



- (51) International Patent Classification:  
*C12P 19/34* (2006.01) *C12Q 1/68* (2006.01)
- (21) International Application Number:  
PCT/US2016/014999
- (22) International Filing Date:  
26 January 2016 (26.01.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
62/108,161 27 January 2015 (27.01.2015) US  
14/788,670 30 June 2015 (30.06.2015) US
- (71) Applicant: **BIOSPYDER TECHNOLOGIES, INC.**  
[US/US]; 5758 San Elijo, Rancho Santa Fe, CA 92067 (US).
- (72) Inventors: **STEVENS, Anthony**; 7276 Golferest Drive, San Diego, CA 92119 (US). **SELIGMANN, Bruce**; 6290 North Nirvana Place, Tuscon, AZ 85750 (US). **YEAKEY, Joanne, M.**; 157 Townwood Way, Encinitas, CA 92024 (US). **MCCOMB, Joel**; Post Office Box 402, Rancho Santa Fe, CA 92067 (US).
- (74) Agent: **FAN, Calvin**; 4564 Pescadero Avenue, San Diego, CA 92107 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,

BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

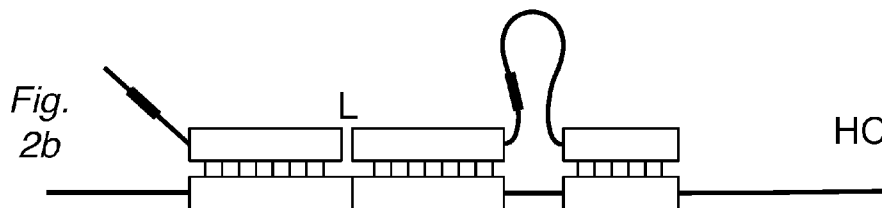
**Declarations under Rule 4.17:**

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

**Published:**

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: LIGATION ASSAYS IN LIQUID PHASE



(57) Abstract: Ligation assays in liquid phase for detecting nucleic acid sequences.



## Ligation Assays in Liquid Phase

### Cross-Reference to Related Applications

This application claims the benefit of priority of U.S. provisional application Ser.  
5 62/108,161, filed January 27, 2015, and U.S. application Ser. 14/788,670, filed June 30,  
2015, the contents of both of which are incorporated herein in their entirety.

### Statement of Government Support

This invention was made with government support under grant 1R43HG007815  
10 awarded by the National Institutes of Health. The government has certain rights in the  
invention.

### Technical Field

This invention relates to molecular biology, and more particularly to assays for  
15 detecting nucleic acid sequences in samples.

### Summary of the Invention

This invention provides methods for detecting target sequences of nucleic acid  
sequences of interest in a sample, and also provides kits for performing the method.

20 In a typical ligation assay, the sample is contacted with a pool of detector oligos,  
where a downstream detector (**DD**) and an upstream detector (**UD**) are provided for each  
target sequence. A portion (**DR'**) of the DD is complementary to a region of the target  
sequence designated as a downstream region (**DR**). The upstream detector has a portion  
(**UR'**) complementary to an upstream region (**UR**) of the target sequence.

25 The downstream and upstream detectors are contacted with the sample and allowed to  
hybridize to the corresponding regions of target sequence present in the sample. When the  
detectors are specifically hybridized to a target sequence, they can be ligated at the junction  
between adjacent detectors, whether directly or after an optional extension step. Formation  
of a ligation product thus serves as evidence that the target sequence was present in the

sample, and the ligation product can be detected by various methods such as detectable labels, microarrays, qPCR, flow-through counters, and sequencing.

The invention provides assays where one or more nucleases are provided during steps in the method to selectively degrade unused or excess detectors, or detectors that are not specifically hybridized to target sequences. Accordingly, the detectors and other components of the assay are configured in a number of embodiments to resist the nucleases while detecting target sequences. The configurations enable sensitive detection of nucleic acids, such as mRNAs and miRNAs, at whole-transcriptome or -miRNome multiplexing and at the level of single cells. Moreover, the steps can be performed in a single well or container without the need for transfers, separation, or solid-phase immobilization, and are therefore ideal for microfluidic platforms.

### Brief Description of the Drawings

**Figure 1** illustrates a representative ligation assay for detecting target nucleic acid sequences. Briefly, downstream detector (**DD**) and upstream detector (**UD**) probe oligonucleotides are allowed to (a) hybridize to a target sequence, having **DR** and **UR** regions, in a sample. For convenience of identification, upstream regions are often underlined herein. While hybridized to the DR and UR of the target sequence, the DD is (b2) ligated selectively to the UR. Optionally, the DD is (b0) extended prior to (b2) ligation. The ligation product is optionally (c) amplified via amplification regions **P1** and **P2'** by one or more primers, such as **P1** and **P2**.

**Figure 2a** shows an “anchored” assay design of the invention where the UD is configured with a second complementary region (**UR2'** or “anchor”) separated by a noncomplementary region (**CP1**). The DD and UD can hybridize to a target sequence as in **Figure 2b**, forming a hybridization complex (**HC**) providing a substrate for ligation at the junction (**L**) between DR' and UR'. After ligation, **Figure 2c** shows the ligation product (**LP**) can be amplified by primers to yield amplification products (**AP**) in **Figure 2d**.

Treatment with an exonuclease, such as an exonuclease with single-stranded 3'-to-5' activity, can be used at various stages of the method to remove undesired components, such as nonbound or excess DD and UD detectors as in **Figure 2e**. Detectors that are nonspecifically or incompletely hybridized to target sequences can be degraded by the exonuclease or will not result in ligation or amplification product, as in **Figure 2f**.

As shown in **Figure 2g**, it may be desirable to provide predetermined quantities of attenuator oligonucleotides such as UR2' (or alternatively UR2) to lessen the formation of product resulting from certain high-abundance target sequences (**HATs**).

**Figure 2h** shows a pair of detectors that are configured to have a modification at one end to resist exonucleases that degrade single-stranded (ss) DNA. The UD has a modification at the 3' end that resists degradation of the detector by an exonuclease having 3' activity on DNA single strands. Alternatively, the DD can have a 5' modification to resist degradation by a 5'-ss-exonuclease.

**Figures 2i** and **2j** illustrate detectors that are configured to resist exonucleases by being hybridized to a protector oligo, such as ones having sequence **DR2** or **UR2** that bind to corresponding DR2' and UR2' sequences of the detectors, presenting double-stranded structures at either end. The protectors can themselves be 5'- or 3'-modified to resist exonucleases, as shown. Figure 2j also illustrates a target sequence (3'-DR-UR-5') that is relatively short, such as a microRNA, where the target has been polyadenylated at its 3' end. The DD features a complementary poly-T portion adjacent to the DR'.

**Figure 3a** depicts a circularizable assay design of the invention using a detector oligo probe (**DO**) that can (a) hybridize via DR' and UR' regions to a target sequence, forming a (noncovalently) circularized structure. After treatment with a nuclease and ligase, a circularized ligation product can then be (c) amplified. **Figures 3b**, **3c**, and **3d** illustrate partially hybridized DO detectors, detectors hybridized to non-target sequences, or nonspecifically hybridized detectors, which can be digested by nucleases or be unsuitable for exponential amplification.

**Figure 4** shows an assay of the invention where a (universal) second strand (**2S**) is provided during hybridization so that the target (DR-UR), DO, and the 2S form a circularized, double-stranded structure. Treatment with ligase results in a covalently circularized ligation product. Optionally, ss-nucleases can be used to degrade excess detectors and hybridization complexes that are not specific for the target. The nucleases can be inactivated. If desired, the circularized structure can be linearized, for example by a restriction endonuclease.

**Figure 5a** shows a detailed view of a hybridization complex using a variant circularizable DO having a short noncomplementary flap (**CP5**) on its 5' end, and optionally a short noncomplementary sequence (**CP3**) on the 3' end. **Figure 5b** shows the hybridization complex after the CP5 is removed by a flap nuclease, such as Fen-1. If desired, the 5' end can be phosphorylated, as in **Figure 5c**. **Figure 5d** illustrates how CP3 can fill in the gap left

by Fen-1, so that the DO can be ligated into circularized form as in **Figure 5e**. The noncomplementary CP5 and/or CP3 flaps can be incorporated in any of the DD and UD designs.

**Figure 6a** provides target sequences (SEQ ID NOs: 33–56) used to design detectors for mRNA expression products for 24 human genes of interest. The genes were selected to demonstrate detection over an expected range of 6 orders of magnitude in abundance, with 10, 1, and 0.1 ng sample RNA input. The number of amplified ligation products, confirmed by sequencing, are shown for anchored detector designs (**Figures 6b, 6c, and 6d**) and circularizable designs (**Figures 6e, 6f, and 6g**). The *x*-axis is for the first technical replicate; the *y*-axis is for the second replicate.

## Detailed Description of the Invention

### ligation assays, generally

A typical ligation assay is illustrated schematically in **Figure 1**, which is discussed in more detail in **Example 1**. A sample that may contain target sequences is contacted with a pool of detector oligonucleotide probes (“probes” or “detectors”). For each target sequence, a pair of detectors is provided: a downstream detector (**DD**) and an upstream detector (**UD**). A downstream detector can have a portion (**DR'**) that is complementary to a region of the target sequence designated as a downstream region (**DR**). An upstream detector can have a portion (**UR'**) that is complementary to a region of the target sequence designated as the upstream region (**UR**). Here, the terms “downstream” and “upstream” are used relative to the 5'-to-3' direction of transcription when the target sequence is a portion of an mRNA, and for convenience the regions designated as upstream are often shown underlined.

As shown in Figure 1, the **DR'** of the **DD** and the **UR'** of the **UD** for each target sequence are allowed to hybridize to the corresponding **DR** and **UR** of the target sequence, if present in the sample. When the **DR** and **UR** of a target sequence are adjacent and the **DR'** and **UR'** of the pair of detector oligos are specifically hybridized to the target sequence to form a hybridization complex, the adjacent detectors **DD** and **UD** can be ligated. Thus, formation of a **DD–UD** ligation product serves as evidence that the target sequence (**DR–UR**) was present in the sample. In cases where the **DR** and **UR** of a target sequence are separated by at least one nucleotide, the ligation step can be preceded by (**b0**) extending the **DR'** using the sample as a template so the extended **DR'** and **UR'** become adjacent and can be ligated.

The ligation product can then be detected by a variety of means; if desired, the products can be amplified prior to detection.

The present invention provides methods where hybridization complexes are exposed at one or more steps to at least one nuclease that can degrade single strands of nucleic acid.

5 As discussed in more detail below, the invention provides detectors and other components of the assay that are configured to selectively resist the nucleases when detecting target sequences. The nucleases can degrade excess or unused detectors, or detectors that are nonspecifically or nonproductively bound to components in the sample that are not of interest. The strategic use of nucleases enables the ligation assay to be performed by adding  
10 one reagent after another in a single reaction container, starting with the sample.

### **samples**

The samples used in the method can be any substance where it is desired to detect whether a target sequence of a nucleic acid of interest is present. Such substances are  
15 typically biological in origin, but can be from artificially created or environmental samples. Biological samples can be from living or dead animals, plants, yeast and other microorganisms, prokaryotes, or cell lines thereof. Particular examples of animals include human, primates, dog, rat, mouse, zebrafish, fruit flies (such as *Drosophila melanogaster*), various worms (such as *Caenorhabditis elegans*) and any other animals studied in  
20 laboratories or as animal models of disease. The samples can be in the form of whole organisms or systems, tissue samples, cell samples, or samples that are cell-free. Particular examples are cancer cells, induced pluripotent stem cells (iPSCs), primary hepatocytes, and lymphocytes and subpopulations thereof. The samples can be provided in liquid phase, such as cell-free homogenates or liquid media from tissue cultures, or nonadherent cells in  
25 suspension, tissue fragments or homogenates, or in solid phase, such as when the sample is mounted on a slide or in the form of formalin-fixed paraffin-embedded (FFPE) tissue or cells, or when cells are grown on or in a surface, as long as detectors can be put into contact for potential hybridization with the sample nucleic acids.

### **30 nucleic acids**

The nucleic acids of interest to be detected in samples include the genome, transcriptome, and other functional sets of nucleic acids, and subsets and fractions thereof. The nucleic acids of interest can be DNA, such as nuclear or mitochondrial DNA, or cDNA

that is reverse transcribed from RNA. The sequence of interest can also be from RNA, such as mRNA, rRNA, tRNA, siRNAs (*e.g.*, small interfering RNAs, small inhibitory RNAs, and synthetic inhibitory RNAs), antisense RNAs, circular RNAs, or long noncoding RNAs. The nucleic acid of interest can be a microRNA (miRNA) at any stage of processing, such as a  
5 primary microRNA (pri-miRNA), precursor microRNA (pre-miRNA), a hairpin-forming microRNA variant (miRNA\*), or a mature miRNA. Detection of microRNAs is discussed in **Example 3a**.

Relatively short nucleic acids of interest, such as mature miRNAs, can be lengthened to enhance hybridization to the detectors. For example, many microRNAs are  
10 phosphorylated at one end, and can be lengthened by chemical or enzymatic ligation with a supplementary oligo. The supplemental oligo can be single-stranded, double-stranded, or partially double-stranded, depending on the ligation method to be used. If desired, the supplemental oligo can be unique to each target sequence, or can be generic to some or all of the target sequences being ligated. The detectors can then be designed with extended DR'  
15 and/or UR' regions that include a portion that hybridizes to the supplemental sequence. A target sequence can also be supplemented by adding nucleotides, such as by polyadenylation, where the extended detectors include at least a portion to hybridize to the supplemental polyA tail. Detection of a family of mature miRNA sequences using extended detectors is discussed in **Example 3b** and illustrated in **Figure 2j**.

