g. SEQ ID NO: 1 (mRNA), NCBI Gene ID: 238).

In some embodiments, the presence of a gene rearrangement of RET in a sample obtained from a tumor in a subject can indicate that the tumor is susceptible to treatment with a treatment selected from the group consisting of: a RET inhibitor; DP-2490, DP-3636, SU5416; BAY 43-9006, BAY 73-4506 (regorafenib), ZD6474, NVP-AST487, sorafenib, RPI-1, XL184, vandetanib, sunitinib, imatinib, pazopanib, axitinib, motesanib, gefitinib, and withaferin A (see, e.g. Samadi et al. Surgery 2010 148:1228-36; Cuccuru et al. JNCI 2004 13:1006-1014; Akeno-Stuart et al. Cancer Research 2007 67:6956; Grazma et al. J Clin Oncol 2010 28:15 s 5559; Mologni e tal. J Mol Endocrinol 2006 37:199-212; Calmomagno et al. Journal NCl 2006 98:326-334; Mologni. Curr Med Chem 2011 18:162-175 and the compounds disclosed in WO 06/034833; US Patent Publication 2011/0201598 and U.S. Pat. No. 8,067,434). All of the foregoing references are incorporated by reference herein in their entireties. A RET inhibitor can include any agent that reduces the expression and/or kinase activity of RET or a portion thereof, including, e.g. oligonucleotides, small molecules, and/or peptides that reduce the expression and/or activity of RET or a portion thereof. As used herein "rearranged during transfection" or "RET" refers to a receptor

tyrosine kinase of the cad herein superfamily which is involved in neural crest development and recognizes glial cell line-derived neurotrophic factor family signaling molecules. Nucleotide sequences of the RET gene and mRNA are known for a number of species, including human (e.g. SEQ ID NOs: 3-4 (mRNA), NCBI Gene ID: 5979).

Further non-limiting examples of applications of methods described herein include detection of hematological malignancy markers and panels thereof (e.g. including those to detect genomic rearrangements in lymphomas and leukemias), detection of sarcoma-related genomic rearrangements and panels thereof; and detection of IGH/TCR gene rearrangements and panels thereof for lymphoma testing.

In some embodiments, methods described herein relate to treating a subject having or diagnosed as having, e.g. cancer with a treatment for cancer. Subjects having cancer can be identified by a physician using current methods of diagnosing cancer. For example, symptoms and/or complications of lung cancer which characterize these conditions and aid in diagnosis are well known in the art and include but are not limited to, weak breathing, swollen lymph nodes above the collarbone, abnormal sounds in the lungs, dullness when the chest is tapped, and chest pain. Tests that may aid in a diagnosis of, e.g. lung cancer include, but are not limited to, x-rays, blood tests for high levels of certain substances (e.g. calcium), CT scans, and tumor biopsy. A family history of lung cancer, or exposure to risk factors for lung cancer (e.g. smoking or exposure to smoke and/or air pollution) can also aid in determining if a subject is likely to have lung cancer or in making a diagnosis of lung cancer.

Cancer can include, but is not limited to, carcinoma, including adenocarcinoma, lymphoma, blastoma, melanoma, sarcoma, leukemia, squamous cell cancer. small-cell lung cancer, non-small cell lung cancer, gastrointestinal cancer, Hodgkin's and non Hodgkin's lymphoma, pancreatic cancer, glioblastoma, basal cell carcinoma, biliary tract cancer, bladder cancer, brain cancer including glioblastomas and medulloblastomas; breast cancer, cervical cancer, choriocarcinoma; colon cancer, colorectal cancer, endometrial carcinoma, endometrial cancer; esophageal cancer, gastric cancer; various types of head and neck cancers, intraepithelial neoplasms including Bowen's disease and Paget's disease; hematological neoplasms including acute lymphocytic and myelogenous leukemia; Kaposi's sarcoma, hairy cell leukemia; chromic myelogenous leukemia, AIDS-associated leukemias and adult T-cell leukemia lymphoma; kidney cancer such as renal cell carcinoma, T-cell acute lymphoblastic leukemia/lymphoma, lymphomas including Hodgkin's disease and lymphocytic lymphomas; liver cancer such as hepatic carcinoma and hepatoma, Merkel cell carcinoma, melanoma, multiple myeloma; neuroblastomas; oral cancer including squamous cell carcinoma; ovarian cancer including those arising from epithelial cells, sarcomas including leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibROS 1 arcoma, and osteosarcoma; pancreatic cancer; skin cancer including melanoma, stromal cells, germ cells and mesenchymal cells; pROS ltate cancer, rectal cancer; vulval cancer, renal cancer including adenocarcinoma; testicular cancer including germinal tumors such as seminoma, non-seminoma (teratomas, choriocarcinomas), stromal tumors, and germ cell tumors; thyroid cancer including thyroid adenocarcinoma and medullar carcinoma; esophageal cancer, salivary gland carcinoma, and Wilms' tumors. In some embodiments, the cancer can be lung cancer.

In some embodiments, methods described herein comprise administering an effective amount of compositions described herein, e.g. a treatment for cancer to a subject in order to alleviate a symptom of a cancer. As used herein, "alleviating a symptom of a cancer" is ameliorating any condition or symptom associated with the cancer. As compared with an equivalent untreated control, such reduction is by at least 5%, 10%, 20%, 40%, 50%, 60%, 80%, 90%, 95%, 99% or more as measured by any standard technique. A variety of means for administering the compositions described herein to subjects are known to those of skill in the art. Such methods can include, but are not limited to oral, parenteral, intravenous, intramuscular, subcutaneous, transdermal, airway (aerosol), pulmonary, cutaneous, topical, injection, or intratumoral administration. Administration can be local or systemic. The term "effective amount" as used herein refers to the amount of a treatment needed to alleviate at least one or more symptom of the disease or disorder, and relates to a sufficient amount of pharmacological composition to provide the desired effect. The term "therapeutically effective amount" therefore refers to an amount that is sufficient to effect a particular anti-cancer effect when administered to a typical subject. An effective amount as used herein, in various contexts, would also include an amount sufficient to delay the development of a symptom of the disease, alter the course of a symptom disease (for example but not limited to, slowing the progression of a symptom of the disease), or reverse a symptom of the disease. Thus, it is not generally practicable to specify an exact "effective amount". However, for any given case, an appropriate "effective amount" can be determined by one of ordinary skill in the art using only routine experimentation. The effects of any particular dosage can be monitored by a suitable bioassay. The dosage can be determined by a physician and adjusted, as appropriate, to suit observed effects of the treatment.

Non-limiting examples of a treatment for cancer can include radiation therapy, surgery, gemcitabine, cisplastin, paclitaxel, carboplatin, bortezomib, AMG479, vorinostat, rituximab, temozolomide, rapamycin, ABT-737, PI-103; alkylating agents such as thiotepa and CYTOXAN® cyclosphosphamide; alkyl sulfonates such as busulfan, improsulfan and piposulfan; aziridines such as benzodopa, carboquone, meturedopa, and uredopa; ethylenimines and methylamelamines including altretamine, triethylenemelamine, trietylenephosphoramide, triethiylenethiophosphoramide and trimethylolomelamine; acetogenins (especially bullatacin and bullatacinone); a camptothecin (including the synthetic analogue topotecan); bryostatin; callystatin; CC-1065 (including its adozelesin, carzelesin and bizelesin synthetic analogues); cryptophycins (particularly cryptophycin 1 and cryptophycin 8); dolastatin; duocarmycin (including the synthetic analogues, KW-2189 and CB1-TM1); eleutherobin; pancratistatin; a sarcodictyin; spongistatin; nitrogen mustards such as chlorambucil, chlornaphazine, cholophosphamide, estramustine, ifosfamide, mechlorethamine, mechlorethamine oxide hydrochloride, melphalan, novembichin, phenesterine, prednimustine, trofosfamide, uracil mustard; nitrosureas such as carmustine, chlorozotocin, fotemustine, lomustine, nimustine, and ranimnustine; antibiotics such as the enediyne antibiotics (e.g., calicheamicin, especially calicheamicin gamma1 and calicheamicin omega1 (see, e.g., Agnew, Chem. Intl. Ed. Engl., 33: 183-186 (1994)); dynemicin, including dynemicin A; bisphosphonates, such as clodronate; an esperamicin; as well as neocarzinostatin chromophore and related chromoprotein enediyne antiobiotic chromophores), aclacinomysins, actinomycin, authramycin, azaserine, bleomycins, cactinomycin, carabicin, caminomycin, carzinophilin, chromomycinis. dactinomycin, daunorubicin, detorubicin, 6-di-

JX-003.25

Appx153

azo-5-oxo-L-norleucine, ADRIAMYCIN® doxorubicin (including morpholino-doxorubicin, cyanomorpholino-doxorubicin, 2-pyrrolino-doxorubicin and deoxydoxorubicin), epirubicin, esorubicin, idarubicin, marcellomycin, mitomycins such as mitomycin C, mycophenolic acid, nogalamycin, olivomycins, peplomycin, potfiromycin, puromycin, quelamycin, rodorubicin, streptonigrin, streptozocin, tubercidin, ubenimex, zinostatin, zorubicin; anti-metabolites such as methotrexate and 5-fluorouracil (5-FU); folic acid analogues such as denopterin, methotrexate, pteropterin, trimetrexate; purine analogs such as fludarabine, 6-mercaptopurine, thiamiprine, thioguanine; pyrimidine analogs such as ancitabine, azacitidine, 6-azauridine, carmofur, cytarabine, dideoxyuridine, doxifluridine, enocitabine, floxuridine; androgens such as calusterone, dromostanolone propionate, epitiostanol, mepitiostane, testolactone; anti-adrenals such as aminoglutethimide, mitotane, trilostane; folic acid replenisher such as frolinic acid; aceglatone; aldophosphamide glycoside; aminolevulinic acid; eniluracil; amsacrine; bestrabucil; bisantrene; edatraxate; defofamine; demecolcine; diaziquone; elformithine; elliptinium acetate; an epothilone; etoglucid; gallium nitrate; hydroxyurea; lentinan; lonidainine; maytansinoids such as maytansine and ansamitocins; mitoguazone; mitoxantrone; mopidanmol; nitraerine; pentostatin; phenamet; pirarubicin; losoxantrone; podophyllinic acid; 2-ethylhydrazide; procarbazine; PSK® polysaccharide complex (JHS Natural Products, Eugene, Oreg.); razoxane; rhizoxin; sizofuran; spirogermanium; tenuazonic acid; triaziquone; 2,2',2"-trichlorotriethylamine; trichothecenes (especially T-2 toxin, verracurin A, roridin A and anguidine); urethan; vindesine; dacarbazine; mannomustine; mitobronitol; mitolactol; pipobroman; gacytosine; arabinoside ("Ara-C"); cyclophosphamide; thiotepa; taxoids, e.g., TAXOL® paclitaxel (Bristol-Myers Squibb Oncology, Princeton, N.J.), ABRAXANE® Cremophor-free, albumin-engineered nanoparticle formulation of paclitaxel (American Pharmaceutical Partners, Schaumberg, Ill.), and TAXOTERE® doxetaxel (Rhone-Poulenc Rorer, Antony, France); chloranbucil; GEMZAR® gemcitabine; 6-thioguanine; mercaptopurine; methotrexate; platinum analogs such as cisplatin, oxaliplatin and carboplatin; vinblastine; platinum; etoposide (VP-16); ifosfamide; mitoxantrone; vincristine; NAVELBINE® vinorelbine; novantrone; teniposide; edatrexate; daunomycin; aminopterin; xeloda; ibandronate; irinotecan (Camptosar, CPT-11) (including the treatment regimen of irinotecan with 5-FU and leucovorin); topoisomerase inhibitor RFS 2000; difluoromethylornithine (DMFO); retinoids such as retinoic acid; capecitabine; combretastatin; leucovorin (LV); oxaliplatin, including the oxaliplatin treatment regimen (FOLFOX); lapatinib (Tykerb®); inhibitors of PKC-alpha, Raf, H-Ras, EGFR (e.g., erlotinib (Tarceva®)) and VEGF-A that reduce cell proliferation and pharmaceutically acceptable salts, acids or derivatives of any of the above. In addition, methods of treatment can further include the use of radiation or radiation therapy. Further, methods of treatment can further include the use of surgical treatments.

In some embodiments, methods described herein can be applicable for resequencing, e.g. for confirming particularly relevant, low-quality, and/or complex sequences obtained by non-directed sequencing of a large amount of nucleic acids. By way of non-limiting examples, methods described herein can allow the directed and/or targeted resequencing of targeted disease gene panels (e.g. 10-100 genes), resequencing to confirm variants obtained in large scale sequencing projects, whole exome resequencing, and/or targeted resequencing for detection of single nucleotide variants, multiple nucleotide variants, insertions, deletions, copy number changes, and methylation status.

In some embodiments, methods described herein can allow microbiota sequencing, ancient sample sequencing, and/or new variant virus genotyping.

For convenience, the meaning of some terms and phrases used in the specification, examples, and appended claims, are provided below. Unless stated otherwise, or implicit from context, the following terms and phrases include the meanings provided below. The definitions are provided to aid in describing particular embodiments, and are not intended to limit the claimed invention, because the scope of the invention is limited only by the claims. Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. If there is an apparent discrepancy between the usage of a term in the art and its definition provided herein, the definition provided within the specification shall prevail.

For convenience, certain terms employed herein, in the specification, examples and appended claims are collected here.

The terms "decrease", "reduced", "reduction", or "inhibit" are all used herein generally to mean a decrease by a statistically significant amount. However, for avoidance of doubt, "reduced", "reduction", "decrease", or "inhibit" means a decrease by at least 10% as compared to a reference level, for example a decrease by at least about 20%, or at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90% or up to and including a 100% decrease (e.g. absent level or non-detectable level as compared to a reference level), or any decrease between 10-100% as compared to a reference level. In the context of a marker or symptom is meant a statistically significant decrease in such level. The decrease can be, for example, at least 10%, at least 20%, at least 30%, at least 40% or more, and is preferably down to a level accepted as within the range of normal for an individual without such disorder.