20 The amount of nucleic acid in the sample will vary on the type of sample, the complexity, and relative purity of the sample. Because of the sensitivity of the assay, the sample can be taken from a small number of cells, for example from fewer than 100,000, 10,000, 1000, 100, 50, 20, 10, 5, or even from a single cell. The total amount of nucleic acid in the sample can also be quite small: less than 100, 50, 20, 10, 5, 2, 1 micrograms, 500, 200,  
25 100, 50, 20, 10, 5, 2, 1, 0.5, 0.2, 0.1 nanogram, 50, 20, 10, 5, 2, 1 picogram or less of nucleic acid (see **Figure 6d**). The copy number of a particular target sequence can be less than 100,000, 10,000, 1000, 100, 50, 20, 10, 5, or even a single copy present in the sample, particularly when coupled with representative amplification of the ligation product for detection. The amount of input nucleic acid will also vary, of course, depending on the  
30 complexity of the sample and the number of target sequences to be detected.

**detectors**

Based on the particular target sequences, the invention provides pools of detector oligos where a target sequence has a pair of upstream and downstream detectors (**UD** and **DD**) that correspond to DR and UR, which are typically subsequences of the entire nucleic acid sequence of interest. Detectors can be provided to detect targets that contain mutations including individual single-nucleotide polymorphisms (SNPs), gene fusions, and exon-splicing variants. An individual target sequence can have more than one set of DRs and URs, which can be selected by the user to optimize the performance of the assay. Multiple sets of DRs and URs can provide multiple measurements of the same target sequence or of different portions of the target sequence, such as different exons or exon junctions, or provide measurement of a portion of sequence that is not mutated versus a portion of sequence that may harbor a mutation.

***target sequences***

The target sequences can be selected from any combination of sequences or subsequences in the genome or transcriptome of a species or an environment. The set can be specific for a sample type, such as a cell or tissue type. For some sample types, the number of target sequences can range in any combination of upper and lower limits of 1, 2, 5, 10, 20, 50, 100, 200, 500, 1000, 2000, 5000, 10,000, 20,000, 23,000, 30,000, 38,000, 40,000, 50,000, or more. The number of target sequences can also be expressed as a percentage of the total number of a defined set of sequences, such as the RNAs in the human transcriptome or genes in the human genome, ranging in any combination of upper and lower limits of 0.1%, 0.2%, 0.5%, 1%, 2%, 5%, 10%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 65%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, and 100%. Where large sets of detector oligos are used, it can be useful to check the full sequence of each oligo for potential cross-hybridization to other oligos in the set, where, for example, one oligo may inadvertently serve as a template to other detectors. While such non-specific artifacts can be identified by sequence, and are typically discarded from detection results, they may represent noninformative hybridization events that compete for reaction resources.

The detector oligos themselves can be DNA, RNA, or a mixture or hybrid of both. If desired, they can have a modified nucleotide such as dideoxy nucleotides, deoxyUridine (dU), 5-methylCytosine (5mC), 5-hydroxymethylCytosine (5hmC), 5-formylCytosine (5fC), 5-carboxylCytosine (5caC), and Inosine. Yet other modifications to detector oligos include modified bases such as 2,6-diaminopurine, 2-aminopurine, 2-fluro bases, 5-bromoUracil, or

5-nitroindole. Other detector oligos can have a modified sugar-phosphate backbone at one or more positions. Such modifications include a 3'-3' or 5'-5' linkage inversion, a locked nucleic acid (LNA), or a peptide nucleic acid (PNA) backbone. LNAs can be useful for their stronger hybridization properties to complementary bases, enhancing the selectivity or the overall binding affinity for the detector oligo as a whole. The modified bases or bonds can also be used at positions 1, 2, or 3 away from the point of ligation.

As shown schematically in Figure 1, a downstream detector (DD) has a complementary downstream region (DR'), which can be at least 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 35, 40, 45, or 50 nucleotides in length. Similarly, an upstream detector (UD) has a complementary upstream region (UR'), which can be at least 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 35, 40, 45, or 50 nucleotides in length. In a given pair of DD and UD for a target sequence, the DR' and UR' need not be exactly the same length, but will typically be similar so they can hybridize to the target under similar conditions and stringency.

As discussed in more detail below, the detectors can be optimized for ligation, such as by providing a 5'-phosphate on the UD, although this is not necessary, depending on the selection of ligase or other ligation methods. Ribonucleotides can also be substituted at the ligatable ends of the DD and UD to increase the specificity and efficiency of ligation, as when an RNA ligase is used.

### *detector labels*

Where the ligation assay proceeds directly to a detection step, either or both detectors can be designed to be labeled appropriately for detection. For example, the detector can be labeled with a color or fluorescent dye, latex bead, quantum dots, or nanodots. The label can also take the form of an additional nucleotide sequence that serves to enable detection and identification, such as a barcode sequence. For example, a useful barcode sequence can uniquely identify the specific gene or target sequence, or a group of select genes or target sequences within the sample that are being measured. Such sequences can be positioned between the UR' and P2' sequence, and/or between the DR' and P1 sequence, so they are amplified when using flanking primers. This sequence can also be a random sequence, useful for identifying the number of copies of the target gene in the sample, independent of the particular efficiency of any amplification step.

**hybridization**

Returning to the steps of the assay, the detectors are provided so that they contact the sample to allow the detectors to hybridize specifically to the target nucleic acids.

Hybridization conditions can be selected by the skilled artisan to allow and optimize for hybridization between the polynucleotides with the desired degree of specificity or mismatches, and such conditions will vary with the lengths and compositions of sequences present in the hybridization reaction, the nature of any modifications, as well as conditions such as the concentrations of the polynucleotides and ionic strength. Particular hybridization temperatures include 30°, 32.5°, 35°, 37.5°, 40°, 42.5°, 45°, 47.5°, 50°, 52.5°, 55°, 57.5°, 60°, 62.5°, 65°, 67.5°, 70°, 72.5°, 75°, 77.5°, 80°, 82.5°, 85°, 87.5°, and/or 90°. Particular hybridization temperatures can be achieved by ramping the temperature up or down at various rates and profiles, such as timed temperature plateaus, one or more incremental increases or decreases of 5C°, 10C°, or 15C°, and repeated cycling between two or more temperatures. Ions such as Li<sup>+</sup>, Na<sup>+</sup>, K<sup>+</sup>, Ca<sup>2+</sup>, Mg<sup>2+</sup> and/or Mn<sup>2+</sup> can also be present from 0, 1, 2, 5, 10, 20, 50, 100, 200, and 500 mM, and such ions can affect the selection of the other hybridization conditions. Hybridization is also affected by steric crowding components such as branched polysaccharides, glycerol, and polyethylene glycol. Further additives can be present in the hybridization (and subsequent) reactions, such as DMSO, non-ionic detergents, betaine, ethylene glycol, 1,2-propanediol, formamide, tetramethyl ammonium chloride (TMAC), and/or proteins such as bovine serum albumin (BSA), according to the desired specificity.

Optionally, the conditions for hybridization can be adjusted or fine-tuned to permit other steps to be performed in the same environment. For example, the same buffers used for hybridization can be used for lysing cells in a sample, promoting hybridization of certain cell types, facilitating removal or permeation of cell walls, cell membranes, or subcellular fractions, as desired. Depending on the ligation method used in the assay, hybridization conditions can be selected to be compatible with conditions for ligation as is, or with the addition of one or more components and preferably without requiring a change of the reaction container when transitioning from hybridization to ligation steps.

**ligation**

The ligation reaction can occur by chemical ligation or by using a ligase enzyme or a ligation-facilitating co-factor. A variety of nick-repairing ligases are commercially available

to catalyze the formation of a phosphodiester bond between adjacent single-stranded polynucleotides when hybridized to another single-stranded template, such as to join DNA to RNA when hybridized to template. An example is bacteriophage T4 DNA ligase, which is generally understood to use ATP as a co-factor. The ATP can be supplied during the ligase  
5 reaction. In other reactions, the ligase can be pre-adenylated. In yet other reactions, the UD must be pre-adenylated at the 5' end, as with a 5' App DNA/RNA ligase. The UD in a typical reaction will have a 5'-phosphate to facilitate ligation to the DD, although this is not necessary, depending on the selection of ligase and ligation conditions. (Where a 5'-phosphate on the DD is required for efficient ligation, using a comparable oligonucleotide  
10 without 5'-phosphorylation can be used to inhibit or reduce undesired ligation.) Preferred ligation conditions include 10, 25, 50, 100 mM Tris-HCl (pH 7.5, 8.0, or 8.5); at least 10mM, 5mM, 2mM, 1mM MgCl<sub>2</sub>; at least 2mM, 1 mM, 0.7mM, 0.5mM, 0.2mM, 0.1mM, 0.05 mM ATP; or at least 10mM, 7 mM, 5 mM, 2 mM, 1 mM, 0.5 mM DTT or other antioxidant. T3 DNA ligase can also be used, which can ligate a broader range of substrates and has a wider  
15 tolerance for salt concentration. As with other steps, the temperature can be selected according to the characteristics of the reaction components and conditions such as ionic strength.

As discussed above, the ligation step can be preceded by an optional extension step, as in Figure 1, step (b0). The ligation step can also be preceded by an optional cleavage step,  
20 such as by a nuclease, to remove any overhangs. In other cases, a portion of the DD can overlap with the UR sequence to which the UD hybridizes, so that after hybridization of the UD and the DD, there is an overhang sequence of 1, 2, 3, or more bases. A useful enzyme for removing an overhang is a Flap endonuclease, such as Fen-1.

## 25 **amplification**

If desired, the ligation product can be amplified (for example by PCR or qPCR) to facilitate detection. Amplification methods and instruments are commercially available, including PCR plate and droplet formats, and the amplification enzymes (such as *Taq* and its commercial variants) and reaction conditions can be selected and tailored to the particular  
30 platform. Optionally, the polymerase selected for amplification can have strand-displacing activity. As illustrated in Figure. 1, the detectors can have additional sequences (“tails”) including primer hybridization sequences (*e.g.* **P1**, **P2'**) or complements thereof, that serve as amplification sequences, so that after ligation, the ligation product can be amplified with a

pair of amplification primers (**P1**, **P2**). An exemplary downstream amplification sequence (**P1**) is

5' -CAAGCAGAAGACGGCATA**C**GAG-3' (SEQ ID NO:1),

5

which can be used with a primer having the same sequence (**P1**). An exemplary upstream amplification sequence (**P2'**) is

5' -**A**TCTCGGTGGTCGCCGTATCATT-3' (SEQ ID NO:2),

10

which can be used with primer **P2** (shown in 3'-to-5' orientation):

3' -**T**AGAGCCACCAGCGGCATAGTAA-5' (SEQ ID NO:3).

15

If desired, the amplification primer can incorporate a barcode sequence, for example a barcode sequence that uniquely identifies the sample in a multi-sample experiment, and optionally has redundant and/or error-correction features. In some experiments, for example, different sample barcodes can be used for 96, 384, 1536, or more generally  $2^n$  or  $4^n$  different samples that are prepared with different barcodes separately for some steps, such as

20

hybridization, ligation, and amplification, and combined for others, such as detection. The barcode sequence can be incorporated into the primer, such as 3' to the amplification sequence, so that the barcode becomes part of the amplified strand. In other instances, the amplification sequence of the primer can be extended by an additional sequence to provide a primer hybridization sequence that can be used for use in subsequent sequencing steps. The

25

barcode may also be interposed between the amplification sequence, and if desired, the extended amplification sequence, and another sequence that can be used for capture, such as capture onto a surface as part of a sequencing process, and/or for yet another primer hybridization sequence that is used for sequencing. In each case the barcode will be

30

amplified with the rest of the detector sequences, for instance forming a single amplified, elongated molecule that contains sequencing primer hybridization sequences, sample barcode, and a gene-specific sequence, which may include a gene-specific barcode or a target molecule-specific barcode as well as sequence or complement to the sequence of the target gene. In the case where the targeted oligo is a cDNA, a gene-specific sequence or a sample-

specific sequence can be added as part of the primer used for reverse transcription, and be a part of the sequence targeted by the UD and DD.

In other instances, methods known in the art can be used to amplify the ligated DD and UD sequences, such as by repetitive cycles of (1) ligation, (2) heating to melt off the  
5 ligated product, (3) cooling to permit hybridization of DD and UD to the target, (4) ligation, then repeating the heating (2), cooling (3), and ligation (4) steps. These additional amplification steps can be performed before amplification step (c), during which the sample barcodes and other sequences are added to the ligated UD and DD sequence. The target of the UD and DD hybridization may also be amplified by whole transcriptome amplification of  
10 RNA or amplification of cDNA.

### **detection**

The ligation product (or its amplicons) can optionally be detected by methods such as sequencing, qPCR, or labeling for detection on an array or other molecule detection. Other  
15 detection methods include flow-through systems for counting labeled molecules. Depending on the detection method, the skilled user will be able to modify the design of the detectors and amplification primers to include functional features that are appropriate, such as for bridge amplification on a sequencing flow cell. The experimental resources used for amplification and detection can be limited and are often among the most expensive, and their  
20 consumption can be optimized by reducing the number of non-informative assay components present at various stages of the assay.

### **nucleases**

Accordingly, the invention provides nucleases and assay components that are  
25 configured to resist degradation to enable more efficient use of resources and more sensitive detection. As a further advantage, the invention enables a simpler assay workflow that can be performed in a single reaction container or entirely in liquid phase.

The nuclease can be an enzyme that digests or degrades single strands of nucleic acids. Preferably the nuclease does not digest (or has significantly less activity on) double  
30 strands, including DNA:RNA hybrids. For example, the nuclease can have less than 10%, 5%, 2%, 1%, 0.5%, 0.2%, or 0.1% the activity on double strands compared to single-strands on a molar substrate ratio under the same conditions. Similarly, the nuclease can be selected so it does not appreciably digest at single-stranded nicks in a double-strand. The nuclease

can be an endonuclease that degrades single strands, such as mung bean nuclease under certain conditions. The nuclease can also be an exonuclease that degrades single strands, which can be single strands of DNA. For example, a nuclease having single-stranded 3'-to-5' exonuclease (3' exo) activity includes Exonuclease I from *E. coli* (exo I) and T3 exonuclease.

5 Enzymes such as exonuclease T (RNase T), which has 3' exo activity on DNA and RNA single strands, can be used as long as the detectors have been ligated and the RNA strands are no longer needed in the assay. Nucleases having single-stranded 5'-to-3' exonuclease activity include exonuclease VIII and RecJ<sub>f</sub>. The nuclease can be an enzyme that digests 5' overhangs or flaps, such as Flap endonuclease 1. Nucleases can be used singly or in a  
10 cocktail of nucleases, such as a pair of 3' and 5' exonucleases.

The nucleases can be used at various stages of the assay. For example, a nuclease can be provided (b2) after the ligation step (b1) to remove unligated or excess detectors, as in **Figure 2e**. The nuclease can also degrade detectors that are only partially or nonspecifically hybridized to target sequences, as in **Figure 2f**. If compatible with the ligation conditions  
15 used, the nuclease can also be provided during the ligation step (b1 and b2 together), or even before the ligation step (b2, then b1) as long as it does not interfere with the intended detection of target sequences. Depending on the assay design, the nuclease can be provided before, during, or after the optional (b0) extension and (d) amplification steps, or at multiple steps to effect the desired purpose of removing undesired target, detectors, other oligos, or  
20 any products.

When the nuclease activity is no longer desired, the nucleases can be removed or inactivated, such as after the ligation step. Nucleases can be inactivated by methods selected for a particular nuclease but will not substantially interfere with the rest of the assay. For some nucleases, a nuclease inhibitor (as in Figure 4, lower right) or chelating agent, such as  
25 EDTA, can be added as long as it does not interfere with (or can be removed prior to) a subsequent step that may require Mg<sup>++</sup> for example. Other nucleases can be inactivated by heat, for example single or repeated incubation at 70°C, 75°C, 80°C, 85°C, 90°C, 95°C or 98°C, for 1, 2, 5, 10, 15, 20, 25, 30, 45 minutes, or 1 hour. If more than one nuclease is used, either or both may be inactivated individually or by the same means. To resist the activity of  
30 nucleases provided at one or more steps of the invention, components of the assay are provided by the invention in various configurations that permit detection of target sequences. Selection of the configuration method will depend, of course, on the particular nuclease being used.

### anchored detectors

In one configuration, the upstream detector has a second region (UR2') that is complementary to a second region of the target sequence (UR2), as illustrated in **Figure 2a**.

5 Because the tail of the UD can hybridize to a separate portion of the target, this configuration can be described as an “anchored” detector, as in **Figure 2b**. The anchor at the 3' end of the UD hybridizes with the target to form a double-strand and is thus configured to resist digestion to nucleases that degrade single strands, such as 3' exonucleases like *exo I*.

As a separate target-binding region, the anchor UR2' can be used to provide  
10 additional discrimination between similar sequences, such as isoforms of a family of genes where sequence differences between isoforms are found beyond the range of the DR and UR target sequence.

The UR2' can be at least 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 35, 40, 45, or 50 nucleotides in length. The UR2' can be separated from the UR' by a  
15 noncomplementary region (CP1), which can be at least 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 35, 40, 45, 50, 60, 70, 80, 90, or 100 nucleotides in length. In general, the UR2' will be upstream relative to the UR'. If an amplification region (such as P2') is present, it can be upstream of the UR', such as within the CP1 or part of UR2' to allow amplification of the UR' portion as shown in **Figure 2c** to generate the amplification products (AP) in **Figure 2d**.

20 In a mirror-image configuration, it is the downstream detector that has the anchor region (DR2') complementary to a second region of the target sequence. The DR2' anchor hybridizes to a DR2 on the target so that the configuration resists the action of 5' ss-exonucleases. The UR2' of the DD will generally be downstream relative to the UR'. If an amplification region (such as P1) is present, it can be downstream of the DR' to allow  
25 amplification of the DR' after ligation. Anchored DDs and UDs can be used separately or in combination to resist a cocktail of nucleases.

Because the separate anchor region of the detector can affect the hybridization characteristics of the detector via monomolecular kinetics, the compositions and relative lengths of the DR2', CP1(s), DR', UR' and UR2' can be tuned to optimize target selectivity  
30 between the detector pair and among the pairs of the detector pool.

Detectors that are not used in the ligation reaction can be degraded as shown in **Figure 2e**. Moreover, incompletely bound detectors, such as those in **Figure 2f**, can also be degraded, for example when the UR' of a UD binds to the UR of a target, but the UR2' does

not bind, whether because the UR' is bound to a non-target sequence or to a target that was related to the intended target UR but lacked a UR2. Similarly, an anchored DD that binds a DR2 but not the DR of a target will be susceptible to a 3' ss-exonuclease (or will not generate a valid ligation product with a corresponding UD). Other detectors will fail to be amplified, for example detectors in excess of target sequence in the sample or detectors that are bound nonspecifically to nontarget sequences. The use of anchored detectors can therefore increase the specificity of the ligation assay for target sequences while allowing nucleases to degrade excess or unused detectors.

## 10 **blocked detectors**

Another configuration has detectors that are nuclease-resistant by having a nuclease-blocking group at or adjacent to one end. **Figure 2h** shows a DD, having a 5'-blocking group, that can be used in combination with a 5' exonuclease. Also shown is a UD having a 3'-blocking group for use with a 3' exonuclease. Preferably when a 5' or 3' exonuclease is used where there are multiple targets and pairs of detectors, all of the downstream or upstream detectors have a 5' or 3' block, respectively.

Useful configurations for resisting nucleases include termination with an inverted nucleotide such as deoxythymidine (idT), a dideoxynucleotide such as dideoxythymidine (ddT or iddT), or 2'/3'-O-acetylation of the terminal nucleotide. Depending on the substrate preferences of the nuclease selected, one or more of the other modified nucleotides described earlier can be used as a blocking group. Alternatively, one or more of the terminal nucleotides are attached to the rest of the oligo via one or more phosphorothioate bonds instead of naturally occurring phosphodiester bonds. Other modifications that may resist a nuclease include the LNA or PNA backbones discussed earlier. In some configurations, a hairpin loop or other secondary structure on the detector can serve as the nuclease-blocking group for a detector. One end of the hairpin can have a blocking group. In other configurations, prior to hybridization, a protein or other component can be bound the 5' end of a DD or the 3' end of a UD, such as a sequence-specific single-strand-binding protein like a far upstream element (FUSE) binding protein (FUBP) via a ssFUSE sequence incorporated into a detector. If the 5' end of a DD or the 3' end of a UD detector is configured to be immobilized, whether permanently or reversibly, to a solid phase, the solid phase itself can serve as a block against nuclease activity on the detector. It can be useful to combine any of

the preceding features in a single detector or both detectors to resist the action of the nuclease selected and to provide other advantages, such as stability and hybridization properties.

### protectors

5 Yet another configuration provides one or more oligos that protect the detectors by hybridizing to the DD or UD at a region that will not interfere with hybridization of the DR' or UR' regions complementary to the target sequence. For example in **Figure 2i**, a DR2 protector oligo is provided to hybridize to a DR2' region at the 5' end of the DD, forming a double-stranded configuration (indicated by a brace) that is resistant to  
10 5' exonucleases. If a 3' exonuclease is to be used, then a UR2 protector can be provided to form a double-strand at the 3' end of the UD. The protector oligos can themselves be protected from exonuclease activity by a blocking group or bond as described above. For example, a 3'-blocked UR2 protector is shown in Figure 2i, and a 5'-blocked DR2 protector is shown in **Figure 2j**. If a cocktail of 5' and 3' exonucleases is to be used, then both DR2 and  
15 UR2 protectors can be provided, optionally with 5'- or 3'-blocking groups, respectively.

### circularizable detectors

In a circularizable configuration with one detector, the upstream complementary region (UR') and downstream complementary region (DR') are on a single, circularizable  
20 detector oligo (DO), as shown in **Figure 3a**. The DO can have in the 5'-to-3' direction: (B) an upstream complementary region (UR'); (C) an optional amplification region (P2'); (D) a noncomplementary region (CP2) having a sequence that is not complementary to the target sequence; (F) a downstream complementary region (DR'); and (E) an optional amplification region (P1). The DO can be at least 50, 60, 70, 80, 90, 100, 110, 120, 130, 140,  
25 150, 200 bases in length to allow the molecule flexibility to circularize.

An alternate circularizable configuration with two detectors has a DD with a CS portion at the 5' end, and an UD with a reverse complementary CS' portion at the 3' end, so that the DD and UD are partially hybridized to each other via the CS and CS' portions. Optionally there are blocking groups at the 5' end of the CS portion or the 3' end of the  
30 CS' portion. Another circularizable configuration has three oligos: two detectors and a bridge oligo: the DD has a CS1 portion at the 5' end; the bridge oligo has a CS1' portion and a CS2' portion; and the UD has a CS2 portion at the 3' end. The bridge oligo optionally has

blocking groups at the 5' end and/or the 3' end.

In the presence of a target sequence DR-UR, the circularizable detector(s) can (a) circularize on the target, forming a hybridization complex (**HC**) that is resistant to single-stranded exonucleases and that can be (b2) ligated.

5 If the amplification regions are provided in the appropriate orientation, the ligation product (**LP**) can be (c) amplified with P1 and P2 primers to form amplification product (**AP**) that contains the joined DR' and UR' regions.

The DOs that are not specifically hybridized to the target or are bound incompletely to the target are susceptible to degradation by nucleases (**Figure 3d**) or the P1 and P2' amplification regions will not be in the correct orientations for primer amplification, as illustrated in **Figure 3b** or **3c**. In some instances, the detector may be amplified, but it will be amplified linearly, rather than exponentially. In such cases, the minor sequences can be detected and discounted or removed from the detection results computationally.

#### *second single-strand (2S)*

15 Still another configuration provides a single-stranded DNA oligonucleotide (**2S**) to hybridize to the single-stranded portion of the detector to form a double-stranded hybridization complex, as illustrated in **Figure 4**. The 2S oligo can be complementary to the CP1 so that the entire structure becomes double-stranded. Where the assay is intended to detect multiple target sequences, the same 2S can be used generically to form the circular structure since it does not rely on hybridization to target sequences. The structure can then be ligated, completing the circular, double-stranded structure and resistant to exonucleases, ss-  
20 endonucleases, and nick-endonucleases.

Optionally, the circular structure can be deliberately nicked or cut, for example by a nicking endonuclease. The DO can have a restriction endonuclease recognition site so the circular structure can be linearized if desired. To avoid digesting target sequences or  
25 detectors, the recognition site selected for CP1 can be a relatively rare site such as for *AscI*, *FseI*, *AsiSI*. If desired, linearized structures can be separated from circular structures by conventional methods.

#### *flaps*

30 The circularizable DO can be configured so that it has a (A) a noncomplementary region (CP5) in the 5' direction of the UR' and (G) an optional noncomplementary region (CP3) in the 3' direction of the DR', as shown in **Figure 5a** and discussed in **Example 4**.

A second strand can be provided that has, in the 5'-to-3' direction: P2, CP2', P1' so that, together, the target nucleic acid, a detector oligo, and the second strand form a hybridization complex having a 5' flap. A nuclease, such as Fen-1 can be used to remove the 5' flap (**Figure 5b**). The 5' end of the circularizable detector can be phosphorylated (**Figure 5c**).

- 5 If desired, the optional CP3 region can then hybridize to the target sequence, forming a 3' terminus that can be ligated (**Figure 5d**) to the adjacent UR' to form a ligated product (**Figure 5e**).

### steps in solid, liquid phases

- 10 In some embodiments, the hybridization, ligation, or extension steps can be performed while the target sequence is *in situ*. This can be particularly useful, for example, when the sample is on a histological slide, so that the ligation is known to occur at a recordable location and can be compared to similar reactions at other locations on the slide. The ligated probes can remain at the location while other steps are performed, such as
- 15 imaging or detection of other analytes at or near the location. If desired, the ligated probes can remain *in situ* more securely by a variety of chemical or enzymatic methods for cross-linking to the site, which can be permanent or reversible, such as by a photocleavable link as with using a cyanovinylcarbazole nucleoside analog (<sup>CNV</sup>K). In a particular embodiment, the ligation products can be eluted from the sample *in situ* for collection and further processing,
- 20 preferably eluting from small areas to preserve the location information and morphological context of the ligation reaction products.