The terms "increased", "increase", "enhance", or "activate" are all used herein to generally mean an increase by a statically significant amount; for the avoidance of doubt, the terms "increased", "increase", "enhance", or "activate" mean an increase of at least 10% as compared to a reference level, for example an increase of at least about 20%, or at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90% or up to and including a 100% increase or any increase between 10-100% as compared to a reference level, or at least about a 2-fold, or at least about a 3-fold, or at least about a 4-fold, or at least about a 5-fold or at least about a 10-fold increase, or any increase between 2-fold and 10-fold or greater as compared to a reference level.

As used herein, a "subject" means a human or animal. Usually the animal is a vertebrate such as a primate, rodent, domestic animal or game animal. Primates include chimpanzees, cynomologous monkeys, spider monkeys, and macaques, e.g., Rhesus. Rodents include mice, rats, woodchucks, ferrets, rabbits and hamsters. Domestic and game animals include cows, horses, pigs, deer, bison, buffalo, feline species, e.g., domestic cat, canine species, e.g., dog, fox, wolf, avian species, e.g., chicken, emu, ostrich, and fish, e.g., trout, catfish and salmon. In some embodiments, the subject is a mammal, e.g., a primate, e.g., a human. The terms, "individual," "patient" and "subject" are used interchangeably herein.

Preferably, the subject is a mammal. The mammal can be a human, non-human primate, mouse, rat, dog, cat, horse, or cow, but is not limited to these examples. Mammals other than humans can be advantageously used as subjects that represent animal models of, e.g. lung cancer. A subject can be male or female.

A subject can be one who has been previously diagnosed with or identified as suffering from or having a condition in need of treatment (e.g. cancer) or one or more complications related to such a condition, and optionally, have already undergone treatment for the condition or the one or more complications related to the condition. Alternatively, a subject can also be one who has not been previously diagnosed as having the condition (e.g. cancer) or one or more complications related to the condition. For example, a subject can be one who exhibits one or more risk factors for the condition or one or more complications related to the condition or a subject who does not exhibit risk factors.

A "subject in need" of treatment for a particular condition can be a subject having that condition, diagnosed as having that condition, or at risk of developing that condition.

As used herein, a "disease associated with a genetic alteration" refers to any disease which is caused by, at least in part, by an alteration in the genetic material of the subject as compared to a healthy wildtype subject, e.g. a deletion, an insertion, a SNP, a gene rearrangement. A disease can be caused by, at least in part, an alteration in the genetic material of the subject if the alteration increases the risk of the subject developing the disease, increases the subject's susceptibility to a disease (including infectious diseases, or diseases with an infectious component), causes the production of a disease-associated molecule, or causes cells to become diseased or abnormal (e.g. loss of cell cycle regulation in cancer cells). Diseases can be associated with multiple genetic alterations, e.g. cancers.

As used herein, the term "nucleic acid" refers to any molecule, preferably a polymeric molecule, incorporating units of ribonucleic acid, deoxyribonucleic acid or an analog thereof. The nucleic acid can be either single-stranded or double-stranded. A single-stranded nucleic acid can be one strand nucleic acid of a denatured double-stranded DNA. Alternatively, it can be a single-stranded nucleic acid not derived from any double-stranded DNA. In one aspect, the template nucleic acid is DNA. In another aspect, the template is RNA. Suitable nucleic acid molecules are DNA, including genomic DNA or cDNA. Other suitable nucleic acid molecules are RNA, including mRNA.

The term "isolated" or "partially purified" as used herein refers, in the case of a nucleic acid, to a nucleic acid separated from at least one other component (e.g., nucleic acid or polypeptide) that is present with the nucleic acid as found in its natural source and/or that would be present with the nucleic acid when expressed by a cell. A chemically synthesized nucleic acid or one synthesized using in vitro transcription/translation is considered "isolated."

As used herein, the term "complementary" refers to the ability of nucleotides to form hydrogen-bonded base pairs. In some embodiment, complementary refers to hydrogen-bonded base pair formation preferences between the nucleotide bases G, A, T, C and U, such that when two given polynucleotides or polynucleotide sequences anneal to each other, A pairs with T and G pairs with C in DNA, and G pairs with C and A pairs with U in RNA. As used herein, "substantially complementary" refers to a nucleic acid molecule or portion thereof (e.g. a primer) having at least 90% complementarity over the entire length of the molecule or portion thereof with a second nucleotide sequence, e.g. 90% complementary, 95% complementary, 98% complementary, 99% complementary, or 100% complementary. As used herein, "substantially identical" refers to a nucleic acid molecule or portion thereof having at least 90% identity over the entire length of a the molecule or portion thereof with a second nucleotide sequence, e.g. 90% identity, 95% identity, 98% identity, 99% identity, or 100% identity.

As used herein, "specific" when used in the context of a primer specific for a target nucleic acid refers to a level of complementarity between the primer and the target such that there exists an annealing temperature at which the primer will anneal to and mediate amplification of the target nucleic acid and will not anneal to or mediate amplification of non-target sequences present in a sample.

As used herein, "amplified product", "amplification product", or "amplicon" refers to oligonucleotides resulting from an amplificationreaction that are copies of a portion of a particular target nucleic acid template strand and/or its complementary sequence, which correspond in nucleotide sequence to the template nucleic acid sequence and/or its complementary sequence. An amplification product can further comprise sequence specific to the primers and which flanks sequence which is a portion of the target nucleic acid and/or its complement. An amplified product, as described herein will generally be double-stranded DNA, although reference can be made to individual strands thereof.

As used herein, a "portion" of a nucleic acid molecule refers to contiguous set of nucleotides comprised by that molecule. A portion can comprise all or only a subset of the nucleotides comprised by the molecule. A portion can be double-stranded or single-stranded.

As used herein, the terms "treat," "treatment," "treating," or "amelioration" refer to therapeutic treatments, wherein the object is to reverse, alleviate. ameliorate, inhibit, slow down or stop the progression or severity of a condition associated with a disease or disorder, e.g. lung cancer. The term "treating" includes reducing or alleviating at least one adverse effect or symptom of a condition, disease or disorder associated with a condition. Treatment is generally "effective" if one or more symptoms or clinical markers are reduced. Alternatively, treatment is "effective" if the progression of a disease is reduced or halted. That is, "treatment" includes not just the improvement of symptoms or markers, but also a cessation of, or at least slowing of, progress or worsening of symptoms compared to what would be expected in the absence of treatment. Beneficial or desired clinical results include, but are not limited to, alleviation of one or more symptom(s), diminishment of extent of disease, stabilized (i.e., not worsening) state of disease, delay or slowing of disease progression, amelioration or palliation of the disease state, remission (whether partial or total), and/or decreased mortality, whether detectable or undetectable. The term "treatment" of a disease also includes providing relief from the symptoms or side-effects of the disease (including palliative treatment).

The term "statistically significant" or "significantly" refers to statistical significance and generally means a two standard deviation (2SD) below normal, or lower, concentration of the marker.

Other than in the operating examples, or where otherwise indicated, all numbers expressing quantities of ingredients or reaction conditions used herein should be understood as modified in all instances by the term "about." The term "about" when used in connection with percentages can mean±1%.

As used herein the term "comprising" or "comprises" is used in reference to compositions, methods, and respective

component(s) thereof, that are essential to method or composition, yet open to the inclusion of unspecified elements, whether essential or not.

The term "consisting of" refers to compositions, methods, and respective components thereof as described herein, which are exclusive of any element not recited in that description of the embodiment.

As used herein the term "consisting essentially of" refers to those elements required for a given embodiment. The term permits the presence of elements that do not materially affect the basic and novel or functional characteristic(s) of that embodiment.

The singular terms "a," "an," and "the" include plural referents unless context clearly indicates otherwise. Similarly, the word "or" is intended to include "and" unless the context clearly indicates otherwise. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of this disclosure, suitable methods and materials are described below. The abbreviation, "e.g." is derived from the Latin exempli gratia, and is used herein to indicate a non-limiting example. Thus, the abbreviation "e.g." is synonymous with the term "for example."

Definitions of common terms in cell biology and molecular biology can be found in "The Merck Manual of Diagnosis and Therapy", 19th Edition, published by Merck Research Laboratories, 2006 (ISBN 0-911910-19-0); Robert S. Porter et al. (eds.), The Encyclopedia of Molecular Biology, published by Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9). Definitions of common terms in molecular biology can also be found in Benjamin Lewin, Genes X, published by Jones & Bartlett Publishing, 2009 (ISBN-10: 0763766321); Kendrew et al. (eds.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8) and Current Protocols in Protein Sciences 2009, Wiley Intersciences, Coligan et al., eds.

Unless otherwise stated, the present invention was performed using standard procedures, as described, for example in Sambrook et al., Molecular Cloning: A Laboratory Manual (3 ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., USA (2001); and Davis et al., Basic Methods in Molecular Biology, Elsevier Science Publishing, Inc., New York, USA (1995) which are all incorporated by reference herein in their entireties.

Other terms are defined herein within the description of the various aspects of the invention.

All patents and other publications; including literature references, issued patents, published patent applications, and co-pending patent applications; cited throughout this application are expressly incorporated herein by reference for the purpose of describing and disclosing, for example, methodologies described in such publications that might be used in connection with technology described herein. These publications are provided solely for their disclosure prior to the filing date of the present application. Nothing in this regard should be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention or for any other reason. All statements as to the date or representation as to the contents of these documents is based on the information available to the applicants and does not constitute any admission as to the correctness of the dates or contents of these documents.

The description of embodiments of the disclosure is not intended to be exhaustive or to limit the disclosure to the precise form disclosed. While specific embodiments of, and examples for, the disclosure are described herein for illustrative purposes, various equivalent modifications are possible within the scope of the disclosure, as those skilled in the relevant art will recognize. For example, while method steps or functions are presented in a given order, alternative embodiments may perform functions in a different order, or functions may be performed substantially concurrently. The teachings of the disclosure provided herein can be applied to other procedures or methods as appropriate. The various embodiments described herein can be combined to provide further embodiments. Aspects of the disclosure can be modified, if appropriate, to employ the compositions, functions and concepts of the above references and application to provide yet further embodiments of the disclosure. These and other changes can be made to the disclosure in light of the detailed description. All such modifications are intended to be included within the scope of the appended claims.

Specific elements of any of the foregoing embodiments can be combined or substituted for elements in other embodiments. Furthermore, while advantages associated with certain embodiments of the disclosure have been described in the context of these embodiments, other embodiments may also exhibit such advantages, and not all embodiments need necessarily exhibit such advantages to fall within the scope of the disclosure.

Technology described herein is further illustrated by the following examples which in no way should be construed as being further limiting.

Some embodiments of the technology described herein can be defined according to any of the following numbered paragraphs:

1. A method of determining the nucleotide sequence contiguous to a known target nucleotide sequence, the method comprising,
   (a) contacting a target nucleic acid molecule comprising the known target nucleotide sequence with an initial target-specific primer under hybridization conditions;
   (b) performing a template-dependent extension reaction that is primed by a hybridized initial target-specific primer and that uses the target nucleic acid molecule as a template;
   (c) contacting the product of step (b) with a population of tailed random primers under hybridization conditions;
   (d) performing a template-dependent extension reaction that is primed by a hybridized tailed random primer and that uses the portion of the target nucleic acid molecule downstream of the site of hybridization as a template;
   (e) amplifying a portion of the target nucleic acid molecule and the tailed random primer sequence with a first tail primer and a first target-specific primer;
   (f) amplifying a portion of the amplicon resulting from step (e) with a second tail primer and a second target-specific primer;
   (g) sequencing the amplified portion from step (f) using a first and second sequencing primer;
   wherein the population of tailed random primers comprises single-stranded oligonucleotide molecules having a 5' nucleic acid sequence identical to a first sequencing primer and a 3' nucleic acid sequence comprising from about 6 to about 12 random nucleotides;
   wherein the first target-specific primer comprises a nucleic acid sequence that can specifically anneal to the

known target nucleotide sequence of the target nucleic acid at the annealing temperature;

wherein the second target-specific primer comprises a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (e), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer and the second target-specific primer is nested with respect to the first target-specific primer;

wherein the first tail primer comprises a nucleic acid sequence identical to the tailed random primer; and

wherein the second tail primer comprises a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first tail primer.

2. A method of determining the nucleotide sequence contiguous to a known target nucleotide sequence, the method comprising;

(a) contacting a target nucleic acid molecule comprising the known target nucleotide sequence with a population of tailed random primers under hybridization conditions;

(b) performing a template-dependent extension reaction that is primed by a hybridized tailed random primer and that uses the portion of the target nucleic acid molecule downstream of the site of hybridization as a template;

(c) contacting the product of step (b) with an initial target-specific primer under hybridization conditions;

(d) performing a template-dependent extension reaction that is primed by a hybridized initial target-specific primer and that uses the target nucleic acid molecule as a template;

(e) amplifying a portion of the target nucleic acid molecule and the tailed random primer sequence with a first tail primer and a first target-specific primer;

(f) amplifying a portion of the amplicon resulting from step (e) with a second tail primer and a second target-specific primer;

(g) sequencing the amplified portion from step (f) using a first and second sequencing primer;

wherein the population of tailed random primers comprises single-stranded oligonucleotide molecules having a 5' nucleic acid sequence identical to a first sequencing primer and a 3' nucleic acid sequence comprising from about 6 to about 12 random nucleotides;

wherein the first target-specific primer comprises a nucleic acid sequence that can specifically anneal to the known target nucleotide sequence of the target nucleic acid at the annealing temperature;

wherein the second target-specific primer comprises a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (c), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer and the second target-specific primer is nested with respect to the first target-specific primer;

wherein the first tail primer comprises a nucleic acid sequence identical to the tailed random primer; and

wherein the second tail primer comprises a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first tail primer.

3. The method of any of paragraphs 1-2, further comprising a step of contacting the sample and products with RNase after extension of the initial target-specific primer.