- In other embodiments, one or more of the steps can be performed in liquid phase, such as in a microfluidic system, so that one or more of the steps does not involve capture to a solid phase, such as to a bead or a plate surface. For example, any one or combination of
- 25 the hybridization, extension, ligation, nuclease digestion, amplification, or detection steps can be performed in liquid phase. In a mixed phase assay, a solid phase can be used to immobilize one or more of the sample, the detector oligos, the hybridization complex, the extension product, the ligation product, or the amplification product. In particular, the target nucleic acid can be attached to a solid surface during the hybridization step, the ligation step,
- 30 or both. The solid surface can be a bead, such as a magnetic, nonmagnetic, polymeric, reversible immobilization, or latex bead, or compound beads thereof, or a relatively flat surface such as a plate or flowcell surface, optionally with coatings of similar materials. The

mixed phase format allows the components to be transferred from one reaction environment to another, or the conditions to be changed as the components remain in one container.

***adding successively to the same reaction container***

Alternatively, the reactions can be optimized so that at least one of steps is performed  
5 by adding reagent, such as an enzyme or buffer component, successively, so that a reaction  
takes place in the same container as the preceding step, optionally without requiring an  
intervening wash or transfer step. Preferably, the sequence of additions does not require  
significant additions of liquid volumes to dilute the components for the next reaction, for  
example no more than 1-, 1.5-, 2-, 2.5-, 3-, 5-, 10-, 15-, or 20-fold dilution between the initial  
10 sample and preparation for detection. The components to be added can be provided in a kit,  
as described below.

**attenuators**

In cases where there is more than one target sequence in a given sample, it is likely  
15 that they will be present in different amounts. Moreover, the amount of a target sequence can  
vary among similar samples. Ideally, a detection assay will have sufficient dynamic range to  
measure the presence of the different target sequences quantitatively in a single experiment.  
For some types of samples, however, the range of abundance for various target sequences can  
span several orders of magnitude. For example, when profiling the RNA expression products  
20 of a cell, individual sequences of particular interest may be present in very few copies, while  
others are highly abundant target sequences (**HATs**). The HATs can be present in a sample  
in such large numbers that they may diminish the ability of a method to detect the presence of  
less abundant target sequences.

Depending on the cell or tissue type, such highly abundant HATs can include  
25 sequences encoding what are generally referred to as housekeeping genes. Examples of  
HATs include sequences that encode all or a portion of myoglobins, actins, tubulins,  
ubiquitins, heat-shock proteins (HSPs), ribosomal proteins, ribosomal RNAs (rRNAs),  
micro-RNAs (miRNAs), or small nuclear RNAs (snRNAs). Other examples of HATs can  
encode all or a portion of cytochrome *c*, glyceraldehyde 3-phosphate dehydrogenase  
30 (GAPDH), ribosomal protein L7 (RPL7), ribosomal protein S6 (rpS6), snRNA RNUs,  
phosphoglycerokinase (PGK), tyrosine 3-monooxygenase / tryptophan 5-monooxygenase  
activation protein zeta (YWHAZ),  $\beta$ -actin, or  $\beta$ -tubulin. Further examples include sequences  
encoding all or a portion of  $\alpha$ -2-microglobulin, vimentin, and fibronectins. Yet other

examples of HATs encode all or part of a cytochrome such as mitochondrially encoded cytochrome *b* (MT-CYB), outer mitochondrial membrane cytochrome *b5* type B, microsomal cytochrome *b5* type A (ACYB5A), and ascorbate-dependent cytochrome *b3* (CYBASC3).

Because which sequences are highly abundant can differ from one sample type to another,

5 such as between different tissues or cell types, certain target sequences can be designated as a predetermined set of potential HATs based on a search of the literature for that type of sample, or can be determined by performing preliminary assays to determine the more abundant sequences in the sample type. Various attenuator oligonucleotides (“attenuators”) can be used to attenuate the overall number of HAT-related ligation products to be detected.  
10 Some attenuators are provided that can to provide positive detection of the HAT in the sample, but at a lower level of signal.

An attenuator useful in the invention is shown in **Figure 2g**, where a UR2' oligo is provided to hybridize to UR2 targets in competition with detectors. Similarly, UR2, DR2', and DR2 oligos can be provided to compete with the binding of portions of anchored  
15 detectors to HATs, thereby attenuating the total number of detectors that form HAT-related ligation products. Particularly useful attenuators can have a portion of DR2 and a portion of DR; or have a portion of UR and a portion of UR2, thereby competing for two portions of the same anchored detector.

For circularizable detector designs, an attenuator can be an oligonucleotide that has a  
20 portion that is identical or complementary to UR or DR, or both. Attenuators can also take the form of oligos that fill a gap, such as shown in **Figure 5b**, but are blocked from yielding a ligatable product.

### **cleavable detectors**

25 It can be desirable for a detector oligo to contain one or other modifications that can be selectively cleaved by treatment after the ligation or optional amplification step. For example, a detector oligo can have a dU located so that it will not interfere with hybridization or ligation steps. After ligation, however, products incorporating the dU oligo can then be  
30 cleaved by dU-specific enzymes, such as uracil-DNA glycosylase followed by endonuclease VIII. Another selectively cleavable site can be a restriction enzyme cleavage site that is not present in the target sequences to be detected.

**kits**

The invention provides kits for performing the methods described above, comprising detector oligos, and optionally a nuclease, a ligase, and/or a polymerase. The kits can further provide reaction buffers for the enzymes in the kit or buffer components to be added to reactions suitable for the enzymes. The component can be suitable for addition to a container for an enzyme reaction to prepare a suitable reaction buffer for the enzyme. The component can also be selected to be compatible with the reaction buffer for the preceding step of the method so that the component can be added to the same container to form a reaction buffer for the next enzyme to be used. Thus, the components can be selected to enable an “add-add-add” strategy for multiple steps of the assay to minimize transfers of sample, oligos, enzymes and/or solutions between separate containers.

The kits can also have eluent solutions suitable for removing oligonucleotides, such as ligated oligonucleotides, from a tissue sample for further analysis. The kits can further have amplification primers suitable for use with the detectors of the kit.

15

**Examples****Example 1: Representative Ligation Assay**

A representative method is provided to illustrate ligation assays. Here, over 100 RNA expression products were detected in a sample of cells using a multiplex assay format. For each expression product, the assay was designed to detect one or more target sequences within the full sequence of the product. For example, in human cells, a *GAPDH* gene of interest encodes the enzyme glyceraldehyde 3-phosphate dehydrogenase; three different portions within the RNA transcript of the *GAPDH* gene were independently detected as target sequences. One such RNA target sequence, identified here as *GAPDH\_2*, was

25

5' - CGACCACUUUGUCAAGCUCAUUUCCUGGUAUGACAACGAAUUUGGCUACA - 3' (SEQ ID NO:4)

where a 5' end was designated “upstream” (underlined) and the 3' end was designated “downstream” for the direction of transcription and translation. The same *GAPDH\_2* target sequence can be shown in the 3'-to-5' direction for later convenience of discussion:

30

3' - ACAUCGGUUUAAGCAACAGUAU GGUCCUUUACUCGAACUGUUUCACCAGC - 5' (SEQ ID NO:5)

A downstream region (DR) was defined as the downstream 25 bases of GAPDH\_2:

3' -ACAUCGGUUUAAGCAACAGUAU**GGU**-5' (SEQ ID NO:6)

5

which has a complementary DNA sequence of DR':

5' -TGTAGCCAAATTCGTTGTCATAC**CA**-3' (SEQ ID NO:7)

10 The upstream region (UR) was defined as the upstream 25 bases of GAPDH\_2:

3' -CCUUUACUCGAACUGUUUCACCAGC-5' (SEQ ID NO:8)

which has a complementary DNA sequence of UR':

15

5' -GGAAATGAGCTTGACAAAGTGGTCG-3' (SEQ ID NO:9)

For GAPDH\_2, a pair of detectors was designed: a downstream detector (DD) having the DR' sequence, and an upstream detector (UD) having the UR' sequence. Similar pairs were designed for each of the target sequences to provide a pool of detectors for the assay. In this example, all the upstream detectors were phosphorylated at the 5' end.

In this particular example, an amplification step was to be performed later in the experiment using two primers, P1 and P2, so all UD's in the experiment included a primer sequence (P1) and all UR's included a complementary primer sequence (P2'). Because amplification is not necessary to the practice of the invention, however, the sequence of the specific primers and primer sequences is a matter of selection to suit the particular amplification method, if used.

At least 10 ng of RNA isolated from human kidney or liver cell lines was placed in a well of a microtiter plate for each assay experiment. To each well was added 20  $\mu$ L of 2X Binding Cocktail, which contained 5 nM of each detector (providing a final input of 0.1 pmoles per oligo), 100 nM biotinylated oligo(dT)<sub>25</sub>, and 5  $\mu$ L streptavidin-coated magnetic beads in a Wash Buffer (40 mM Tris-Cl pH 7.6, 1 M NaCl, 2 mM EDTA disodium, 0.2% SDS).

The plate was heated for 10 min at 65°C to denature the RNA, then the temperature was ramped down over 40 min to 45°C to allow the detectors to anneal to the target sequences in the RNA sample. The plate was then transferred to a magnetic base to immobilize the beads, allowing the supernatant, containing unbound and excess detectors, to be aspirated from the wells. The beads were washed at least three times with 50 µL Wash Buffer.

To each well was added 5 Weiss units of T4 DNA ligase in 20 µL of 1X ligation buffer, as provided by the supplier. After the beads were resuspended by pipette, the plates were incubated for 60 min at 37°C to allow target-dependent ligation of DDs to UDAs as appropriate. After the ligation reaction, the beads were immobilized and washed twice with 50µL Wash Buffer. To release the ligated detectors from their RNA targets, the beads were resuspended in 30µL and incubated for 5 min at 65°C. After incubation, the beads were immobilized, and the supernatant was removed and transferred to a storage plate.

For the optional amplification step, 5 µL of the supernatant, containing the ligation products, was transferred to a well of a PCR plate. Then 10 µL of a PCR cocktail was added, containing 0.45 U *Taq* polymerase, 0.6 µM P1 primer, 0.6 µM P2 primer, 1.5 mM MgCl<sub>2</sub>, and 200 µM dNTPs. The thermocycler used the following program: 10 min at 94°C, followed by 20 to 25 cycles of 30 sec at 94°C, 30 sec at 58°C, and 30 sec at 72°C. The amplification products were then sequenced according to manufacturer's instructions.

This representative ligation assay can be modified as in the following examples.

### Example 2: Anchored and Circularizable Detector Designs

Upstream and downstream detector probe oligonucleotides were prepared as in **Figure 2a** and **3a** for 24 target sequences identified as breast cancer targets: ACTB\_1, TFF1\_1, GATA3\_3, GAPDH\_3, CDH1\_1, KRT19\_2, TIMP1\_2, NFKBIA\_1, ESR1\_1, VEGFA\_3, LAMP1\_2, MUC1\_3, BAD\_3, PTEN\_1, BRCA2\_1, BCAT2\_3, ICAM1\_2, IGF2\_3, BRCA1\_2, EGFR\_1, BMP4\_1, KIT\_3, WNT1\_1, and EGF\_3 (in descending order of expected counts). The targets were selected for a range of expression covering 6 orders of magnitude from ACTB\_1 to EGF\_3. The target sequences used for the DRs and URs are shown in **Figure 6a**.

The assay was performed in triplicate with 100, 10, 1, and 0.1 and 0 (control) nanograms of MCF7 total RNA as sample. The detectors were added to the sample in a

volume of 1 or 2  $\mu$ L and allowed to hybridize by incubating at 65°C for 10 minutes, ramping down over 20 minutes from 65° to 45°C, then held for 20 minutes at 45°C. Exonuclease I (*E. coli*) was added to the hybridization mixture in 6  $\mu$ L of 0.5 Units and incubated for 1 hour at 37°C. T4 ligase was added to the mixture in 6  $\mu$ L of 5 Units and incubated for 1 hour at 37°C. A heat step was performed for 30 minutes at 80°C. The mixture was amplified by adding 2X PCR master mix. The amplification products corresponding to the target sequences were detected and quantificated by qPCR and sequencing. The results are provided in **Figures 6b–6g**.

### 10 **Example 3a: Circularizable Detector Design for microRNAs**

Circularizable DO detectors were designed for the Let-7 family of miRNAs. These miRNAs are initially transcribed as relatively long transcripts (pri-miRNAs), but are processed into pre-miRNAs, and subsequently processed into a relatively short mature form. In mature form, the highly homologous Let-7 family is shown 5'-to-3', with variants from the let-7a sequence bolded).

Hsa let-7a	<u>ugagguaguaggu</u> uguauaguu	SEQ ID NO:10
Hsa let-7b	<u>ugagguaguaggu</u> ug <u>g</u> gguu	SEQ ID NO:11
Hsa let-7c	<u>ugagguaguaggu</u> uguau <u>g</u> guu	SEQ ID NO:12
20 Hsa let-7d	<u>agagguaguaggu</u> ug <u>c</u> auaguu	SEQ ID NO:13
Hsa let-7e	<u>ugagguag</u> <u>g</u> agguuguauaguu	SEQ ID NO:14
Hsa let-7f	<u>ugagguaguag</u> <u>a</u> uuguauaguu	SEQ ID NO:15
Hsa let-7g	<u>ugagguaguag</u> <u>u</u> uuguacaguu	SEQ ID NO:16
Hsa let-7h	<u>ugagguaguag</u> <u>u</u> uugug <u>c</u> u	SEQ ID NO:17

25

Using Hsa let-7a as an example, the DR' was 5'-AACTATAACAAC-3' (SEQ ID NO:18) and the UR' was 5'-CTACTACCTCA-3' (SEQ ID NO:19). A single-stranded DNA oligonucleotide (2S), about 80 nucleotides, is provided to hybridize to the single-stranded portion of the DO to form a double-stranded hybridization complex, as illustrated in

### 30 **Figure 4.**

After hybridization, the region of the DR and UR can be represented as

5'-...TAAGAG-AACTATAACAAC CTACTACCTCA-CGGAAC...-3' SEQ ID NO:20

||||| ||||||| ||||||| |||||||  
 3'-...**ATTCTC** uugauauguug-gaugauggagu **GCCTTG**...-5' SEQ ID NO:21

where the target miRNA is in lowercase. Part of the DO is shown as the upper sequence,  
 5 with the DR' in roman and the UR' underlined roman, flanked by sequence, partially shown,  
 in italics, such as P1 or P2'. The bases in bolded italics represent the 3' end (on the left) and  
 the 5' end (on the right) of the same 2S oligonucleotide.

After ligation, the portion shown forms a double-stranded structure without any nicks

10 5'-...*TAAGAG*-AACTATAACAAC-CTACTACCTCA-CGGAAC...-3' SEQ ID NO:22

||||| ||||||| ||||||| |||||||  
 3'-...**ATTCTC**-uugauauguug-gaugauggagu-**GCCTTG**...-5' SEQ ID NO:23

15 which is resistant to attack by exonucleases.

If the DO for let-7a becomes hybridized to similar let-7c, the following structure is  
 formed:

5'-...*TAAGAG*-AACTATAACAAC-CTACTACCTCA-CGGAAC...-3' SEQ ID NO:24

20 ||||| ||| ||||||| ||||||| |||||||  
 3'-...**ATTCTC**-uugguauauguug-gaugauggagu-**GCCTTG**...-5' SEQ ID NO:25

The complex, which contains a mismatch, can be nicked with a variety of enzymes, such as  
 T4 endonuclease VII, T7 endonuclease I, or in combinations of exonuclease I and *E. coli*  
 25 exonuclease III, S1 nuclease, or nuclease BAL-31. The nicked complex can then be  
 degraded by treatment with a nuclease in step (b1) so that no ligation product is formed.

As illustrated, the covalently circularized, double-stranded structure can be linearized  
 by treatment with a restriction endonuclease, if desired, where the 2S contains an appropriate  
 restriction site. The linearized product can be amplified with primers.

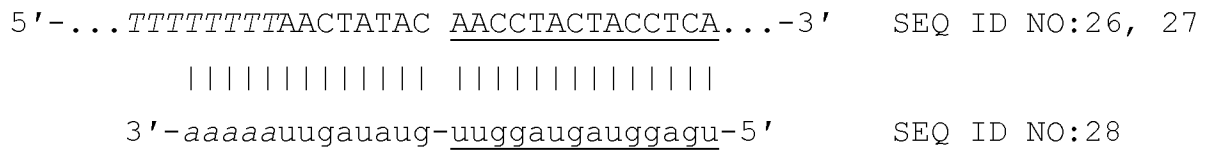
30

**Example 3b: Extended Detector Design for microRNAs**

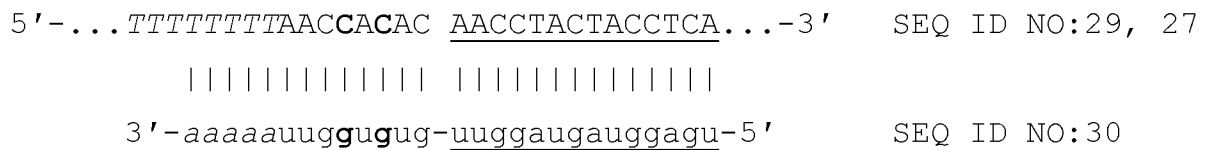
Extended detectors were designed for Let-7 family microRNAs that have been  
 polyadenylated. The microRNAs are extended using polynucleotide adenylyltransferase to

add a 3' polyadenine tail. For a Hsa let-7a microRNA (SEQ ID NO:10), a polyadenylated sequence is shown below (SEQ ID NO:28) in italics. An upstream detector is provided having SEQ ID NO:27 and an extended downstream detector is provided having SEQ ID NO:26, which has an italicized poly-T region (usually poly-dT if the detector is DNA).

5



10 The combination of the supplemental 3' polyadenine tail and the extended poly-T region provides a longer complementary region for hybridization of the target to the detector, and allows greater freedom of designing DRs and URs for the target. For instance, the lengths of the complementary regions for the DD and UD can be more similar in length. When a family of related target sequences is being detected, a DD or UD can be used to detect more than one  
 15 family member (a "generic detector"). Thus for Hsa let-7b,



20

the same upstream detector can be used to detect let-7a and let-7b (and let-7c), since the 14 bases in the 5' direction are identical. Skilled artisans will be able to design various combinations of specific and generic detectors for related sequences, such as the let-7 family, depending on the number of detectors and hybridization properties desired.

25

After the extended detectors are allowed to hybridize to the polyadenylated microRNAs, the detectors are ligated to form the ligation product for detection or optional amplification. If the number of supplemental adenosines added is fewer than the number of dTs in the DD, this does not interfere with the ligation and subsequent steps. If the number of supplemental As is greater, then excess portion of the 3' tail need not hybridize entirely to  
 30 the remaining 5' portion of the DD for specific and target-valid ligation to occur.



**We claim:**

1. A method for detecting target nucleic acid sequences in a sample, wherein a target sequence has a downstream region (DR) and an upstream region (UR), comprising

5 (a) contacting the sample with

a downstream detector oligo (DD) having a complementary downstream region (DR') and

an upstream detector oligo (UD) having a complementary upstream region (UR'),

10 thereby allowing the detectors to hybridize specifically to the target nucleic acids;

(b1) ligating the DR' and UR' if both are specifically hybridized to the DR and UR of a target sequence; and

15 (b2) exposing hybridization complexes to at least one nuclease that degrades single strands but does not significantly degrade double strands, wherein at least one of the DD or UD is configured to resist the nuclease,

whereby nonspecifically hybridized DDs and UDUs are degraded by the nuclease;

20 whereby the ligation product indicates the presence of the target sequence in the sample.

2. The method of claim 1, wherein step (a), step (b1), or both are performed in liquid phase; or wherein step (a) or step (b1) is performed while the target nucleic acid is attached to a solid surface.