4. The method of any of paragraphs 1-3, wherein the tailed random primer can form a hair-pin loop structure.

5. The method of any of paragraphs 1-4, wherein the initial target-specific primer and the first target-specific primer are identical.

6. The method of any of paragraphs 1-5, wherein the tailed random primer further comprises a barcode portion comprising 6-12 random nucleotides between the 5' nucleic acid sequence identical to a first sequencing primer and the 3' nucleic acid sequence comprising 6-12 random nucleotides.

7. A method of determining the nucleotide sequence contiguous to a known target nucleotide sequence, the method comprising;

(a) contacting a target nucleic acid molecule comprising the known target nucleotide sequence with a population of tailed random primers under hybridization conditions;

(b) performing a template-dependent extension reaction that is primed by a hybridized tailed random primer and that uses the portion of the target nucleic acid molecule downstream of the site of hybridization as a template;

(c) amplifying a portion of the target nucleic acid molecule and the tailed random primer sequence with a first tail primer and a first target-specific primer;

(d) amplifying a portion of the amplicon resulting from step (c) with a second tail primer and a second target-specific primer;

(e) sequencing the amplified portion from step (d) using a first and second sequencing primer;

wherein the population of tailed random primers comprises single-stranded oligonucleotide molecules having a 5' nucleic acid sequence identical to a first sequencing primer; a middle barcode portion comprising; and a 3' nucleic acid sequence comprising from about 6 to about 12 random nucleotides;

wherein the first target-specific primer comprises a nucleic acid sequence that can specifically anneal to the known target nucleotide sequence of the target nucleic acid at the annealing temperature;

wherein the second target-specific primer comprises a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (c), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer and the second target-specific primer is nested with respect to the first target-specific primer;

wherein the first tail primer comprises a nucleic acid sequence identical to the tailed random primer; and

wherein the second tail primer comprises a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first tail primer.

8. The method of paragraph 7, wherein the each tailed random primer further comprises a spacer nucleic acid sequence between the 5' nucleic acid sequence identical to a first sequencing primer and the 3' nucleic acid sequence comprising about 6 to about 12 random nucleotides.

9. The method of paragraph 7 or 8, wherein the unhybridized primers are removed from the reaction after an extension step.
10. The method of any of paragraphs 7-9, wherein the second tail primer is nested with respect to the first tail primer by at least 3 nucleotides.
11. The method of any of paragraphs 7-10, wherein the first target-specific primer further comprises a 5' tag sequence portion comprising a nucleic acid sequence of high GC content which is not substantially complementary to or substantially identical to any other portion of any of the primers.
12. The method of any of paragraphs 7-11, wherein the second tail primer is identical to the full-length first sequencing primer.
13. The method of any of paragraphs 7-12, wherein the portions of the target-specific primers that specifically anneal to the known target will anneal specifically at a temperature of about 65° C. in a PCR buffer.
14. The method of any of paragraphs 7-13, wherein the sample comprises genomic DNA.
15. The method of any of paragraphs 7-14, wherein the sample comprises RNA and the method further comprises a first step of subjecting the sample to a reverse transcriptase regimen.
16. The method of any of paragraphs 7-15, wherein the nucleic acids present in the sample have not been subjected to shearing or digestion.
17. The method of any of paragraphs 7-16, wherein the sample comprises single-stranded gDNA or cDNA. 18. The method of any of paragraphs 7-17, wherein the reverse transcriptase regimen comprises the use of random hexamers.
19. The method of any of paragraphs 7-18, wherein a gene rearrangement comprises the known target sequence.
20. The method of paragraph 19, wherein the gene rearrangement is present in a nucleic acid selected from the group consisting of: genomic DNA; RNA; and cDNA.
21. The method of any of paragraphs 19-20, wherein the gene rearrangement comprises an oncogene.
22. The method of paragraph 21, wherein the gene rearrangement comprises a fusion oncogene.
23. The method of any of paragraphs 7-22, wherein the nucleic acid product is sequenced by a next-generation sequencing method.
24. The method of paragraph 23, wherein the next-generation sequencing method comprises a method selected from the group consisting of:

Ion Torrent, Illumina, SOLiD, 454; Massively Parallel Signature Sequencing solid-phase, reversible dye-terminator sequencing; and DNA nanoball sequencing.

25. The method of any of paragraphs 7-24, wherein the first and second sequencing primers are compatible with the selected next-generation sequencing method.
26. The method of any of paragraphs 7-25, wherein the method comprises contacting the sample, or separate portions of the sample, with a plurality of sets of first and second target-specific primers.
27. The method of any of paragraphs 7-26, wherein the method comprises contacting a single reaction mixture comprising the sample with a plurality of sets of first and second target-specific primers.
28. The method of any of paragraphs 7-27, wherein the plurality of sets of first and second target-specific primers specifically anneal to known target nucleotide sequences comprised by separate genes.
29. The method of any of paragraphs 7-28, wherein at least two sets of first and second target-specific primers specifically anneal to different portions of a known target nucleotide sequence.
30. The method of any of paragraphs 7-29, wherein at least two sets of first and second target-specific primers specifically anneal to different portions of a single gene comprising a known target nucleotide sequence.
31. The method of any of paragraphs 7-30, wherein at least two sets of first and second target-specific primers specifically anneal to different exons of a gene comprising a known nucleotide target sequence.
32. The method of any of paragraphs 7-31, wherein the plurality of first target-specific primers comprise identical 5' tag sequence portions.
33. The method of any of paragraphs 7-32, wherein each tailed random primer in a population of tailed random primers further comprises an identical sample barcoding portion.
34. The method of paragraph 33, wherein multiple samples are each contacted with a separate population of tailed random primers with a sample barcoding portion; wherein each population of tailed random primers has a distinct sample barcoding portion; and wherein the samples are pooled after step (b).
35. The method of any of paragraphs 7-34, wherein each amplification step comprises a set of cycles of a PCR amplification regimen from 5 cycles to 20 cycles in length.
36. The method of any of paragraphs 7-35, wherein the target-specific primers and the tail primers are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of from about 61 to 72° C.
37. The method of any of paragraphs 7-36, wherein the target-specific primers and the tail primers are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of about 65° C.
38. The method of any of paragraphs 7-37, wherein the target nucleic acid molecule is from a sample, optionally which is a biological sample obtained from a subject.
39. The method of paragraph 38, wherein the sample is obtained from a subject in need of treatment for a disease associated with a genetic alteration.
40. The method of paragraph 39, wherein the disease is cancer.
41. The method of paragraph 38, wherein the sample comprises a population of tumor cells.
42. The method of paragraph 38, wherein the sample is a tumor biopsy.
43. The method paragraph 40, wherein the cancer is lung cancer.
44. The method of any of paragraphs 7-43, wherein a disease-associated gene comprises the known target sequence.
45. The method of 38, wherein a gene rearrangement product in the sample comprises the known target sequence.
46. The method of paragraph 45, wherein the gene rearrangement product is an oncogene.
47. A method of preparing nucleic acids for analysis, the method comprising:
(a) contacting a nucleic acid template comprising a first strand of a target nucleic acid with a complementary target-specific primer that comprises a target-specific

hybridization sequence, under conditions to promote template-specific hybridization and extension of the target-specific primer; and

(b) contacting a nucleic acid template comprising a second strand that is complementary to the first strand of the target nucleic acid with a plurality of different primers that share a common sequence that is 5' to different hybridization sequences, under conditions to promote template-specific hybridization and extension of at least one of the plurality of different primers,

wherein an extension product is generated to contain both a sequence that is characteristic of the target-specific primer and a sequence that is characteristic of the at least one of the plurality of different primers.

48. The method of paragraph 47, wherein the target nucleic acid is a ribonucleic acid.

49. The method of paragraph 47, wherein the target nucleic acid is a deoxyribonucleic acid.

50. The method of any of paragraphs 47 to 49 wherein steps (a) and (b) are performed sequentially.

51. The method of any of paragraphs 47 to 50, wherein the nucleic acid template in step (a) comprises an extension product resulting from the hybridization and extension of the at least one of the plurality of different primers in step (b).

52. The method of any one of paragraphs 47 to 50, wherein the nucleic acid template in step (b) comprises an extension product resulting from the hybridization and extension of the target-specific primer in step (a).

53. The method of paragraph 48, wherein the target nucleic acid is a messenger RNA encoded from a chromosomal segment that comprises a genetic rearrangement.

54. The method of paragraph 49, wherein the target nucleic acid is a chromosomal segment that comprises a portion of a genetic rearrangement.

55. The method of paragraph 8, wherein the genetic rearrangement is an inversion, deletion, or translocation.

56. The method of any one of paragraphs 47 to 55 further comprising amplifying the extension product.

57. The method of any one of paragraphs 47 to 55 further comprising contacting the extension product or amplified extension product with an immobilized oligonucleotide under conditions in which hybridization occurs between the extension product and immobilized oligonucleotide.

58. The method of any preceding paragraph wherein the target nucleic acid comprises a target portion having a known sequence and a flanking portion having an unknown sequence.

59. The method of paragraph 58, wherein different hybridization sequences are complementary to the flanking portion.

60. The method of paragraph 58 or 59, wherein the target-specific hybridization sequence is complementary to the target portion.

61. The method of any of paragraphs 47 to 60, wherein the target-specific primer further comprises, 5' to the target-specific hybridization sequence, at least one of an index sequence, a barcode sequence and an adaptor sequence.

62. The method of any of paragraphs 47 to 60, wherein the common sequence comprises at least one of an index sequence, barcode sequence and an adaptor sequence.

63. The method of any of paragraphs 1-62, wherein the adaptor sequence is a cleavable adaptor sequence for immobilizing oligonucleotides in a flow cell.

Some embodiments of the technology described herein can be defined according to any of the following numbered paragraphs:

1. A method of preparing nucleic acids for analysis, the method comprising:

(a) contacting a nucleic acid template comprising a first strand of a target nucleic acid with a complementary target-specific primer that comprises a target-specific hybridization sequence, under conditions to promote template-specific hybridization and extension of the target-specific primer; and

(b) contacting a nucleic acid template comprising a second strand that is complementary to the first strand of the target nucleic acid with a plurality of different primers that share a common sequence that is 5' to different hybridization sequences, under conditions to promote template-specific hybridization and extension of at least one of the plurality of different primers,

wherein an extension product is generated to contain both a sequence that is characteristic of the target-specific primer and a sequence that is characteristic of the at least one of the plurality of different primers.

2. The method of paragraph 1, wherein the target nucleic acid is a ribonucleic acid.

3. The method of paragraph 1, wherein the target nucleic acid is a deoxyribonucleic acid.

4. The method of any of paragraphs 1 to 3 wherein steps (a) and (b) are performed sequentially.

5. The method of any of paragraphs 1 to 4, wherein the nucleic acid template in step (a) comprises an extension product resulting from the hybridization and extension of the at least one of the plurality of different primers in step (b).

6. The method of any one of paragraphs 1 to 4, wherein the nucleic acid template in step (b) comprises an extension product resulting from the hybridization and extension of the target-specific primer in step (a).

7. The method of paragraph 2, wherein the target nucleic acid is a messenger RNA encoded from a chromosomal segment that comprises a genetic rearrangement.

8. The method of paragraph 3, wherein the target nucleic acid is a chromosomal segment that comprises a portion of a genetic rearrangement.

9. The method of paragraph 8, wherein the genetic rearrangement is an inversion, deletion, or translocation.

10. The method of any one of paragraphs 1 to 9 further comprising amplifying the extension product.

11. The method of any one of paragraphs 1 to 9 further comprising contacting the extension product or amplified extension product with an immobilized oligonucleotide under conditions in which hybridization occurs between the extension product and immobilized oligonucleotide.

12. The method of any preceding paragraph wherein the target nucleic acid comprises a target portion having a known sequence and a flanking portion having an unknown sequence.

13. The method of paragraph 12, wherein different hybridization sequences are complementary to the flanking portion.

14. The method of paragraph 12 or 13, wherein the target-specific hybridization sequence is complementary to the target portion.

15. The method of any of paragraphs 1 to 14, wherein the target-specific primer further comprises, 5' to the target-specific hybridization sequence, at least one of an index sequence, a barcode sequence and an adaptor sequence.
16. The method of any of paragraphs 1 to 14, wherein the common sequence comprises at least one of an index sequence, barcode sequence and an adaptor sequence.
17. The method of paragraph 15 or 16, wherein the adaptor sequence is a cleavable adaptor sequence for immobilizing oligonucleotides in a flow cell.
18. A method of determining the nucleotide sequence contiguous to a known target nucleotide sequence, the method comprising;
(a) contacting a target nucleic acid molecule comprising the known target nucleotide sequence with an initial target-specific primer under hybridization conditions;
(b) performing a template-dependent extension reaction that is primed by a hybridized initial target-specific primer and that uses the target nucleic acid molecule as a template;
(c) contacting the product of step (b) with a population of tailed random primers under hybridization conditions;
(d) performing a template-dependent extension reaction that is primed by a hybridized tailed random primer and that uses the portion of the target nucleic acid molecule downstream of the site of hybridization as a template;
(e) amplifying a portion of the target nucleic acid molecule and the tailed random primer sequence with a first tail primer and a first target-specific primer;
(f) amplifying a portion of the amplicon resulting from step (e) with a second tail primer and a second target-specific primer;
(g) sequencing the amplified portion from step (f) using a first and second sequencing primer;
wherein the population of tailed random primers comprises single-stranded oligonucleotide molecules having a 5' nucleic acid sequence identical to a first sequencing primer and a 3' nucleic acid sequence comprising from about 6 to about 12 random nucleotides;
wherein the first target-specific primer comprises a nucleic acid sequence that can specifically anneal to the known target nucleotide sequence of the target nucleic acid at the annealing temperature:
wherein the second target-specific primer comprises a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (e), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer and the second target-specific primer is nested with respect to the first target-specific primer;
wherein the first tail primer comprises a nucleic acid sequence identical to the tailed random primer; and
wherein the second tail primer comprises a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first tail primer.
19. A method of determining the nucleotide sequence contiguous to a known target nucleotide sequence, the method comprising;
(a) contacting a target nucleic acid molecule comprising the known target nucleotide sequence with a population of tailed random primers under hybridization conditions;
(b) performing a template-dependent extension reaction that is primed by a hybridized tailed random primer and that uses the portion of the target nucleic acid molecule downstream of the site of hybridization as a template;
(c) contacting the product of step (b) with an initial target-specific primer under hybridization conditions;
(d) performing a template-dependent extension reaction that is primed by a hybridized initial target-specific primer and that uses the target nucleic acid molecule as a template;
(e) amplifying a portion of the target nucleic acid molecule and the tailed random primer sequence with a first tail primer and a first target-specific primer;
(f) amplifying a portion of the amplicon resulting from step (e) with a second tail primer and a second target-specific primer;
(g) sequencing the amplified portion from step (f) using a first and second sequencing primer;
wherein the population of tailed random primers comprises single-stranded oligonucleotide molecules having a 5' nucleic acid sequence identical to a first sequencing primer and a 3' nucleic acid sequence comprising from about 6 to about 12 random nucleotides;
wherein the first target-specific primer comprises a nucleic acid sequence that can specifically anneal to the known target nucleotide sequence of the target nucleic acid at the annealing temperature;
wherein the second target-specific primer comprises a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (c), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer and the second target-specific primer is nested with respect to the first target-specific primer;
wherein the first tail primer comprises a nucleic acid sequence identical to the tailed random primer; and
wherein the second tail primer comprises a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first tail primer.
20. The method of any of paragraphs 18-19, further comprising a step of contacting the sample and products with RNase after extension of the initial target-specific primer.
21. The method of any of paragraphs 1-3, wherein the tailed random primer can form a hair-pin loop structure.
22. The method of any of paragraphs 1-4, wherein the initial target-specific primer and the first target-specific primer are identical.
23. The method of any of paragraphs 1-5, wherein the tailed random primer further comprises a barcode portion comprising 6-12 random nucleotides between the 5' nucleic acid sequence identical to a first sequencing primer and the 3' nucleic acid sequence comprising 6-12 random nucleotides. 7. A method of determining the nucleotide sequence contiguous to a known target nucleotide sequence, the method comprising;

(a) contacting a target nucleic acid molecule comprising the known target nucleotide sequence with a population of tailed random primers under hybridization conditions;

(b) performing a template-dependent extension reaction that is primed by a hybridized tailed random primer and that uses the portion of the target nucleic acid molecule downstream of the site of hybridization as a template;

(c) amplifying a portion of the target nucleic acid molecule and the tailed random primer sequence with a first tail primer and a first target-specific primer;

(d) amplifying a portion of the amplicon resulting from step (c) with a second tail primer and a second target-specific primer;

(e) sequencing the amplified portion from step (d) using a first and second sequencing primer;

wherein the population of tailed random primers comprises single-stranded oligonucleotide molecules having a 5' nucleic acid sequence identical to a first sequencing primer; a middle barcode portion comprising; and a 3' nucleic acid sequence comprising from about 6 to about 12 random nucleotides;

wherein the first target-specific primer comprises a nucleic acid sequence that can specifically anneal to the known target nucleotide sequence of the target nucleic acid at the annealing temperature;

wherein the second target-specific primer comprises a 3' portion comprising a nucleic acid sequence that can specifically anneal to a portion of the known target nucleotide sequence comprised by the amplicon resulting from step (c), and a 5' portion comprising a nucleic acid sequence that is identical to a second sequencing primer and the second target-specific primer is nested with respect to the first target-specific primer;

wherein the first tail primer comprises a nucleic acid sequence identical to the tailed random primer; and

wherein the second tail primer comprises a nucleic acid sequence identical to a portion of the first sequencing primer and is nested with respect to the first tail primer.

24. The method of paragraph 23, wherein the each tailed random primer further comprises a spacer nucleic acid sequence between the 5' nucleic acid sequence identical to a first sequencing primer and the 3' nucleic acid sequence comprising about 6 to about 12 random nucleotides.

25. The method of paragraph 23 or 24, wherein the unhybridized primers are removed from the reaction after an extension step.

26. The method of any of paragraphs 23-25, wherein the second tail primer is nested with respect to the first tail primer by at least 3 nucleotides.

27. The method of any of paragraphs 23-26, wherein the first target-specific primer further comprises a 5' tag sequence portion comprising a nucleic acid sequence of high GC content which is not substantially complementary to or substantially identical to any other portion of any of the primers.

28. The method of any of paragraphs 23-27, wherein the second tail primer is identical to the full-length first sequencing primer.

29. The method of any of paragraphs 23-28, wherein the portions of the target-specific primers that specifically anneal to the known target will anneal specifically at a temperature of about 65° C. in a PCR buffer.

30. The method of any of paragraphs 23-29, wherein the sample comprises genomic DNA.

31. The method of any of paragraphs 23-30, wherein the sample comprises RNA and the method further comprises a first step of subjecting the sample to a reverse transcriptase regimen.

32. The method of any of paragraphs 23-31, wherein the nucleic acids present in the sample have not been subjected to shearing or digestion or wherein the sample comprises single-stranded gDNA or cDNA.

33. The method of any of paragraphs 23-32, wherein the reverse transcriptase regimen comprises the use of random hexamers.

34. The method of any of paragraphs 23-33, wherein a gene rearrangement comprises the known target sequence.

35. The method of paragraph 34, wherein the gene rearrangement is present in a nucleic acid selected from the group consisting of: genomic DNA; RNA; and cDNA.

36. The method of any of paragraphs 34-35, wherein the gene rearrangement comprises an oncogene.

37. The method of paragraph 36, wherein the gene rearrangement comprises a fusion oncogene.

38. The method of any of paragraphs 23-37, wherein the nucleic acid product is sequenced by a next-generation sequencing method.

39. The method of paragraph 38, wherein the next-generation sequencing method comprises a method selected from the group consisting of:

Ion Torrent, Illumina, SOLiD, 454; Massively Parallel Signature Sequencing solid-phase, reversible dye-terminator sequencing; and DNA nanoball sequencing.

40. The method of any of paragraphs 23-39, wherein the first and second sequencing primers are compatible with the selected next-generation sequencing method.

41. The method of any of paragraphs 23-40, wherein the method comprises contacting the sample, or separate portions of the sample, with a plurality of sets of first and second target-specific primers.

42. The method of any of paragraphs 23-41, wherein the method comprises contacting a single reaction mixture comprising the sample with a plurality of sets of first and second target-specific primers.

43. The method of any of paragraphs 23-42, wherein the plurality of sets of first and second target-specific primers specifically anneal to known target nucleotide sequences comprised by separate genes.

44. The method of any of paragraphs 23-43, wherein at least two sets of first and second target-specific primers specifically anneal to different portions of a known target nucleotide sequence.

45. The method of any of paragraphs 23-44, wherein at least two sets of first and second target-specific primers specifically anneal to different portions of a single gene comprising a known target nucleotide sequence.

46. The method of any of paragraphs 23-45, wherein at least two sets of first and second target-specific primers specifically anneal to different exons of a gene comprising a known nucleotide target sequence.

47. The method of any of paragraphs 23-46, wherein the plurality of first target-specific primers comprise identical 5' tag sequence portions.

48. The method of any of paragraphs 23-47, wherein each tailed random primer in a population of tailed random primers further comprises an identical sample barcoding portion.

49. The method of paragraph 48, wherein multiple samples are each contacted with a separate population of tailed random primers with a sample barcoding portion; wherein each population of tailed random primers has a distinct sample barcoding portion; and wherein the samples are pooled after step (b).
50. The method of any of paragraphs 23-49, wherein each amplification step comprises a set of cycles of a PCR amplification regimen from 5 cycles to 20 cycles in length.
51. The method of any of paragraphs 23-50, wherein the target-specific primers and the tail primers are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of from about 61 to 72° C.
52. The method of any of paragraphs 23-51, wherein the target-specific primers and the tail primers are designed such that they will specifically anneal to their complementary sequences at an annealing temperature of about 65° C.
53. The method of any of paragraphs 23-52, wherein the target nucleic acid molecule is from a sample, optionally which is a biological sample obtained from a subject.
54. The method of paragraph 53, wherein the sample is obtained from a subject in need of treatment for a disease associated with a genetic alteration.
55. The method of paragraph 54, wherein the disease is cancer.
56. The method of paragraph 53, wherein the sample comprises a population of tumor cells.
57. The method of paragraph 53, wherein the sample is a tumor biopsy.
58. The method paragraph 55, wherein the cancer is lung cancer.
59. The method of any of paragraphs 23-58, wherein a disease-associated gene comprises the known target sequence.
60. The method of 53, wherein a gene rearrangement product in the sample comprises the known target sequence.
61. The method of paragraph 60, wherein the gene rearrangement product is an oncogene.

EXAMPLES

Example 1: A Method Using Reverse Transcriptase with Tailed Random Oligonucleotides and Gene-Specific Oligonucleotides to Amplify 3'Fusion Events

First Strand Synthesis

As a first step towards amplifying 3' fusion events for sequence analysis, RNA was obtained from a sample isolated from a subject. The following reaction was assembled on ice to synthesize the first cDNA strand:

12 μL purified RNA and H2O
2 μL 1.2 μg/μL random primer #1 (9mer)
2 μL dNTP

The reaction was transferred to a thermocycler and incubated at 65° C. for 5 minutes. Then, the reaction was centrifuged and incubated on ice for at least one minute.

To the reaction above, the following was prepared on ice:

2 μL 10× M-MuLV reverse transcriptase buffer
1 μL 40 U/μL RNase inhibitor
1 μL 200 U/μL M-MuLV enzyme

The reaction was mixed and centrifuged briefly to collect the reaction contents at the bottom of the tube, then placed back on ice. The reaction was then incubated at 42° C. for 60 minutes, followed by 4° C.

Exol Treatment

The following was added to the reaction above,

1 μL 20 U/μL Exonuclease I

The reaction was mixed and centrifuged briefly to collect the contents at the bottom of the tube, then incubated at 37° C. for 10 minutes. Next, 1.28 μL of 1N NaOH was added and mixed by pipetting up and down then centrifuged to collect the contents. The reaction was incubated at 80° C. for 10 minutes, then 4 μL 10 mM Tris pH 8.3 was added ad mixed by pipetting up and down. 20 μL of the Exonuclease-treated DNA solution was transferred to a fresh 200 μL PCR tube on ice.

Second Strand cDNA Synthesis

The following reaction was prepared:

20 μL DNA solution, from above
11 μL nuclease-free H_2O
4 μL 10×PCR Buffer II
4 μL 3 μM gene-specific primer #1
1 μL 0.5 mM dNTP

The reaction was mixed by pipetting up and down then centrifuged briefly to collect the contents and placed on ice. The reaction was then incubated at 95° C. for 3 minutes, then 22° C. for 10 seconds followed by 4° C. until proceeding to the next step. The reaction was incubated on ice for at least one minutes.

The following was added to the reaction:

1 μL 400 U/μL Manta 1.0 DNA Polymerase (high concentration)

The reaction was incubated at 25° C. for 10 seconds, then 70° C. for 10 minutes and maintained at 4° C. until proceeding to the next step.

DNA Purification with AMPure Beads #1

The following was added to the reaction above,

88.4 μL AMPure beads

The suspension was mixed well and incubated for 5 minutes at room temperature. A magnet was used 2-4 minutes to collect the beads and the solution appeared clear. The supernatant was discarded and the beads were washed twice times with 200 μL 70% ethanol on the magnet. After the second wash, the beads were dried at room temperature for 5 minutes. Finally, the DNA was eluted by removing the tubes from the magnet and resuspending the beads in 12 μL 10 mM Tris-HCl pH 8.3 elution buffer included in the AMPure kit. The RNA-bead solution was placed on the magnet for 2 minutes. Then, the DNA solution was transferred to a fresh PCR tube, being sure to avoid transferring beads to the fresh tube.

It should be appreciated that in some embodiments the ratio of beads to reaction mix can effect the size of fragments returned. In some embodiments, for fusion detection, all or substantially all fragments are longer than 60 nt (e.g., 30 nt on either side of a fusion break point or junction) so each gene can be easily identified.

Amplification #1

The following reaction was prepared:

10 μL purified DNA, from Purification #1 above
4 μL 5×Phoenix Hot Start Buffer
2 μL 2 mM dNTP
2 μL 10 μM P5_barcode primer
2 μL 3 μM gene-specific primer #1
0.5 to 2 Units polymerase (e.g., Pheonix Hot Start Taq, VeraSeq)

The reaction was incubated as follows:

Step 1: 95° C. for 3 minutes

Step 2: 95° C. for 30 seconds

Step 3: 65° C. for 5 minutes, return to step 2 for 14 cycles

Step 4: 72° C. for 2 minutes

Step 5: 4° C., until proceeding with the protocol

DNA Purification with AMPure Beads #2

The following was added to the reaction above,

36.4 μL AMPure beads

The suspension was mixed well and incubated for 5 minutes at room temperature. A magnet was used 2-4 minutes to collect the beads and the solution appeared clear. The supernatant was discarded and the beads were washed twice times with 200 μL 70% ethanol on the magnet. After the second wash, the beads were dried at room temperature for 5 minutes. Finally, the DNA was eluted by removing the tubes from the magnet and resuspending the beads in 9 μL 10 mM Tris-HCl pH 8.3 elution buffer included in the AMPure kit. The RNA-bead solution was placed on the magnet for 2 minutes. Then, the DNA solution was transferred to a fresh PCR tube, being sure to avoid transferring beads to the fresh tube.