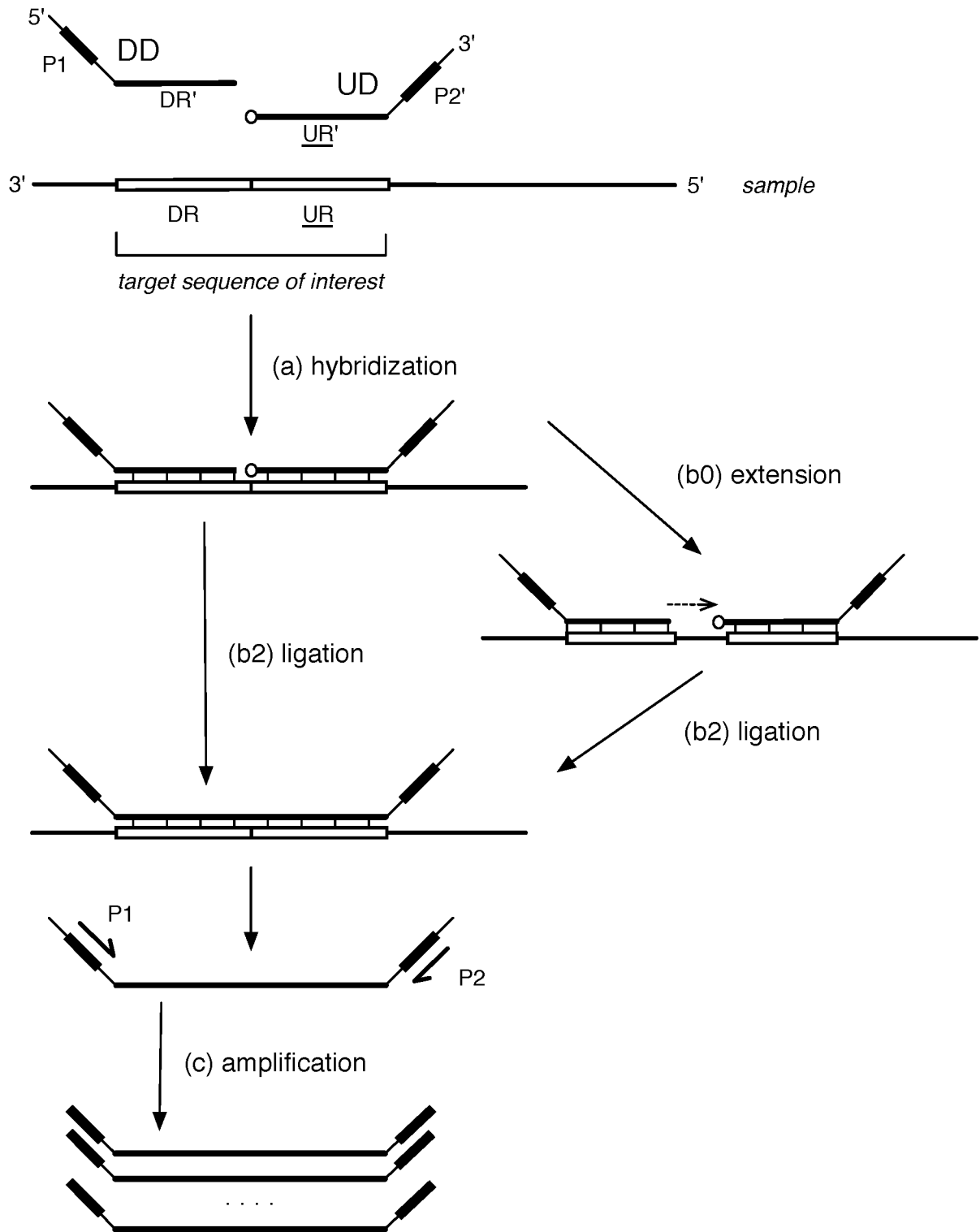
25

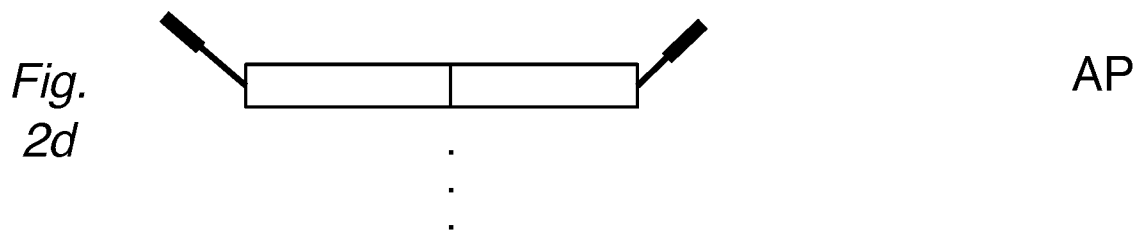
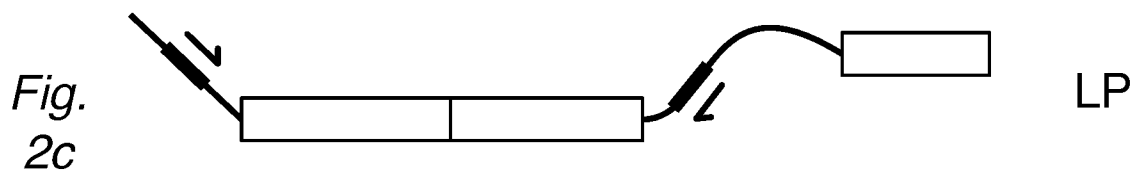
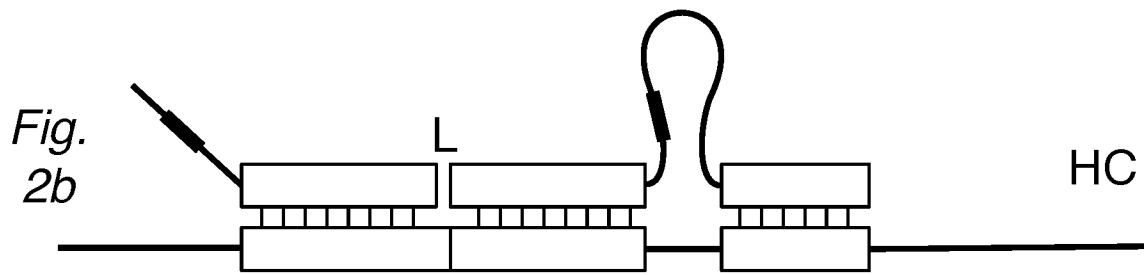
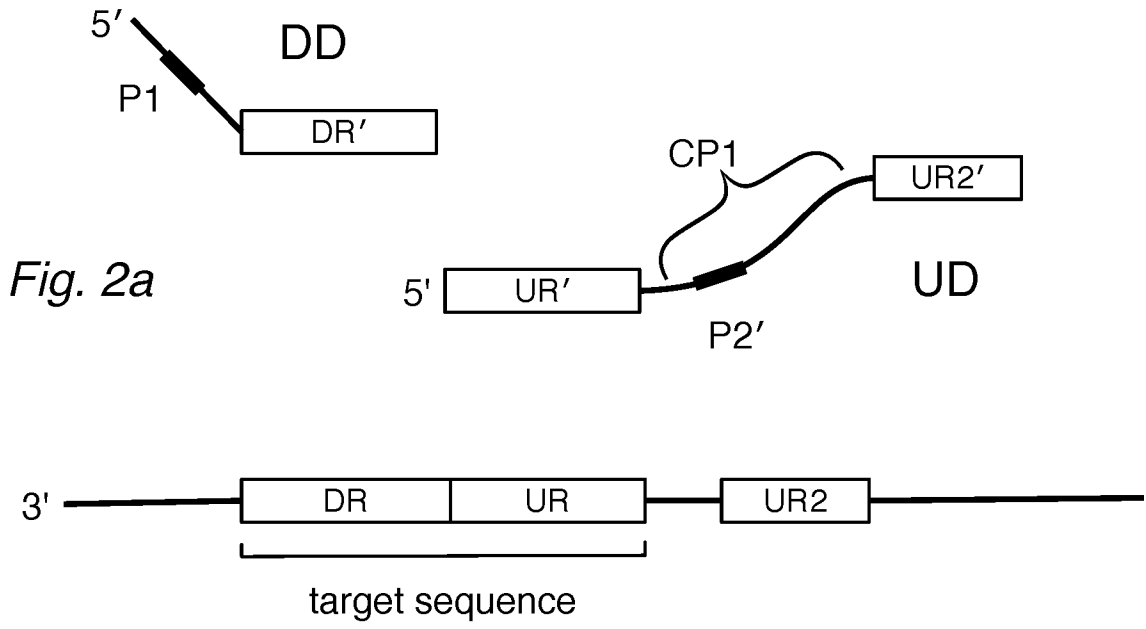
3. The method of claim 1, wherein at least one of steps (a), (b1), or (b2) is performed by adding a reagent to the same reaction container as the preceding step.

4. The method of claim 1, wherein the target nucleic acid is a microRNA or portion thereof, wherein optionally the microRNA is 3'-polyadenylated prior to step (a) and the DD has a poly-T region adjacent to the DR'.
- 5 5. The method of claim 1, wherein the nuclease has 3'-to-5' activity; or has 5'-to-3' activity; or digests 5' flaps.
6. The method of claim 1, wherein the method further comprises any of
  - (c) inactivating the nuclease;
  - 10 (d) amplifying the ligation products; orwherein the DR and UR are separated by at least one nucleotide,
  - (b0) extending the DR' using the sample as template.
7. The method of claim 1, wherein at least one of the downstream detector (DD) or the  
15 upstream detector (UD) is configured by having a second complementary region (DR2' or UR2'), whereby the DR2' or UR2' can specifically hybridize to a DR2 or UR2 of the target sequence.
8. The method of claim 7, wherein the DD has a first amplification region (P1)  
20 downstream of the DR', and the UD has a second amplification region (P2') upstream of the UR'.
9. The method of claim 7, wherein an attenuator oligonucleotide that has a portion of  
DR2, DR2', UR2, or UR2' is provided in step (a) for a high-abundance target  
25 sequence (HAT).
10. The method of claim 9, wherein the attenuator (i) has a portion of DR2 and a portion of DR or (ii) has a portion of UR and a portion of UR2.

11. The method of claim 1, wherein (i) the DD is configured by a 5'-exonuclease blocking group, (ii) the UD is configured by a 3'-exonuclease blocking group, or both (i) and (ii).
- 5
12. The method of claim 11, wherein the blocking group is a solid phase.
13. The method of claim 1, wherein DD is configured by having a 5'-terminal DR2' region that is hybridized to a DR2 protector oligo; or wherein the UD is  
10 configured by having a 3'-terminal UR2' sequence that is hybridized to a UR2 protector oligo.
14. The method of claim 1, wherein at least one detector oligo is configured by forming a circularized structure when hybridized to the target sequence.
- 15
15. A kit for performing the method of claim 1, comprising detectors for target sequences and at least one nuclease; optionally further comprising a ligase; or optionally further comprising a component for a reaction buffer for an enzyme in the kit, wherein the component is suitable for addition for the enzyme reaction in the same container as  
20 the preceding step of the method.