Amplification #2

The following reaction was prepared:

8.5 μL purified DNA, from Purification #2 above

4 μL 5×Phoenix Hot Start Buffer

2 μL 2 mM dNTP

2 μL 10 μM P5_29 bp primer

2 μL 10 μM P7 barcode primer

2 μL 3 μM gene-specific primer #2

0.2 μL 5 U/μL Phoenix Hot Start Taq polymerase

2 μL 10 μM P7 barcode primer

The reaction was incubated as follows:

Step 1: 95° C. for 3 minutes

Step 2: 95° C. for 30 seconds

Step 3: 65° C. for 5 minutes, return to step 2 for 14 cycles

Step 4: 72° C. for 2 minutes

Step 5: 4° C., until proceeding with the protocol

DNA Purification with AMPure Beads #3

The following was added to the reaction above,

37.3 μL AMPure beads

The suspension was mixed well and incubated for 5 minutes at room temperature. A magnet was used 2-4 minutes to collect the beads and the solution appeared clear. The supernatant was discarded and the beads were washed twice times with 200 μL 70% ethanol on the magnet. After the second wash, the beads were dried at room temperature for 5 minutes. Finally, the DNA was eluted by removing the tubes from the magnet and resuspending the beads in 20 μL 10 mM Tris-HCl pH 8.3 elution buffer included in the AMPure kit. The RNA-bead solution was placed on the magnet for 2 minutes. Then, the DNA solution was transferred to a fresh PCR tube, being sure to avoid transferring beads to the fresh tube.

Quantification of Library Concentration

The Kapa Biosystems qPCR kit for Illumina was used to quantitate the concentration of each library prepared using the protocol above. The barcoded libraries were pooled at equimolar concentrations. Then the library was loaded on an Illumina MiSeq at XpM using the MiSeq v2 300 cycle reagent kit following the manufacturer's instruction. The samples were sequenced using 2×150 bp reads with 7 base encoded index reads.

Example 2: A Method Using Reverse Transcriptase with Gene-Specific Oligonucleotides and Tailed Random Oligonucleotides to Amplify 5'Fusion Events

First Strand Synthesis

As a first step towards amplifying 5' fusion events for sequence analysis, RNA was obtained from a sample isolated from a subject. The following reaction was assembled on ice to synthesize the first cDNA strand:

12 μL purified RNA and H_2O

2 μL gene-specific primer #1

2 μL dNTP

The reaction was transferred to a thermocycler and incubated at 65° C. for 5 minutes. Then, the reaction was centrifuged and incubated on ice for at least one minute.

To the reaction above, the following was prepared on ice:

2 μL 10×M-MuLV reverse transcriptase buffer

1 μL 40 U/μL RNase inhibitor

1 μL 200 U/μL M-MuLV enzyme

The reaction was mixed and centrifuged briefly to collect the reaction contents at the bottom of the tube, then placed back on ice. The reaction was then incubated at 42° C. for 60 minutes, followed by 4° C.

Exol Treatment

The following was added to the reaction above,

1 μL 20 U/μL Exonuclease I

The reaction was mixed and centrifuged briefly to collect the contents at the bottom of the tube, then incubated at 37° C. for 10 minutes. Next, 1.28 μL of 1N NaOH was added and mixed by pipetting up and down then centrifuged to collect the contents. The reaction was incubated at 80° C. for 10 minutes, then 4 μL 10 mM Tris pH 8.3 was added ad mixed by pipetting up and down. 20 μL of the Exonuclease-treated DNA solution was transferred to a fresh 200 μL PCR tube on ice.

Second Strand cDNA Synthesis

The following reaction was prepared:

20 μL DNA solution, from above

11 μL nuclease-free H_2O

4 μL 10×PCR Buffer II

4 μL 1.2 μg/μL random primer (9mer)

1 μL 0.5 mM dNTP

The reaction was mixed by pipetting up and down then centrifuged briefly to collect the contents and placed on ice. The reaction was then incubated at 95° C. for 3 minutes, then 22° C. for 10 seconds followed by 4° C. until proceeding to the next step. The reaction was incubated on ice for at least one minutes.

The following was added to the reaction:

1 μL 400 U/μL Manta 1.0 DNA Polymerase (high concentration)

The reaction was incubated at 25° C. for 10 seconds, then 70° C. for 10 minutes and maintained at 4° C. until proceeding to the next step.

DNA Purification with AMPure Beads #1

The following was added to the reaction above,

88.4 μL AMPure beads

The suspension was mixed well and incubated for 5 minutes at room temperature. A magnet was used 2-4 minutes to collect the beads and the solution appeared clear. The supernatant was discarded and the beads were washed twice times with 200 μL 70% ethanol on the magnet. After the second wash, the beads were dried at room temperature for 5 minutes. Finally, the DNA was eluted by removing the tubes from the magnet and resuspending the beads in 12 μL 10 mM Tris-HCl pH 8.3 elution buffer included in the AMPure kit. The RNA-bead solution was placed on the magnet for 2 minutes. Then, the DNA solution was transferred to a fresh PCR tube, being sure to avoid transferring beads to the fresh tube.

Amplification #1

The following reaction was prepared:

10 μL purified DNA, from Purification #1 above

4 μL 5× Phoenix Hot Start Buffer

2 μL 2 mM dNTP

2 μL 10 μM P5_barcode primer

2 μL 3 μM gene-specific primer #1

0.5 to 2 Units polymerase (e.g., Pheonix Hot Start Taq, VeraSeq)

The reaction was incubated as follows:

Step 1: 95° C. for 3 minutes

Step 2: 95° C. for 30 seconds

Step 3: 65° C. for 5 minutes, return to step 2 for 14 cycles

Step 4: 72° C. for 2 minutes

Step 5: 4° C., until proceeding with the protocol

DNA Purification with AMPure beads #2

The following was added to the reaction above,

36.4 μL AMPure beads

The suspension was mixed well and incubated for 5 minutes at room temperature. A magnet was used 2-4 minutes to collect the beads and the solution appeared clear. The supernatant was discarded and the beads were washed twice times with 200 μL 70% ethanol on the magnet. After the second wash, the beads were dried at room temperature for 5 minutes. Finally, the DNA was eluted by removing the tubes from the magnet and resuspending the beads in 9 μL 10 mM Tris-HCl pH 8.3 elution buffer included in the AMPure kit. The RNA-bead solution was placed on the magnet for 2 minutes. Then, the DNA solution was transferred to a fresh PCR tube, being sure to avoid transferring beads to the fresh tube.

Amplification #2

The following reaction was prepared:

8.5 μL purified DNA, from Purification #2 above

4 μL 5× Phoenix Hot Start Buffer

2 μL 2 mM dNTP

2 μL 10 μM P5_29 bp primer

2 μL 10 μM P7 barcode primer

2 μL 3 μM gene-specific primer #2

0.2 μL 5 U/μL Phoenix Hot Start Taq polymerase

2 μL 10 μM P7 barcode primer

The reaction was incubated as follows:

Step 1: 95° C. for 3 minutes

Step 2: 95° C. for 30 seconds

Step 3: 65° C. for 5 minutes, return to step 2 for 14 cycles

Step 4: 72° C. for 2 minutes

Step 5: 4° C., until proceeding with the protocol

DNA Purification with AMPure Beads #3

The following was added to the reaction above,

37.3 μL AMPure beads

The suspension was mixed well and incubated for 5 minutes at room temperature. A magnet was used 2-4 minutes to collect the beads and the solution appeared clear. The supernatant was discarded and the beads were washed twice times with 200 μL 70% ethanol on the magnet. After the second wash, the beads were dried at room temperature for 5 minutes. Finally, the DNA was eluted by removing the tubes from the magnet and resuspending the beads in 20 μL 10 mM Tris-HCl pH 8.3 elution buffer included in the AMPure kit. The RNA-bead solution was placed on the magnet for 2 minutes. Then, the DNA solution was transferred to a fresh PCR tube, being sure to avoid transferring beads to the fresh tube.

Quantification of Library Concentration

The Kapa Biosystems qPCR kit for Illumina was used to quantitate the concentration of each library prepared using the protocol above. The barcoded libraries were pooled at equimolar concentrations. Then the library was loaded on an Illumina MiSeq at XpM using the MiSeq v2 300 cycle reagent kit following the manufacturer's instruction. The samples were sequenced using 2×150 bp reads with 7 base encoded index reads.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 5

<210> SEQ ID NO 1
<211> LENGTH: 7368
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

caagctttca agcattcaaa ggtctaaatg aaaaaggcta agtattattt caaaaggcaa 60

gtatatccta atatagcaaa acaaacaaag caaaatccat cagctactcc tccaattgaa 120

gtgatgaagc ccaaataatt catatagcaa aatggagaaa attagaccgg ccatctaaaa 180

atctgccatt ggtgaagtga tgaagaacat ttactgtctt attccgaagc ttgtcaattt 240

tgcaactctt ggctgcctat ggatttctgt ggtgcagtgt acagttttaa atagctgcct 300

aaagtcgtgt gtaactaatc tgggccagca gcttgacctt ggcacaccac ataatctgag 360

tgaaccgtgt atccaaggat gtcacttttg gaactctgta gatcagaaaa actgtgcttt 420

aaagtgtcgg gagtcgtgtg aggttggctg tagcagcgcg gaaggtgcat atgaagagga 480

agtactggaa aatgcagacc taccaactgc tccctttgct tcttccattg gaagccacaa 540

tatgacatta cgatggaaat ctgcaaactt ctctggagta aaatacatca ttcagtggaa 600

atatgcacaa cttctgggaa gctggactta tactaagact gtgtccagac cgtcctatgt 660

ggtcaagccc ctgcaccctt tcactgagta cattttccga gtggtttgga tcttcacagc 720

-continued

gcagctgcag ctctactccc ctccaagtcc cagttacagg actcatcctc atggagttcc 780

tgaaactgca cctttgatta ggaatattga gagctcaagt cccgacactg tggaagtcag 840

ctgggatcca cctcaattcc caggtggacc tattttgggt tataacttaa ggctgatcag 900

caaaaatcaa aaattagatg cagggacaca gagaaccagt ttccagtttt actccacttt 960

accaaatact atctacaggt tttctattgc agcagtaaat gaagttggtg agggtccaga 1020

agcagaatct agtattacca cttcatcttc agcagttcaa caagaggaac agtggctctt 1080

tttatccaga aaaacttctc taagaaagag atctttaaaa catttagtag atgaagcaca 1140

ttgccttcgg ttggatgcta tataccataa tattacagga atatctgttg atgtccacca 1200

gcaaattgtt tatttctctg aaggaactct catatgggcg aagaaggctg ccaacatgtc 1260

tgatgtatct gacctgagaa ttttttacag aggttcagga ttaatttctt ctatctccat 1320

agattggctt tatcaaagaa tgtatttcat catggatgaa ctggtatgtg tctgtgattt 1380

agagaactgc tcaaacatcg aggaaattac tccaccctct attagtgcac ctcaaaaaat 1440

tgtggctgat tcatacaatg ggtatgtctt ttacctcctg agagatggca tttatagagc 1500

agaccttcct gtaccatctg gccggtgtgc agaagctgtg cgtattgtgg agagttgcac 1560

gttaaaggac tttgcaatca agccacaagc caagcgaatc atttacttca atgacactgc 1620

ccaagtcttc atgtcaacat ttctggatgg ctctgcttcc catctcatcc tacctcgcat 1680

ccccttttgct gatgtgaaaa gttttgcttg tgaaaacaat gactttcttg tcacagatgg 1740

caaggtcatt ttccaacagg atgctttgtc ttttaatgaa ttcatcgtgg gatgtgacct 1800

gagtcacata gaagaatttg ggtttggtaa cttggtcatc tttggctcat cctcccagct 1860

gcaccctctg ccaggccgcc cgcaggagct ttcggtgctg tttggctctc accaggctct 1920

tgttcaatgg aagcctcctg cccttgccat aggagccaat gtcatcctga tcagtgatat 1980

tattgaactc tttgaattag gcccttctgc ctggcagaac tggacctatg aggtgaaagt 2040

atccacccaa gaccctcctg aagtcactca tattttcttg aacataagtg gaaccatgct 2100

gaatgtacct gagctgcaga gtgctatgaa atacaaggtt tctgtgagag caagttctcc 2160

aaagaggcca ggcccctggt cagagccctc agtgggtact accctggtgc cagctagtga 2220

accaccattt atcatggctg tgaaagaaga tgggctttgg agtaaaccat taaatagctt 2280

tggcccagga gagttcttat cctctgatat aggaaatgtg tcagacatgg attggtataa 2340

caacagcctc tactacagtg acacgaaagg cgacgttttt gtgtggctgc tgaatgggac 2400

ggatatctca gagaattatc acctacccag cattgcagga gcaggggctt tagcttttga 2460

gtggctgggt cactttctct actgggctgg aaagacatat gtgatacaaa ggcagtctgt 2520

gttgacggga cacacagaca ttgttaccca cgtgaagcta ttggtgaatg acatggtggt 2580

ggattcagtt ggtggatatc tctactggac cacactctat tcagtggaaa gcaccagact 2640

aaatgggaa agttcccttg tactacagac acagccttgg ttttctggga aaaaggtaat 2700

tgctctaact ttagacctca gtgatgggct cctgtattgg ttggttcaag acagtcaatg 2760

tattcacctg tacacagctg ttcttcgggg acagagcact ggggatacca ccatcacaga 2820

atttgcagcc tggagtactt ctgaaatttc ccagaatgca ctgatgtact atagtggtcg 2880

gctgttctgg atcaatggct ttaggattat cacaactcaa gaaataggtc agaaaaccag 2940

tgtctctgtt ttggaaccag ccagatttaa tcagttcaca attattcaga catcccttaa 3000

gcccctgcca gggaactttt cctttacccc taaggttatt ccagattctg ttcaagagtc 3060

JX-003.37

Appx165

-continued

ttcatttagg attgaaggaa atgcttcaag ttttcaaatc ctgtggaatg gtccccctgc 3120

ggtagactgg ggtgtagttt tctacagtgt agaatttagt gctcattcta agtttcttggc 3180

tagtgaacaa cactctttac ctgtatttac tgtggaagga ctggaacctt atgccttatt 3240

taatctttct gtcactcctt atacctactg gggaaagggc cccaaaacat ctctgtcact 3300

tcgagcacct gaaacagttc catcagcacc agagaacccc agaatattta tattaccaag 3360

tggaaaatgc tgcaacaaga atgaagttgt ggtggaattt aggtggaaca aacctaagca 3420

tgaaaatggg gtgttaacaa aatttgaaat tttctacaat atatccaatc aaagtattac 3480

aaacaaaaca tgtgaagact ggattgctgt caatgtcact ccctcagtga tgtcttttca 3540

acttgaaggc atgagtccca gatgctttat tgccttccag gttagggcct ttacatctaa 3600

ggggccagga ccatatgctg acgttgtaaa gtctacaaca tcagaaatca acccatttcc 3660

tcacctcata actcttcttg gtaacaagat agtttttta gatatggatc aaaatcaagt 3720

tgtgtggacg ttttcagcag aaagagttat cagtgccgtt tgctacacag ctgataatga 3780

gatgggatat tatgctgaag gggactcact ctttcttctg cacttgcaca atcgctctag 3840

ctctgagctt ttccaagatt cactggtttt tgatatcaca gttattacaa ttgactggat 3900

ttcaaggcac ctctactttg cactgaaaga atcacaaaat ggaatgcaag tatttgatgt 3960

tgatcttgaa cacaaggtga aatatcccag agaggtgaag attcacaata ggaattcaac 4020

aataatttct ttttctgtat atcctctttt aagtcgcttg tattggacag aagtttccaa 4080

ttttggctac cagatgttct actacagtat tatcagtcac accttgcacc gaattctgca 4140

acccacagct acaaaccaac aaaacaaaag gaatcaatgt tcttgtaatg tgactgaatt 4200

tgagttaagt ggagcaatgg ctattgatac ctctaaccta gagaaaccat tgatatactt 4260

tgccaaagca caagagatct gggcaatgga tctggaaggc tgtcagtgtt ggagagttat 4320

cacagtacct gctatgctcg caggaaaaac ccttgttagc ttaactgtgg atggagatct 4380

tatatactgg atcatcacag caaaggacag cacacagatt tatcaggcaa agaaaggaaa 4440

tggggccatc gtttcccagg tgaaggccct aaggagtagg catatcttgg cttacagttc 4500

agttatgcag ccttttccag ataaagcgtt tctgtctcta gcttcagaca ctgtggaacc 4560

aactatactt aatgccacta acactagcct cacaatcaga ttacctctgg ccaagacaaa 4620

cctcacatgg tatggcatca ccagccctac tccaacatac ctggtttatt atgcagaagt 4680

taatgacagg aaaaacagct ctgacttgaa atatagaatt ctggaatttc aggacagtat 4740

agctcttatt gaagatttac aaccattttc aacatacatg atacagatag ctgtaaaaaa 4800

ttattattca gatcctttgg aacatttacc accaggaaaa gagatttggg gaaaaactaa 4860

aaatggagta ccagaggcag tgcagctcat taatacaact gtgcggtcag acaccagcct 4920

cattatatct tggagagaat ctcacaagcc aaatggacct aaagaatcag tccgttatca 4980

gttggcaatc tcacacctgg ccctaattcc tgaaactcct ctaagacaaa gtgaatttcc 5040

aaatggaagg ctcactctcc ttgttactag actgtctggt ggaaatattt atgtgttaaa 5100

ggttcttgcc tgccactctg aggaaatgtg gtgtacagag agtcatcctg tcactgtgga 5160

aatgtttaac acaccagaga aaccttattc cttggttcca gagaacacta gtttgcaatt 5220

taattggaag gctccattga atgttaacct catcagattt tgggttgagc tacagaagtg 5280

gaaatacaat gagttttacc atgttaaaac ttcatgcagc caaggtcctg cttatgtctg 5340

taatatcaca aatctacaac cttatacttc atataatgtc agagtagtgg tggtttataa 5400

gacgggagaa aatagcacct cacttccaga aagctttaag acaaaagctg gagtcccaaa 5460

-continued

taaaccaggc attcccaaat tactagaagg gagtaaaaat tcaatacagt gggagaaagc 5520

tgaagataat ggatgtagaa ttacatacta tatccttgag ataagaaaga gcacttcaaa 5580

taatttacag aaccagaatt taaggtggaa gatgacattt aatggatcct gcagtagtgt 5640

ttgcacatgg aagtccaaaa acctgaaagg aatatttcag ttcagagtag tagctgcaaa 5700

taatctaggg tttggtgaat atagtggaat cagtgagaat attatattag ttggagatga 5760

tttttggata ccagaaacaa gtttcatact tactattata gttggaatat ttctggttgt 5820

tacaatccca ctgacctttg tctggcatag aagattaaag aatcaaaaaa gtgccaagga 5880

aggggtgaca gtgcttataa acgaagacaa agagttggct gagctgcgag gtctggcagc 5940

cggagtaggc ctggctaatg cctgctatgc aatacatact cttccaaccc aagaggagat 6000

tgaaaatctt cctgccttcc ctcgggaaaa actgactctg cgtctcttgc tgggaagtgg 6060

agcctttgga gaagtgtatg aaggaacagc agtggacatc ttaggagttg gaagtggaga 6120

aatcaaagta gcagtgaaga ctttgaagaa gggttccaca gaccaggaga agattgaatt 6180

cctgaaggag gcacatctga tgagcaaatt taatcatccc aacattctga agcagcttgg 6240

agtttgtctg ctgaatgaac cccaatacat tatcctggaa ctgatggagg gaggagacct 6300

tcttacttat ttgcgtaaag cccggatggc aacgttttat ggtcctttac tcaccttggt 6360

tgaccttgta gacctgtgtg tagatatttc aaaaggctgt gtctacttgg aacggatgca 6420

tttcattcac agggatctgg cagctagaaa ttgccttgtt tccgtgaaag actataccag 6480

tccacggata gtgaagattg gagactttgg actcgccaga gacatctata aaaatgatta 6540

ctatagaaag agaggggaag gcctgctccc agttcggtgg atggctccag aaagtttgat 6600

ggatggaatc ttcactactc aatctgatgt atggtctttt ggaattctga tttgggagat 6660

tttaactctt ggtcatcagc cttatccagc tcattccaac cttgatgtgt taaactatgt 6720

gcaaacagga gggagactgg agccaccaag aaattgtcct gatgatctgt ggaatttaat 6780

gacccagtgc tgggctcaag aacccgacca aagacctact tttcatagaa ttcaggacca 6840

acttcagtta ttcagaaatt ttttcttaaa tagcatttat aagtccagag atgaagcaaa 6900

caacagtgga gtcataaatg aaagctttga aggtgaagat ggcgatgtga tttgtttgaa 6960

ttcagatgac attatgccag ttgctttaat ggaaacgaag aaccgagaag ggttaaacta 7020

tatggtactt gctacagaat gtggccaagg tgaagaaaag tctgagggtc ctctaggctc 7080

ccaggaatct gaatcttgtg gtctgaggaa agaagagaag gaaccacatg cagacaaaga 7140

tttctgccaa gaaaaacaag tggcttactg cccttctggc aagcctgaag gcctgaacta 7200

tgcctgtctc actcacagtg gatatggaga tgggtctgat taatagcgtt gtttgggaaa 7260

tagagagttg agataaacac tctcattcag tagttactga aagaaaactc tgctagaatg 7320

ataaatgtca tggtggtcta taactccaaa taaacaatgc aacgttcc 7368

<210> SEQ ID NO 2
<211> LENGTH: 6267
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2