Fig. 1





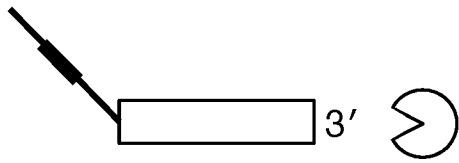


Fig. 2e

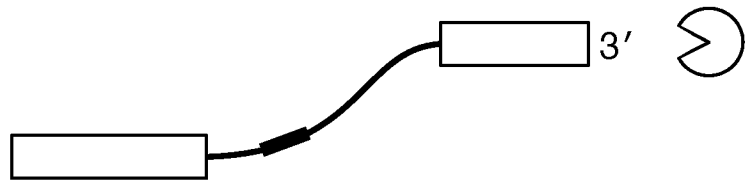


Fig. 2f

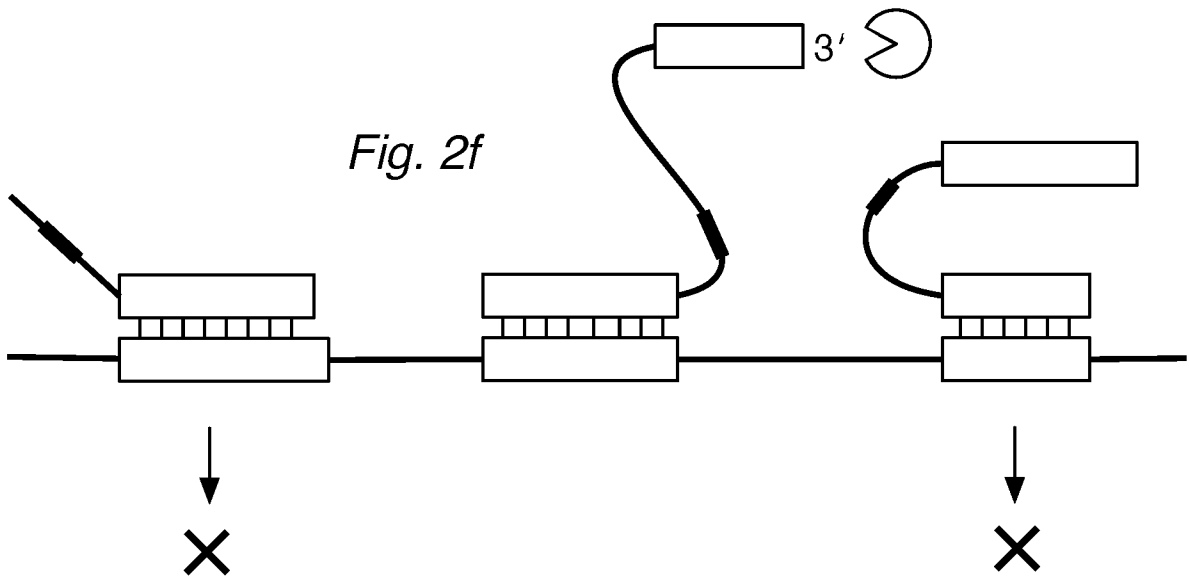


Fig. 2g

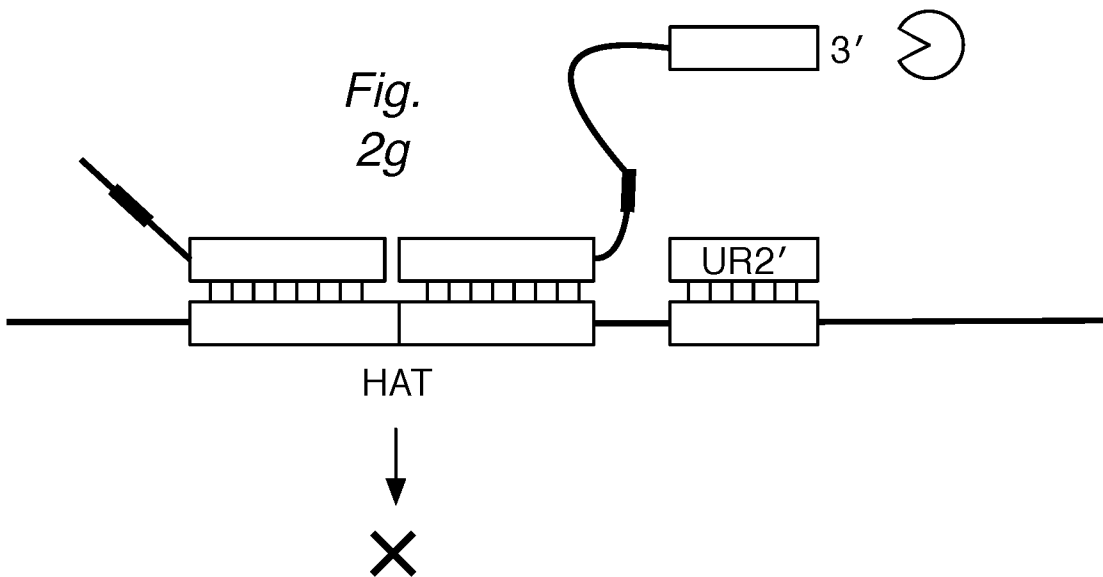




Fig. 3a

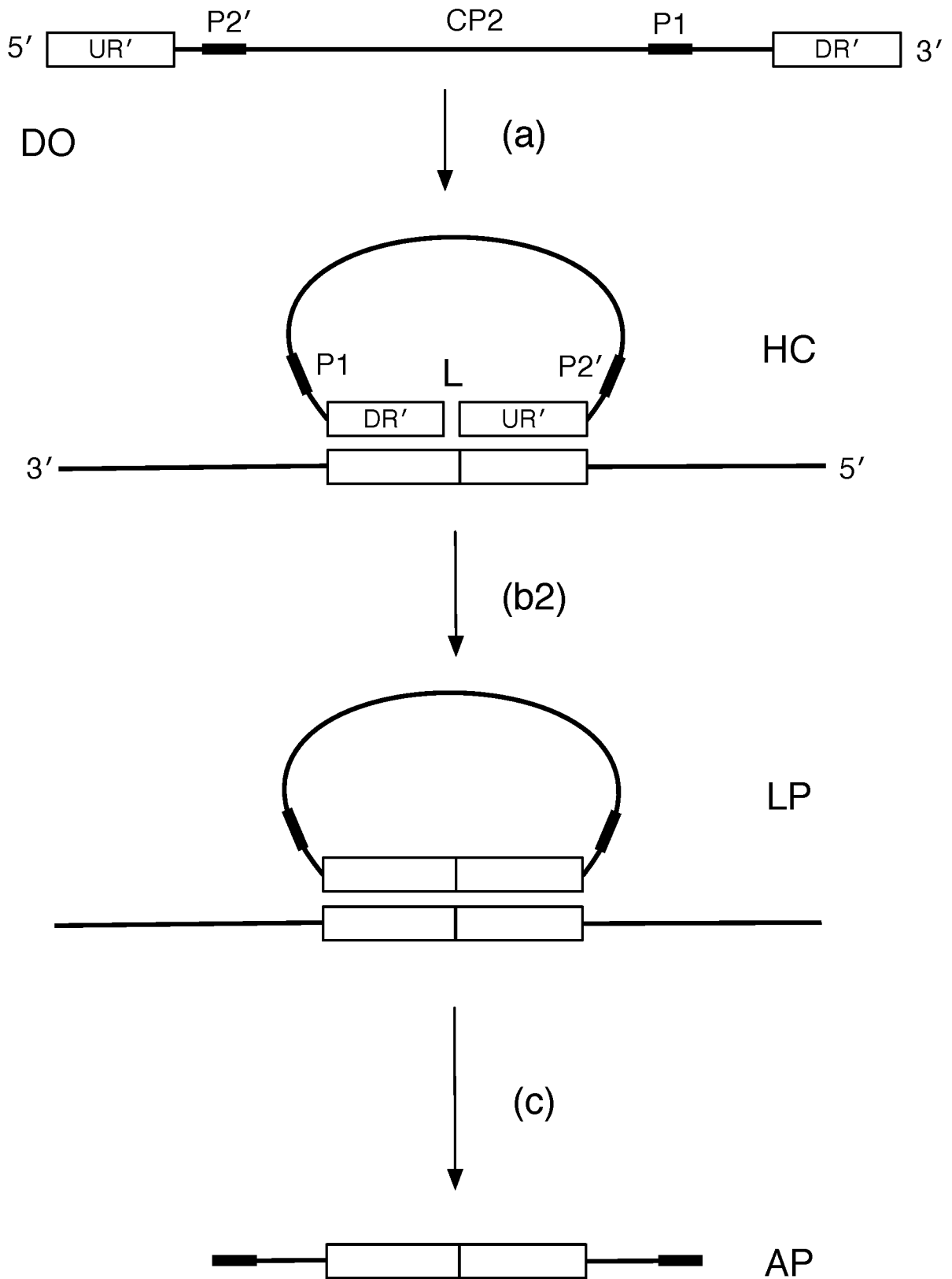


Fig. 3d

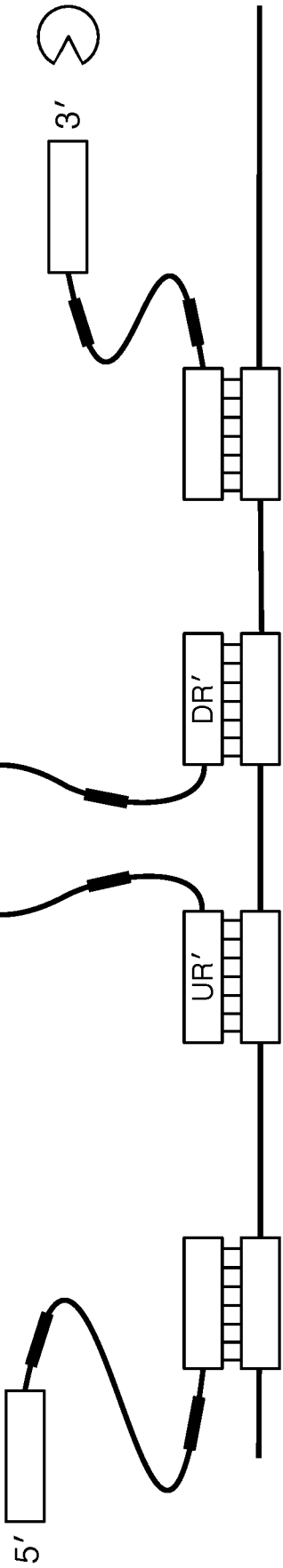
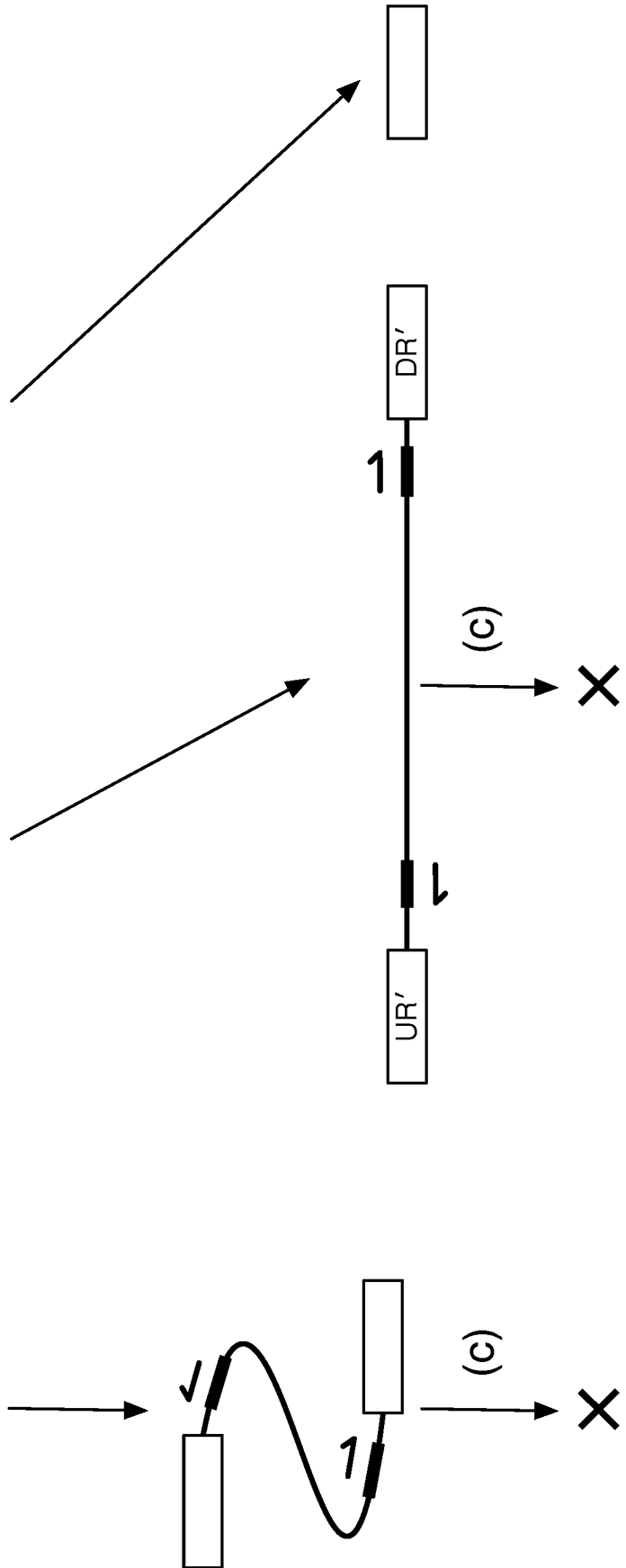


Fig. 3c

Fig. 3b



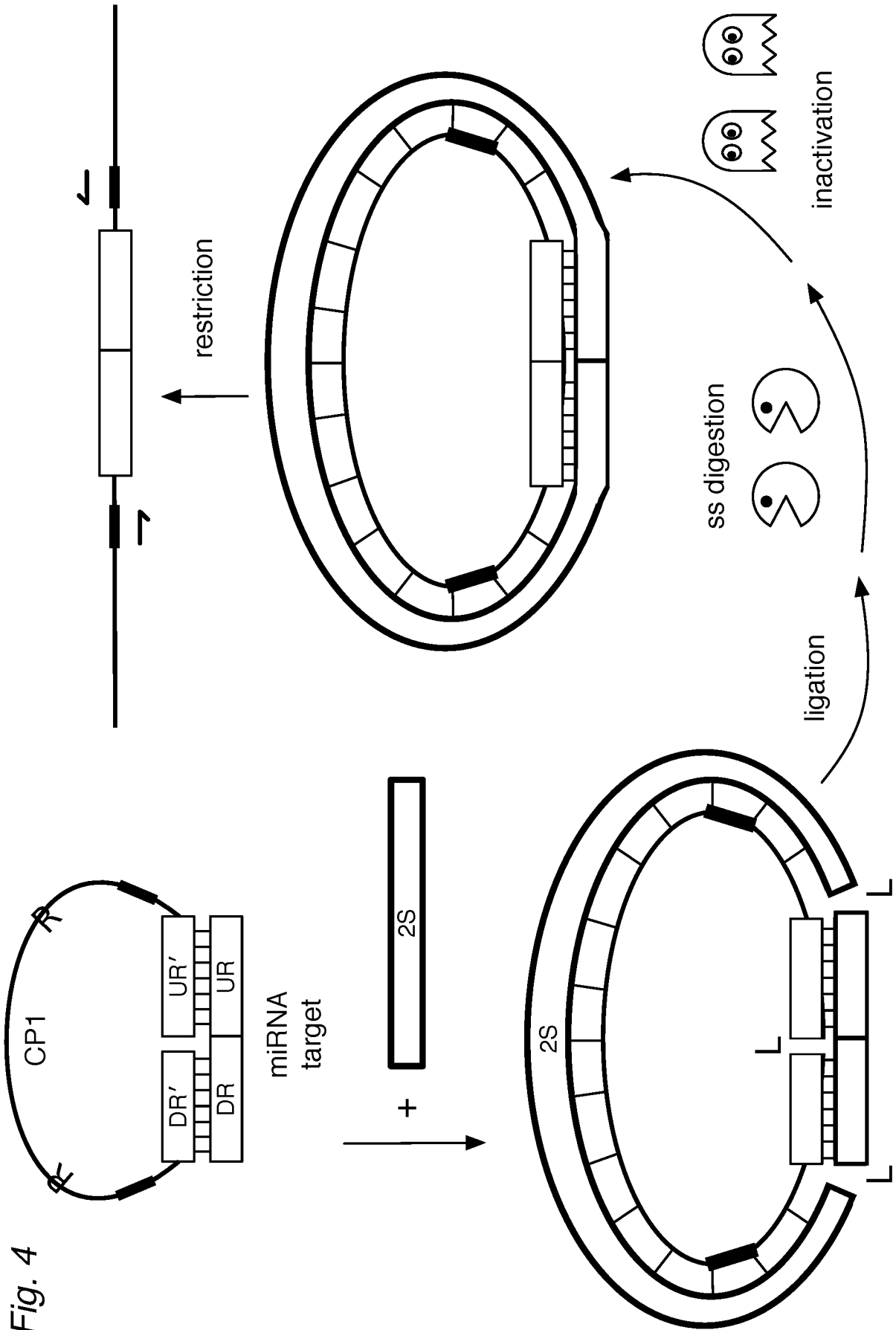
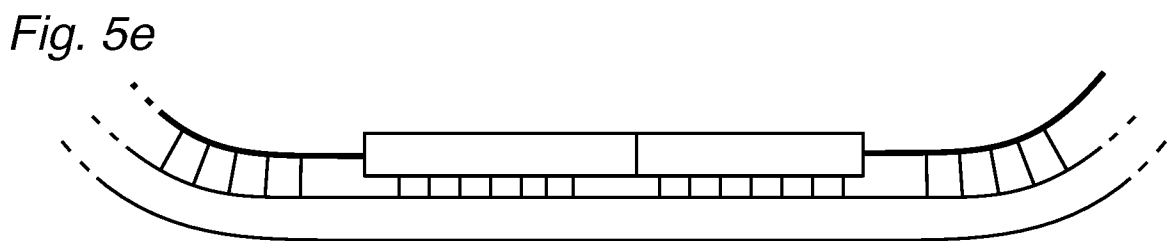
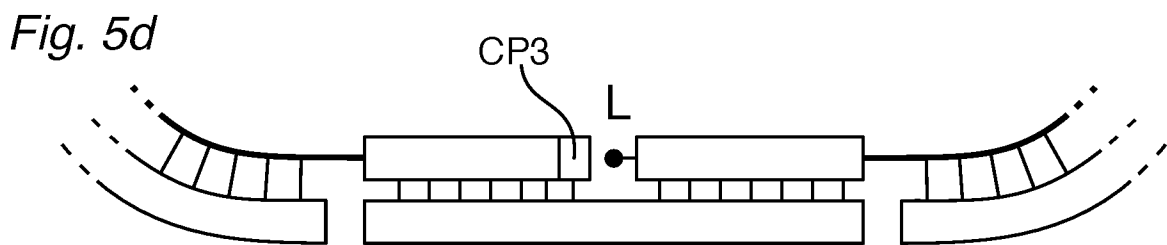
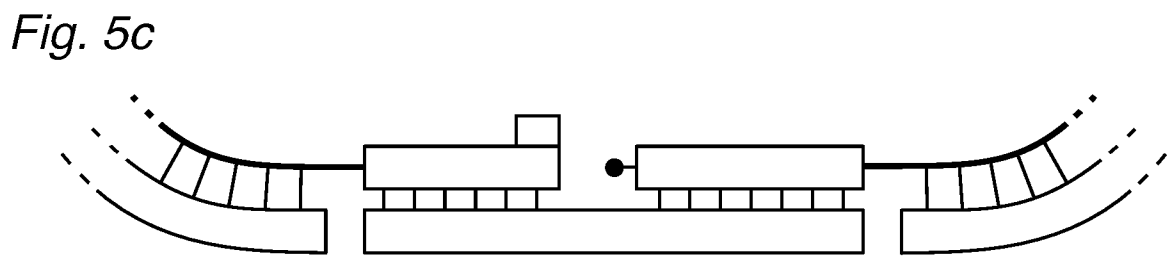
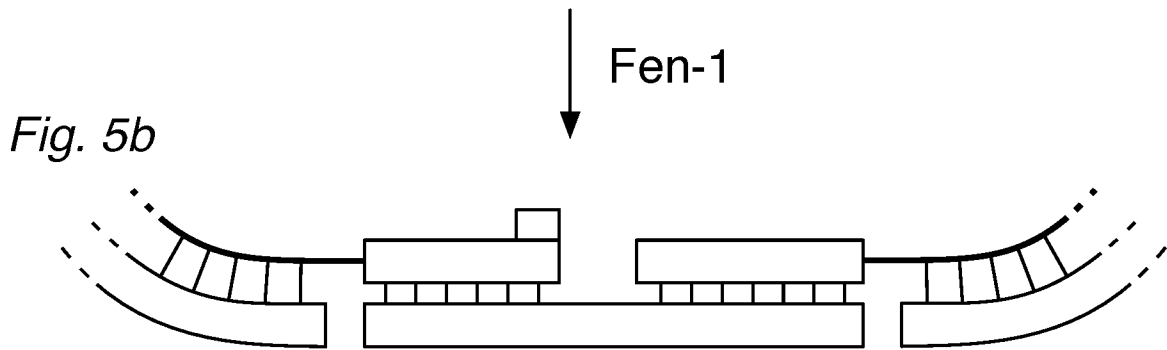
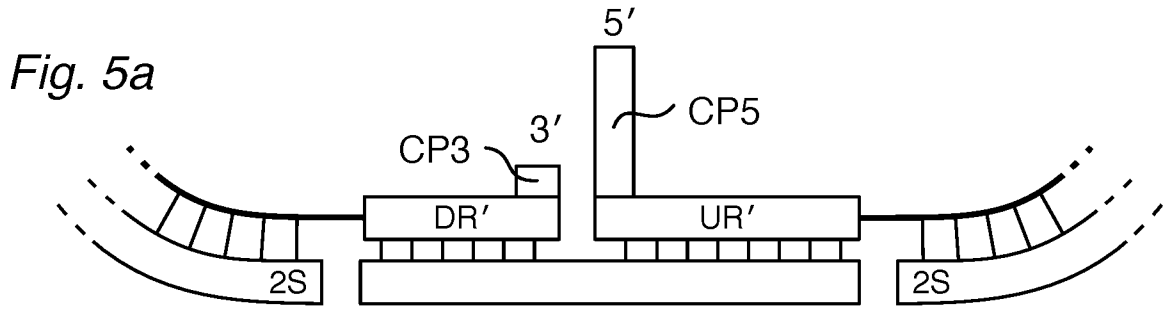