agctgcaagt ggcgggcgcc caggcagatg cgatccagcg gctctggggg cggcagcggt 60

ggtagcagct ggtacctccc gccgcctctg ttcggagggt cgcggggcac cgaggtgctt 120

tccggccgcc ctctggtcgg ccacccaaag ccgcgggcgc tgatgatggg tgaggagggg 180

-continued

gcggcaagat ttcgggcgcc cctgccctga acgccctcag ctgctgccgc cggggccgct 240

ccagtgcctg cgaactctga ggagccgagg cgccggtgag agcaaggacg ctgcaaactt 300

gcgcagcgcg ggggctggga ttcacgccca gaagttcagc aggcagacag tccgaagcct 360

tcccgcagcg gagagatagc ttgagggtgc gcaagacggc agcctccgcc ctcggttccc 420

gcccagaccg ggcagaagag cttggaggag ccaaaaggaa cgcaaaaggc ggccaggaca 480

gcgtgcagca gctgggagcc gccgttctca gccttaaaag ttgcagagat tggaggctgc 540

cccgagaggg gacagacccc agctccgact gcgggggcca ggagaggacg gtacccaact 600

gccacctccc ttcaaccata gtagttcctc tgtaccgagc gcagcgagct acagacgggg 660

gcgcggcact cggcgcggag agcgggaggc tcaaggtccc agccagtgag cccagtgtgc 720

ttgagtgtct ctggactcgc ccctgagctt ccaggtctgt ttcatttaga ctcctgctcg 780

cctccgtgca gttgggggaa agcaagagac ttgcgcgcac gcacagtcct ctggagatca 840

ggtggaagga gccgctgggt accaaggact gttcagagcc tcttcccatc tcggggagag 900

cgaagggtga ggctgggccc ggagagcagt gtaaacggcc tcctccggcg ggatgggagc 960

catcgggctc ctgtggctcc tgccgctgct gctttccacg gcagctgtgg gctccgggat 1020

ggggaccggc cagcgcgcgg gctccccagc tgcggggccg ccgctgcagc cccgggagcc 1080

actcagctac tcgcgcctgc agaggaagag tctggcagtt gacttcgtgg tgccctcgct 1140

cttccgtgtc tacgcccggg acctactgct gccaccatcc tcctcggagc tgaaggctgg 1200

caggcccgag gcccgcggct cgctagctct ggactgcgcc ccgctgctca ggttgctggg 1260

gccggcgccg ggggtctcct ggaccgccgg ttcaccagcc ccggcagagg cccggacgct 1320

gtccagggtg ctgaagggcg gctccgtgcg caagctccgg cgtgccaagc agttggtgct 1380

ggagctgggc gaggaggcga tcttggaggg ttgcgtcggg ccccccgggg aggcggctgt 1440

ggggctgctc cagttcaatc tcagcgagct gttcagttgg tggattcgcc aaggcgaagg 1500

gcgactgagg atccgcctga tgcccgagaa gaaggcgtcg gaagtgggca gagagggaag 1560

gctgtccgcg gcaattcgcg cctcccagcc ccgccttctc ttccagatct tcgggactgg 1620

tcatagctcc ttggaatcac caacaaacat gccttctcct tctcctgatt attttacatg 1680

gaatctcacc tggataatga aagactcctt ccctttcctg tctcatcgca gccgatatgg 1740

tctggagtgc agctttgact tcccctgtga gctggagtat tcccctccac tgcatgacct 1800

caggaaccag agctggtcct ggcgccgcat cccctccgag gaggcctccc agatggactt 1860

gctggatggg cctggggcag agcgttctaa ggagatgccc agaggctcct ttctccttct 1920

caacacctca gctgactcca agcacaccat cctgagtccg tggatgagga gcagcagtga 1980

gcactgcaca ctggccgtct cggtgcacag gcacctgcag ccctctggaa ggtacattgc 2040

ccagctgctg ccccacaacg aggctgcaag agagatcctc ctgatgccca ctccagggaa 2100

gcatggttgg acagtgctcc agggaagaat cgggcgtcca gacaacccat ttcgagtggc 2160

cctggaatac atctccagtg gaaaccgcag cttgtctgca gtggacttct ttgccctgaa 2220

gaactgcagt gaaggaacat ccccaggctc caagatggcc ctgcagagct ccttcacttg 2280

ttggaatggg acagtcctcc agcttgggca ggcctgtgac ttccaccagg actgtgccca 2340

gggagaagat gagagccaga tgtgccggaa actgcctgtg ggtttttact gcaactttga 2400

agatggcttc tgtggctgga cccaaggcac actgtcaccc cacactcctc aatggcaggt 2460

caggacccta aaggatgccc ggttccagga ccaccaagac catgctctat tgctcagtac 2520

cactgatgtc cccgcttctg aaagtgctac agtgaccagt gctacgtttc ctgcaccgat 2580

-continued

caagagctct ccatgtgagc tccgaatgtc ctggctcatt cgtggagtct tgaggggaaa 2640

cgtgtccttg gtgctagtgg agaacaaaac cgggaaggag caaggcagga tggtctggca 2700

tgtcgccgcc tatgaaggct tgagcctgtg gcagtggatg gtgttgcctc tcctcgatgt 2760

gtctgacagg ttctggctgc agatggtcgc atggtgggga caaggatcca gagccatcgt 2820

ggcttttgac aatatctcca tcagcctgga ctgctacctc accattagcg gagaggacaa 2880

gatcctgcag aatacagcac ccaaatcaag aaacctgttt gagagaaacc caaacaagga 2940

gctgaaaccc ggggaaaatt caccaagaca gacccccatc tttgacccta cagttcattg 3000

gctgttcacc acatgtgggg ccagcgggcc ccatggcccc acccaggcac agtgcaacaa 3060

cgcctaccag aactccaacc tgagcgtgga ggtggggagc gagggccccc tgaaaggcat 3120

ccagatctgg aaggtgccag ccaccgacac ctacagcatc tcgggctacg gagctgctgg 3180

cgggaaaggc gggaagaaca ccatgatgcg gtcccacggc gtgtctgtgc tgggcatctt 3240

caacctggag aaggatgaca tgctgtacat cctggttggg cagcagggag aggacgcctg 3300

ccccagtaca aaccagttaa tccagaaagt ctgcattgga gagaacaatg tgatagaaga 3360

agaaatccgt gtgaacagaa gcgtgcatga gtgggcagga ggcggaggag gaggggtgg 3420

agccacctac gtatttaaga tgaaggatgg agtgccggtg cccctgatca ttgcagccgg 3480

aggtggtggc agggcctacg gggccaagac agacacgttc cacccagaga gactggagaa 3540

taactcctcg gttctagggc taaacggcaa ttccggagcc gcaggtggtg gaggtggctg 3600

gaatgataac acttccttgc tctgggccgg aaaatctttg caggagggtg ccaccggagg 3660

acattcctgc ccccaggcca tgaagaagtg ggggtgggag acaagagggg gtttcggagg 3720

gggtggaggg gggtgctcct caggtggagg aggcggagga tatataggcg gcaatgcagc 3780

ctcaaacaat gaccccgaaa tggatgggga agatggggtt tccttcatca gtccactggg 3840

catcctgtac accccagctt taaaagtgat ggaaggccac ggggaagtga atattaagca 3900

ttatctaaac tgcagtcact gtgaggtaga cgaatgtcac atggaccctg aaagccacaa 3960

ggtcatctgc ttctgtgacc acgggacggt gctggctgag gatggcgtct cctgcattgt 4020

gtcacccacc ccggagccac acctgccact ctcgctgatc ctctctgtgg tgacctctgc 4080

cctcgtggcc gccctggtcc tggctttctc cggcatcatg attgtgtacc gccggaagca 4140

ccaggagctg caagccatgc agatggagct gcagagcccc gagtacaagc tgagcaagct 4200

ccgcacctcg accatcatga ccgactacaa ccccaactac tgctttgctg gcaagacctc 4260

ctccatcagt gacctgaagg aggtgccgcg gaaaaacatc accctcattc ggggtctggg 4320

ccatggcgcc tttggggagg tgtatgaagg ccaggtgtcc ggaatgccca acgacccaag 4380

ccccctgcaa gtggctgtga agacgctgcc tgaagtgtgc tctgaacagg acgaactgga 4440

tttcctcatg gaagccctga tcatcagcaa attcaaccac cagaacattg ttcgctgcat 4500

tggggtgagc ctgcaatccc tgccccggtt catcctgctg gagctcatgg cgggggaga 4560

cctcaagtcc ttcctccgag agacccgccc tcgcccgagc cagccctcct ccctggccat 4620

gctggacctt ctgcacgtgg ctcgggacat tgcctgtggc tgtcagtatt tggaggaaaa 4680

ccacttcatc caccgagaca ttgctgccag aaactgcctc ttgacctgtc caggccctgg 4740

aagagtggcc aagattggag acttcgggat ggcccgagac atctacaggg cgagctacta 4800

tagaaaggga ggctgtgcca tgctgccagt taagtggatg ccccagagg ccttcatgga 4860

aggaatattc acttctaaaa cagacacatg gtcctttgga gtgctgctat gggaaatctt 4920

JX-003.41

Appx169

-continued

ttctcttgga tatatgccat accccagcaa aagcaaccag gaagttctgg agtttgtcac 4980

cagtggaggc cggatggacc cacccaagaa ctgccctggg cctgtatacc ggataatgac 5040

tcagtgctgg caacatcagc ctgaagacag gcccaacttt gccatcattt tggagaggat 5100

tgaatactgc acccaggacc cggatgtaat caacaccgct ttgccgatag aatatggtcc 5160

acttgtggaa gaggaagaga aagtgcctgt gaggcccaag gaccctgagg gggttcctcc 5220

tctcctggtc tctcaacagg caaaacggga ggaggagcgc agcccagctg ccccaccacc 5280

tctgcctacc acctcctctg gcaaggctgc aaagaaaccc acagctgcag agatctctgt 5340

tcgagtccct agagggccgg ccgtggaagg gggacacgtg aatatggcat tctctcagtc 5400

caaccctcct tcggagttgc acaaggtcca cggatccaga aacaagccca ccagcttgtg 5460

gaacccaacg tacggctcct ggtttacaga gaaacccacc aaaaagaata atcctatagc 5520

aaagaaggag ccacacgaca ggggtaacct ggggctggag ggaagctgta ctgtcccacc 5580

taacgttgca actgggagac ttccgggggc ctcactgctc ctagagccct cttcgctgac 5640

tgccaatatg aaggaggtac ctctgttcag gctacgtcac ttcccttgtg ggaatgtcaa 5700

ttacggctac cagcaacagg gcttgccctt agaagccgct actgcccctg gagctggtca 5760

ttacgaggat accattctga aaagcaagaa tagcatgaac cagcctgggc cctgagctcg 5820

gtcgcacact cacttctctt ccttgggatc cctaagaccg tggaggagag agaggcaatg 5880

gctccttcac aaaccagaga ccaaatgtca cgttttgttt tgtgccaacc tattttgaag 5940

taccaccaaa aaagctgtat tttgaaaatg ctttagaaag gttttgagca tgggttcatc 6000

ctattctttc gaaagaagaa aatatcataa aaatgagtga taaatacaag gcccagatgt 6060

ggttgcataa ggtttttatg catgtttgtt gtatacttcc ttatgcttct ttcaaattgt 6120

gtgtgctctg cttcaatgta gtcagaatta gctgcttcta tgtttcatag ttggggtcat 6180

agatgtttcc ttgccttgtt gatgtggaca tgagccattt gaggggagag ggaacggaaa 6240

taaaggagtt atttgtaatg actaaaa 6267

<210> SEQ ID NO 3
<211> LENGTH: 4174
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 3

agtcccgcga ccgaagcagg gcgcgcagca gcgctgagtg ccccggaacg tgcgtcgcgc 60

ccccagtgtc cgtcgcgtcc gccgcgcccc gggcggggat ggggcggcca gactgagcgc 120

cgcacccgcc atccagaccc gccggcccta gccgcagtcc ctccagccgt ggccccagcg 180

cgcacgggcg atggcgaagg cgacgtccgg tgccgcgggg ctgcgtctgc tgttgctgct 240

gctgctgccg ctgctaggca aagtggcatt gggcctctac ttctcgaggg atgcttactg 300

ggagaagctg tatgtggacc aggcggccgg cacgcccttg ctgtacgtcc atgccctgcg 360

ggacgccccc gaggaggtgc ccagcttccg cctgggccag catctctacg gcacgtaccg 420

cacacggctg catgagaaca actggatctg catccaggag gacaccggcc tcctctacct 480

taaccggagc ctggaccata gctcctggga gaagctcagt gtccgcaacc gcggctttcc 540

cctgctcacc gtctacctca aggtcttcct gtcacccaca tcccttcgtg agggcgagtg 600

ccagtggcca ggctgtgccc gcgtatactt ctccttcttc aacacctcct ttccagcctg 660

cagctccctc aagccccggg agctctgctt cccagagaca aggccctcct tccgcattcg 720

ggagaaccga cccccaggca ccttccacca gttccgcctg ctgcctgtgc agttcttgtg 780

JX-003.42

Appx170

-continued

ccccaacatc agcgtggcct acaggctcct ggagggtgag ggtctgccct tccgctgcgc 840

cccggacagc ctggaggtga gcacgcgctg ggccctggac cgcgagcagc gggagaagta 900

cgagctggtg gccgtgtgca ccgtgcacgc cggcgcgcgc gaggaggtgg tgatggtgcc 960

cttcccggtg accgtgtacg acgaggacga ctcggcgccc accttccccg cgggcgtcga 1020

caccgccagc gccgtggtgg agttcaagcg gaaggaggac accgtggtgg ccacgctgcg 1080

tgtcttcgat gcagacgtgg tacctgcatc aggggagctg gtgaggcggt acacaagcac 1140

gctgctcccc ggggacacct gggcccagca gaccttccgg gtggaacact ggcccaacga 1200

gacctcggtc caggccaacg gcagcttcgt gcgggcgacc gtacatgact ataggctggt 1260

tctcaaccgg aacctctcca tctcggagaa ccgcaccatg cagctggcgg tgctggtcaa 1320

tgactcagac ttccagggcc caggagcggg cgtcctcttg ctccacttca acgtgtcggt 1380

gctgccggtc agcctgcacc tgcccagtac ctactccctc tccgtgagca ggagggctcg 1440

ccgatttgcc cagatcggga aagtctgtgt ggaaaactgc caggcattca gtggcatcaa 1500

cgtccagtac aagctgcatt cctctggtgc caactgcagc acgctagggg tggtcacctc 1560

agccgaggac acctcgggga tcctgtttgt gaatgacacc aaggccctgc ggcggcccaa 1620

gtgtgccgaa cttcactaca tggtggtggc caccgaccag cagacctcta ggcaggccca 1680

ggcccagctg cttgtaacag tggaggggtc atatgtggcc gaggaggcgg gctgcccccc 1740

gtcctgtgca gtcagcaaga gacggctgga gtgtgaggag tgtggcggcc tgggctcccc 1800

aacaggcagg tgtgagtgga ggcaaggaga tggcaaaggg atcaccagga acttctccac 1860

ctgctctccc agcaccaaga cctgccccga cggccactgc gatgttgtgg agacccaaga 1920

catcaacatt tgccctcagg actgcctccg gggcagcatt gttgggggac acgagcctgg 1980

ggagccccgg gggattaaag ctggctatgg cacctgcaac tgcttccctg aggaggagaa 2040

gtgcttctgc gagcccgaag acatccagga tccactgtgc gacgagctgt gccgcacggt 2100

gatcgcagcc gctgtcctct tctccttcat cgtctcggtg ctgctgtctg ccttctgcat 2160

ccactgctac cacaagtttg cccacaagcc acccatctcc tcagctgaga tgaccttccg 2220

gaggcccgcc caggccttcc cggtcagcta ctcctcttcc ggtgcccgcc ggccctcgct 2280

ggactccatg gagaaccagg tctccgtgga tgccttcaag atcctggagg atccaaagtg 2340

ggaattccct cggaagaact tggttcttgg aaaaactcta ggagaaggcg aatttggaaa 2400

agtggtcaag gcaacggcct tccatctgaa aggcagagca gggtacacca cggtggccgt 2460

gaagatgctg aaagagaacg cctccccgag tgagcttcga gacctgctgt cagagttcaa 2520

cgtcctgaag caggtcaacc acccacatgt catcaaattg tatggggcct gcagccagga 2580

tggcccgctc ctcctcatcg tggagtacgc caaatacggc tccctgcggg gcttcctccg 2640

cgagagccgc aaagtggggc ctggctacct gggcagtgga ggcagccgca actccagctc 2700

cctggaccac ccggatgagc gggccctcac catgggcgac ctcatctcat ttgcctggca 2760

gatctcacag gggatgcagt atctggccga gatgaagctc gttcatcggg acttggcagc 2820

cagaaacatc ctggtagctg aggggcggaa gatgaagatt tcggatttcg gcttgtcccg 2880

agatgtttat gaagaggatt cctacgtgaa gaggagccag ggtcggattc cagttaaatg 2940

gatggcaatt gaatcccttt ttgatcatat ctacaccacg caaagtgatg tatggtcttt 3000

tggtgtcctg ctgtgggaga tcgtgaccct agggggaaac ccctatcctg ggattcctcc 3060

tgagcggctc ttcaaccttc tgaagaccgg ccaccggatg gagaggccag acaactgcag 3120

JX-003.43

Appx171

-continued

cgaggagatg taccgcctga tgctgcaatg ctggaagcag gagccggaca aaaggccggt 3180

gtttgcggac atcagcaaag acctggagaa gatgatggtt aagaggagag actacttgga 3240

ccttgcggcg tccactccat ctgactccct gatttatgac gacggcctct cagaggagga 3300

gacaccgctg gtggactgta ataatgcccc cctccctcga gccctcccct ccacatggat 3360

tgaaaacaaa ctctatggta gaatttccca tgcatttact agattctagc accgctgtcc 3420

cctctgcact atccttcctc tctgtgatgc ttttttaaaaa tgtttctggt ctgaacaaaa 3480

ccaaagtctg ctctgaacct ttttatttgt aaatgtctga ctttgcatcc agtttacatt 3540

taggcattat tgcaactatg ttttttctaaa aggaagtgaa aataagtgta attaccacat 3600

tgcccagcaa cttaggatgg tagaggaaaa aacagatcag ggcggaactc tcaggggaga 3660

ccaagaacag gttgaataag gcgcttctgg ggtgggaatc aagtcatagt acttctactt 3720

taactaagtg gataaatata caaatctggg gaggtattca gttgagaaag gagccaccag 3780

caccactcag cctgcactgg gagcacagcc aggttccccc agacccctcc tgggcaggca 3840

ggtgcctctc agaggccacc cggcactggc gagcagccac tggccaagcc tcagccccag 3900

tcccagccac atgtcctcca tcaggggtag cgaggttgca ggagctggct ggccctggga 3960

ggacgcaccc ccactgctgt tttcacatcc tttcccttac ccaccttcag gacggttgtc 4020

acttatgaag tcagtgctaa agctggagca gttgcttttt gaaagaacat ggtctgtggt 4080

gctgtggtct tacaatggac agtaaatatg gttcttgcca aaactccttc ttttgtcttt 4140

gattaaatac tagaaattta aaaaaaaaaa aaaa 4174

<210> SEQ ID NO 4
<211> LENGTH: 5629
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 4

agtcccgcga ccgaagcagg gcgcgcagca gcgctgagtg ccccggaacg tgcgtcgcgc 60

ccccagtgtc cgtcgcgtcc gccgcgcccc gggcggggat ggggcggcca gactgagcgc 120

cgcacccgcc atccagaccc gccggcccta gccgcagtcc ctccagccgt ggccccagcg 180

cgcacgggcg atggcgaagg cgacgtccgg tgccgcgggg ctgcgtctgc tgttgctgct 240

gctgctgccg ctgctaggca aagtggcatt gggcctctac ttctcgaggg atgcttactg 300

ggagaagctg tatgtggacc aggcggccgg cacgcccttg ctgtacgtcc atgccctgcg 360

ggacgccccc gaggaggtgc ccagcttccg cctgggccag catctctacg gcacgtaccg 420

cacacggctg catgagaaca actggatctg catccaggag gacaccggcc tcctctacct 480

taaccggagc ctggaccata gctcctggga gaagctcagt gtccgcaacc gcggctttcc 540

cctgctcacc gtctacctca aggtcttcct gtcacccaca tcccttcgtg agggcgagtg 600

ccagtggcca ggctgtgccc gcgtatactt ctccttcttc aacacctcct ttccagcctg 660

cagctccctc aagccccggg agctctgctt cccagagaca aggccctcct tccgcattcg 720

ggagaaccga cccccaggca ccttccacca gttccgcctg ctgcctgtgc agttcttgtg 780

ccccaacatc agcgtggcct acaggctcct ggagggtgag ggtctgccct tccgctgcgc 840

cccggacagc ctggaggtga gcacgcgctg ggccctggac cgcgagcagc gggagaagta 900

cgagctggtg gccgtgtgca ccgtgcacgc cggcgcgcgc gaggaggtgg tgatggtgcc 960

cttcccggtg accgtgtacg acgaggacga ctcggcgccc accttccccg cgggcgtcga 1020

caccgccagc gccgtggtgg agttcaagcg gaaggaggac accgtggtgg ccacgctgcg 1080

-continued

tgtcttcgat gcagacgtgg tacctgcatc aggggagctg gtgaggcggt acacaagcac 1140

gctgctcccc ggggacacct gggcccagca gaccttccgg gtggaacact ggcccaacga 1200

gacctcggtc caggccaacg gcagcttcgt gcgggcgacc gtacatgact ataggctggt 1260

tctcaaccgg aacctctcca tctcggagaa ccgcaccatg cagctggcgg tgctggtcaa 1320

tgactcagac ttccagggcc caggagcggg cgtcctcttg ctccacttca acgtgtcggt 1380

gctgccggtc agcctgcacc tgcccagtac ctactccctc tccgtgagca ggagggctcg 1440

ccgatttgcc cagatcggga aagtctgtgt ggaaaactgc caggcattca gtggcatcaa 1500

cgtccagtac aagctgcatt cctctggtgc caactgcagc acgctagggg tggtcacctc 1560

agccgaggac acctcgggga tcctgtttgt gaatgacacc aaggccctgc ggcggcccaa 1620

gtgtgccgaa cttcactaca tggtggtggc caccgaccag cagacctcta ggcaggccca 1680

ggcccagctg cttgtaacag tggaggggtc atatgtggcc gaggaggcgg gctgcccgct 1740

gtcctgtgca gtcagcaaga gacggctgga gtgtgaggag tgtggcggcc tgggctcccc 1800

aacaggcagg tgtgagtgga ggcaaggaga tggcaaaggg atcaccagga acttctccac 1860

ctgctctccc agcaccaaga cctgccccga cggccactgc gatgttgtgg agacccaaga 1920

catcaacatt tgccctcagg actgcctccg gggcagcatt gttgggggac acgagcctgg 1980

ggagccccgg gggattaaag ctggctatgg cacctgcaac tgcttccctg aggaggagaa 2040

gtgcttctgc gagcccgaag acatccagga tccactgtgc gacgagctgt gccgcacggt 2100

gatcgcagcc gctgtcctct tctccttcat cgtctcggtg ctgctgtctg ccttctgcat 2160

ccactgctac cacaagtttg cccacaagcc acccatctcc tcagctgaga tgaccttccg 2220

gaggcccgcc caggccttcc cggtcagcta ctcctcttcc ggtgcccgcc ggccctcgct 2280

ggactccatg gagaaccagg tctccgtgga tgccttcaag atcctggagg atccaaagtg 2340

ggaattccct cggaagaact tggttcttgg aaaaactcta ggagaaggcg aatttggaaa 2400

agtggtcaag gcaacggcct tccatctgaa aggcagagca gggtacacca cggtggccgt 2460

gaagatgctg aaagagaacg cctccccgag tgagcttcga gacctgctgt cagagttcaa 2520

cgtcctgaag caggtcaacc acccacatgt catcaaattg tatggggcct gcagccagga 2580

tggcccgctc ctcctcatcg tggagtacgc caaatacggc tccctgcggg gcttcctccg 2640

cgagagccgc aaagtggggc ctggctacct gggcagtgga ggcagccgca actccagctc 2700

cctggaccac ccggatgagc gggccctcac catgggcgac ctcatctcat ttgcctggca 2760

gatctcacag gggatgcagt atctggccga gatgaagctc gttcatcggg acttggcagc 2820

cagaaacatc ctggtagctg aggggcggaa gatgaagatt tcggatttcg gcttgtcccg 2880

agatgtttat gaagaggatt cctacgtgaa gaggagccag ggtcggattc cagttaaatg 2940

gatggcaatt gaatcccttt ttgatcatat ctacaccacg caaagtgatg tatggtcttt 3000

tggtgtcctg ctgtgggaga tcgtgaccct agggggaaac ccctatcctg ggattcctcc 3060

tgagcggctc ttcaaccttc tgaagaccgg ccaccggatg gagaggccag acaactgcag 3120

cgaggagatg taccgcctga tgctgcaatg ctggaagcag gagccggaca aaaggccggt 3180

gtttgcggac atcagcaaag acctggagaa gatgatggtt aagaggagag actacttgga 3240

ccttgcggcg tccactccat ctgactccct gatttatgac gacggcctct cagaggagga 3300

gacaccgctg gtggactgta ataatgcccc cctccctcga gccctccctt ccacatggat 3360

tgaaaacaaa ctctatggca tgtcagaccc gaactggcct ggagagagtc ctgtaccact 3420

-continued

cacgagagct gatggcacta acactgggtt tccaagatat ccaaatgata gtgtatatgc 3480

taactggatg ctttcaccct cagcggcaaa attaatggac acgtttgata gttaacattt 3540

ctttgtgaaa ggtaatggac tcacaagggg aagaaacatg ctgagaatgg aaagtctacc 3600

ggccctttct ttgtgaacgt cacattggcc gagccgtgtt cagttcccag gtggcagact 3660

cgtttttggt agtttgtttt aacttccaag gtggttttac ttctgatagc cggtgatttt 3720

ccctcctagc agacatgcca caccgggtaa gagctctgag tcttagtggt taagcattcc 3780

tttctcttca gtgcccagca gcacccagtg ttggtctgtg tccatcagtg accaccaaca 3840

ttctgtgttc acatgtgtgg gtccaacact tactacctgg tgtatgaaat tggacctgaa 3900

ctgttggatt tttctagttg ccgccaaaca aggcaaaaaa atttaaacat gaagcacaca 3960

cacaaaaaag gcagtaggaa aaatgctggc cctgatgacc tgtccttatt cagaatgaga 4020

gactgcgggg ggggcctggg ggtagtgtca atgcccctcc agggctggag gggaagaggg 4080

gccccgagga tgggcctggg ctcagcattc gagatcttga gaatgatttt tttttaatca 4140

tgcaaccttt ccttaggaag acatttggtt ttcatcatga ttaagatgat tcctagattt 4200

agcacaatgg agagattcca tgccatcttt actatgtgga tggtggtatc agggaagagg 4260

gctcacaaga cacatttgtc ccccgggccc accacatcat cctcacgtgt tcggtactga 4320

gcagccacta cccctgatga gaacagtatg aagaaagggg gctgttggag tcccagaatt 4380

gctgacagca gaggctttgc tgctgtgaat cccacctgcc accagcctgc agcacacccc 4440

acagccaagt agaggcgaaa gcagtggctc atcctacctg ttaggagcag gtagggcttg 4500

tactcacttt aatttgaatc ttatcaactt actcataaag ggacaggcta gctagctgtg 4560

ttagaagtag caatgacaat gaccaaggac tgctacacct ctgattacaa ttctgatgtg 4620

aaaaagatgg tgtttggctc ttatagagcc tgtgtgaaag gcccatggat cagctcttcc 4680

tgtgtttgta atttaatgct gctacaagat gtttctgttt cttagattct gaccatgact 4740

cataagcttc ttgtcattct tcattgcttg tttgtggtca cagatgcaca acactcctcc 4800

agtcttgtgg gggcagcttt tgggaagtct cagcagctct tctggctgtg ttgtcagcac 4860

tgtaacttcg cagaaaagag tcggattacc aaaacactgc ctgctcttca gacttaaagc 4920

actgatagga cttaaaatag tctcattcaa atactgtatt ttatataggc atttcacaaa 4980

aacagcaaaa ttgtggcatt ttgtgaggcc aaggcttgga tgcgtgtgta atagagcctt 5040

gtggtgtgtg cgcacacacc cagagggaga gtttgaaaaa tgcttattgg acacgtaacc 5100

tggctctaat ttgggctgtt tttcagatac actgtgataa gttcttttac aaatatctat 5160

agacatggta aacttttggt tttcagatat gcttaatgat agtcttacta aatgcagaaa 5220

taagaataaa ctttctcaaa ttattaaaaa tgcctacaca gtaagtgtga attgctgcaa 5280

caggtttgtt ctcaggaggg taagaactcc aggtctaaac agctgaccca gtgatgggga 5340

atttatcctt gaccaattta tccttgacca ataacctaat tgtctattcc tgagttataa 5400

aagtccccat ccttattagc tctactggaa ttttcataca cgtaaatgca gaagttacta 5460

agtattaagt attactgagt attaagtagt aatctgtcag ttattaaaat ttgtaaaatc 5520

tatttatgaa aggtcattaa accagatcat gttccttttt ttgtaatcaa ggtgactaag 5580

aaaatcagtt gtgtaaataa aatcatgtat cataaaaaaa aaaaaaaaa 5629

<210> SEQ ID NO 5
<211> LENGTH: 10
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence

JX-003.46

Appx174

-continued

<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 5

tttttttttt 10

What is claimed herein is:

1. A method of preparing nucleic acids for analysis, the method comprising:

(a) contacting a first nucleic acid template comprising a sequence of a first strand of a double-stranded target nucleic acid with a complementary target-specific primer that comprises a target-specific hybridization sequence, under conditions to promote template-specific hybridization and extension of the target-specific primer;

(b) contacting a second nucleic acid template comprising a sequence of a second strand that is complementary to the sequence of the first strand of the double-stranded target nucleic acid with a plurality of different primers that share a common sequence that is 5' to different hybridization sequences, under conditions to promote template-specific hybridization and extension of at least one of the plurality of different primers, wherein the different hybridization sequences have different 3' ends. and wherein each primer of the plurality of different primers does not anneal to the same sequence of the double-stranded target nucleic acid as any other primer of the plurality of different primers,

wherein, following (a) and (b), an extension product is generated to contain both a sequence that is characteristic of the target-specific primer and a sequence that is characteristic of the at least one of the plurality of different primers; and

(c) subjecting the extension product to an amplification reaction comprising successive rounds of polymerase extension of i) a tail primer that comprises a 3' sequence that specifically anneals to the complement of the common sequence and that comprises a 5' tail sequence, and ii) a primer that specifically anneals to the complement of the target-specific hybridization sequence.

2. The method of claim 1, wherein the first strand or the second strand of the double-stranded target nucleic acid is a ribonucleic acid.

3. The method of claim 1, wherein the first strand or the second strand of the double-stranded target nucleic acid is a deoxyribonucleic acid.

4. The method of claim 1, wherein steps (a) and (b) are performed sequentially.

5. The method of claim 1, wherein the first nucleic acid template comprises an extension product resulting from the hybridization and extension of the at least one of the plurality of different primers in step (b).

6. The method of claim 1, wherein the second nucleic acid template comprises an extension product resulting from the hybridization and extension of the target-specific primer in step (a).

7. The method of claim 2, wherein the ribonucleic acid is a messenger RNA encoded from a chromosomal segment that comprises a genetic rearrangement.

8. The method of claim 3, wherein the double-stranded target nucleic acid is a chromosomal segment that comprises a portion of a genetic rearrangement.

9. The method of claim 8, wherein the genetic rearrangement is an inversion, deletion, or translocation.

10. The method of claim 1, further comprising contacting the extension product or amplified extension product with an immobilized oligonucleotide under conditions in which hybridization occurs between the extension product and immobilized oligonucleotide.

11. The method of claim 1, wherein the double-stranded target nucleic acid comprises a target portion having a known sequence and a flanking portion having an unknown sequence.

12. The method of claim 11, wherein different hybridization sequences are complementary to the flanking portion.

13. The method of claim 11, wherein the target-specific hybridization sequence is complementary to the target portion.

14. The method of claim 1, wherein the target-specific primer further comprises, 5' to the target-specific hybridization sequence, at least one of an index sequence, a barcode sequence and an adaptor sequence.

15. The method of claim 1, wherein the common sequence comprises at least one of an index sequence, barcode sequence and an adaptor sequence.

16. The method of claim 14, wherein the adaptor sequence is a cleavable adaptor sequence for immobilizing oligonucleotides in a flow cell.

17. The method of claim 1, wherein the tail primer comprises a nucleic acid sequence identical to the common sequence.

18. The method of claim 1, wherein the 5' tail sequence comprises at least one of an index sequence, a barcode sequence, an adaptor sequence, and a sequencing primer hybridization sequence.

19. The method of claim 1, wherein the primer that specifically anneals to the complement of the target-specific hybridization sequence comprises a sequence identical to the target-specific primer.

20. The method of claim 1, wherein the primer that specifically anneals to the complement of the target-specific hybridization sequence comprises at least one modified internucleoside linkage.

21. The method of claim 1, wherein the target-specific primer comprises at least one modified internucleoside linkage.

22. The method of claim 1 further comprising purifying the extension product prior to (c).

23. The method of claim 1 further comprising purifying the amplified product of (c).

24. The method of claim 1, wherein the plurality of different primers can specifically anneal to sequences comprised by separate genes.

25. The method of claim 1, wherein the plurality of different primers can specifically anneal to different portions

of the first strand or the second strand of the double-stranded target nucleic acid comprised by a single gene.

26. The method of claim **1**, wherein the plurality of different primers can specifically anneal to different exons of a gene comprising the first strand or the second strand of the double-stranded target nucleic acid.

27. The method of claim **1**, wherein the first strand or the second strand of the double-stranded target nucleic acid is derived from a biological sample.

28. The method of claim **27**, wherein the first strand or the second strand of the double-stranded target nucleic acid is a cDNA produced from RNA in the biological sample.

29. The method of claim **1**, wherein the double-stranded target nucleic acid is derived from a sample comprising nucleic acids that have been ligated to an oligonucleotide comprising a barcode sequence.

* * * * *