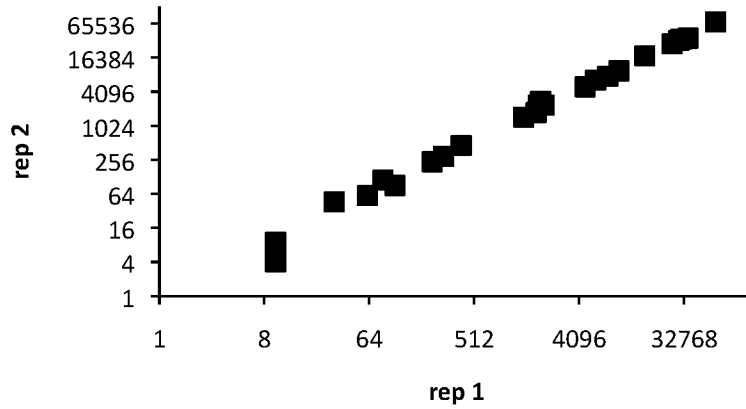
Fig. 4



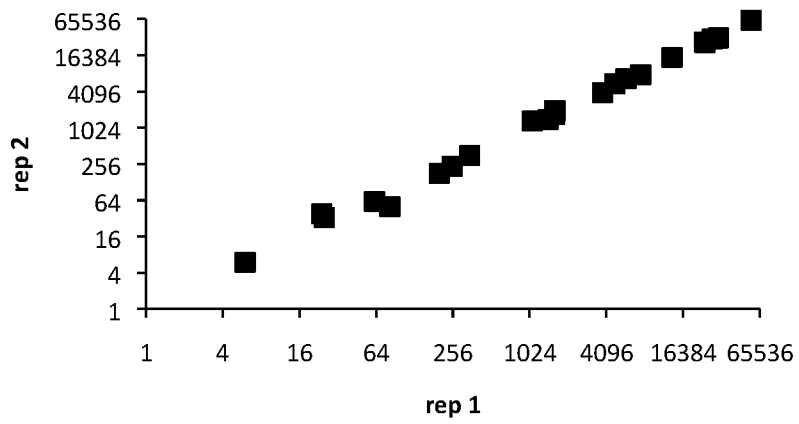
*Figure 6a*

gene	DR' and UR' of detectors	
ACTB_1	AGGTGTGCACTTTTATTCAACTGGTCTCAAGTCAGTGTACAGGTAAGCCC	SEQ ID NO: 33
BAD_3	CGAGGAAGTCCCTTCTTAAAGGAGTCCACAAACTCGTCACTCATCCTCCG	SEQ ID NO: 34
BCAT2_3	CTTGTCATTCCATTCCACCATCAGCATGTGGTCGGTAAATGTCTTCCCAA	SEQ ID NO: 35
BMP4_1	GTGTATATCTGTCTATCCTCAAGGACTGCCTGATCTCAGCGGCACCCACA	SEQ ID NO: 36
BRCA1_2	TGCCCAAGGACTATTCTGACTTTAAGTCACATAATCGATCCCAAGCACTC	SEQ ID NO: 37
BRCA2_1	TTCTTCCGTACTGGCCTGGGAACTCTCCTGTTCTTTGATCAGAGATGTAG	SEQ ID NO: 38
CDH1_1	TATTCTCGGTTTTCTGTGCACACCTGGAATTGGGCAAATGTGTTTCAGCTC	SEQ ID NO: 39
EGF_3	TTTTCCATCCCCAGCAAATCCTTTCAAACACTGACATGTGGCATCCTCTC	SEQ ID NO: 40
EGFR_1	AGCAAAAGGAACATTTTGTATGTGTGTGTGACTGAACATAACTGTAGGCT	SEQ ID NO: 41
ESR1_1	GCGACAAAACCGAGTCACATCAGTAATAGTATGCATCGGCAAAAGGGCAT	SEQ ID NO: 42
GAPDH_3	CCATTGATGACAAGCTTCCCGTTCTCAGCCTTGACGGTGCCATGGAATTT	SEQ ID NO: 43
GATA3_3	CTCTCTGAAACCCTCAACGGCAACTGGTGAACGGTAAACTGATTGCCCA	SEQ ID NO: 44
ICAM1_2	CTGGCATCCGTCAGGAAGTGTGGGCCTTTGTGTTTTGATGCTACACATGT	SEQ ID NO: 45
IGF2_3	CCCTGCCCCAGCCTGATGGAACCCTCTGTTTACACACCTGCTAGCCCCTT	SEQ ID NO: 46
KIT_3	TGAGCCTATTCTCACAGATCTCCTTTTGTGCGCCTTGTTGGGACAACAT	SEQ ID NO: 47
KRT19_2	TCCGTTTCTGCCAGTGTGTCTTCCAAGGCAGCTTTTCATGCTCAGCTGTGA	SEQ ID NO: 48
LAMP1_2	CTTTGAATATATTGACTGAAAACGUCTTCGTGACACGGACGTGCTCCTCC	SEQ ID NO: 49
MUC1_3	ATCGAGAGGCTGCTTCCGTTTTATACTGATTGAACTGTGTCTCCACGTCCG	SEQ ID NO: 50
NFKBIA_1	TACATTATGTACACCATTTACAGGAGGGTAACACAAACCTTGACAGGTAG	SEQ ID NO: 51
PTEN_1	TGCATAGCATTTACACACAGAGCCACTGCTGCACAGCACAAAGAGTATCTG	SEQ ID NO: 52
TFF1_1	AAGCGTGTCTGAGGTGTCCGGTGGAGGTGGCAGCCGAGCTCTGGGACTAA	SEQ ID NO: 53
TIMP1_2	GCATCCCCTAAGGCTTGGAAACCCTUTATAACATCTTGGTCATCTTGATCTC	SEQ ID NO: 54
VEGFA_3	GATGGTGTGGTGGCGGCAGCGTGGTTTCTGTATCGATCGTTCTGTATCAG	SEQ ID NO: 55
WNT1_1	AGGAGCCGCTAATAGCTACAGTGAAGGAAATACTGATTCCAGGAGGCAA	SEQ ID NO: 56

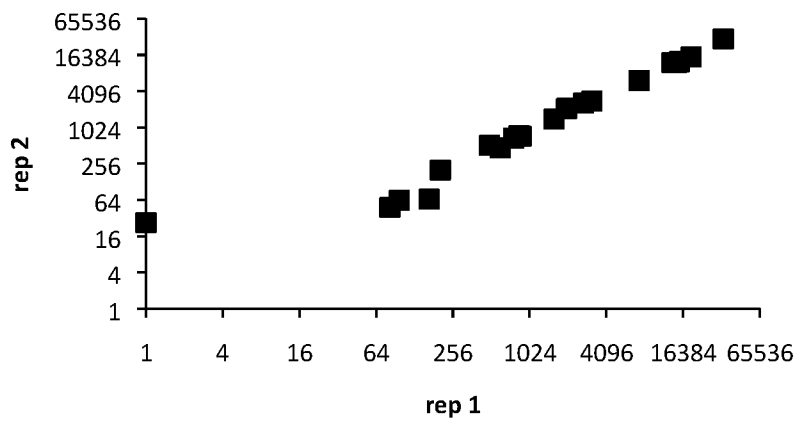
**Fig. 6b: Anchored probes, 10 ng input**



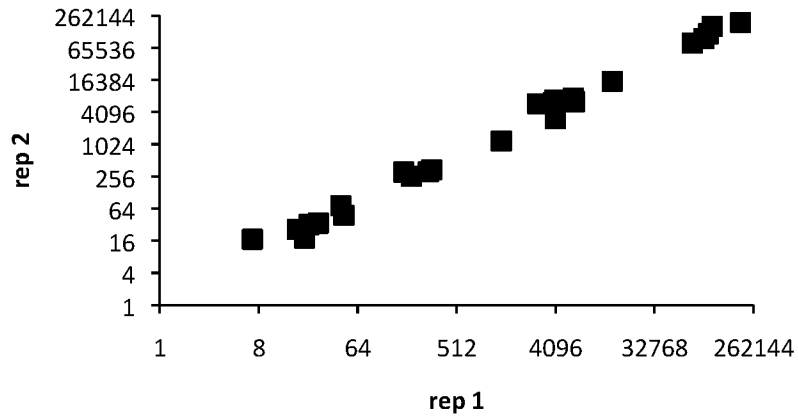
**Fig. 6c: Anchored probes, 1 ng input**



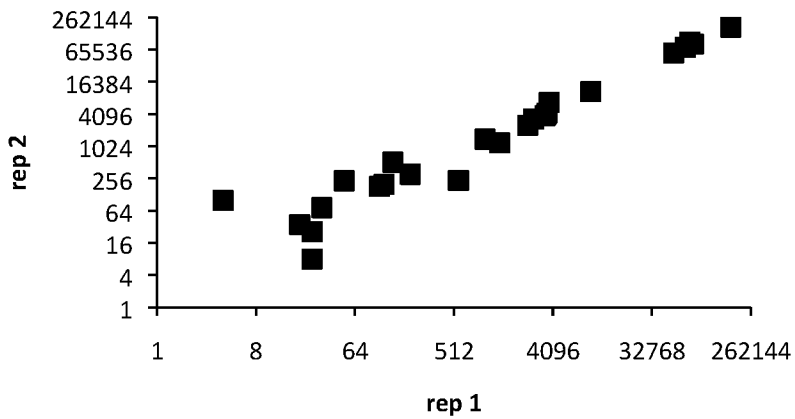
**Fig. 6d: Anchored probes, 0.1 ng input**



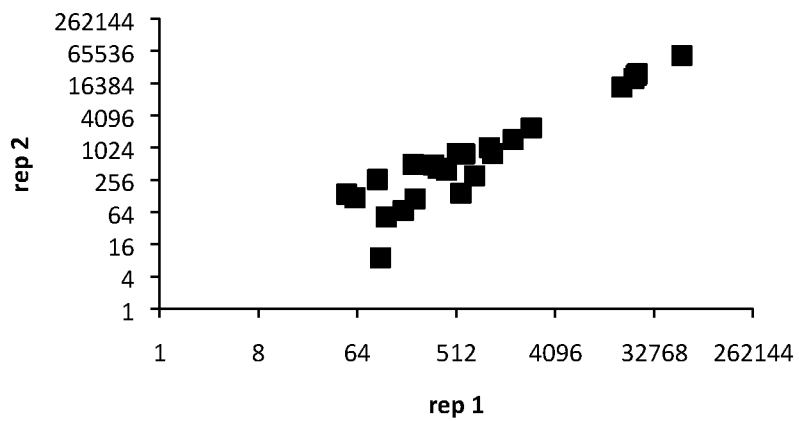
**Fig. 6e: Circular probe, 10 ng input**



**Fig. 6f: Circular probe, 1 ng input**



**Fig. 6g: Circular probe, 0.1 ng input**



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US16/14999

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12P 19/34; C12Q 1/68 (2016.01)

CPC - C12Q 1/6816, 1/6818, 1/6827

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(8) Classification(s): C12P 19/34; C12Q 1/68 (2016.01)

CPC Classification(s): C12Q 1/6816, 1/6818, 1/6827

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PatSeer (US, EP, WO, JP, DE, GB, CN, FR, KR, ES, AU, IN, CA, INPADOC Data); The Lens; Google/Google Scholar;

NCBI/BLAST/PubMed; EBSCO;

Search terms used: nucleic acid, target, detect, oligo, probe, adapter, upstream, downstream, domain, region, single strand, nuclease, amplify

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2006/0246475 A1 (PETERSON, P et al.) November 2, 2006; paragraphs [0006], [0018], [0065]-[0068], [0103], [0108], [0114], [0186]-[0187], [0192], [0222], [0278], [0280]	1-15
Y	US 2005/0026166 A1 (BI, W) February 3, 2005; paragraph [0077]	1-15
Y	CA 2914367 A1 (ATHENA DIAGNOSTICS, INC., et al.) December 11, 2014; paragraph [0011]; SEQ ID NO: 115)	7-10

 Further documents are listed in the continuation of Box C. See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search

28 March 2016 (28.03.2016)

Date of mailing of the international search report

14 APR 2016

Name and mailing address of the ISA/

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774