

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
19 September 2002 (19.09.2002)

PCT

(10) International Publication Number
WO 02/072889 A2

- (51) International Patent Classification⁷: **C12Q 1/68**
- (21) International Application Number: PCT/US02/00624
- (22) International Filing Date: 11 January 2002 (11.01.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/261,526 12 January 2001 (12.01.2001) US
- (71) Applicant: **PE CORPORATION (NY)** [US/US]; 850
Lincoln Centre Drive, Foster City, CA 94404 (US).
- (72) Inventors: **FEREA, Tracy, L.**; 595 Maybell Avenue, Palo
Alto, CA 94306 (US). **SCHROTH, Gary, P.**; 100 Monas-
terio Court, San Ramon, CA 94583 (US).
- (74) Agents: **GARRETT, Arthur, S.** et al.; Finnegan, Hen-
derson, Farabow, Garrett & Dunner, L.L.P., 1300 I Street,
N.W., Washington, DC 20005-3315 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN,
YU, ZA, ZM, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).

Published:

— without international search report and to be republished
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.



WO 02/072889 A2

(54) Title: METHODS AND COMPOSITIONS FOR MICROARRAY CONTROL

(57) Abstract: Assays for detecting target nucleic acids in a sample in hybridisation assays using labelled control probes, in order to accurately quantitate and compare signals obtained in different hybridisation experiments.

METHODS AND COMPOSITIONS FOR MICROARRAY CONTROL

Related Application Information

This application claims the filing date benefit of U.S. Provisional Patent Application Serial No. 60/261,526, filed January 12, 2001, incorporated by reference in its entirety for any purpose.

Background of the Invention

This invention relates to assays for detecting or determining target molecules in a sample. In certain embodiments, it relates to nucleic acid arrays and controls used in such arrays.

Substrate-bound oligonucleotide arrays, also known as microarrays, enable one to test the hybridization of different nucleic acid sequences in a sample to many different oligonucleotide probes. These arrays can be composed of hundreds of thousands of probes deposited or synthesized within specific regions, defined as features, on a substrate such as a glass microscope slide or other materials. (See, e.g., Figure 1.) In some procedures, one may use target nucleic acid directly from a sample (as mRNA, for example). In some procedures, one may use target nucleic acid replicated or amplified from a sample (as cDNA, for example). Hybridization assays on such arrays may be used for profiling of gene expression levels, identification of genetic variants of infectious diseases, identification of genetic diseases, or any assay that identifies different nucleic acid sequences.

Target nucleic acid is typically labeled with a detectable marker, such as a fluorescent molecule. Hybridization between a target and a probe is detected by a fluorescent signal at various features within the array. The amount of signal can be dependent on the amount of target available for hybridization, as well as the thermal stability of the probe-target hybrids. Thermal stability is a function of several factors. For example, the length of the hybridizing region, the accuracy of the match in hybridization, the total length of the oligonucleotides, as well as the actual sequence composition (A-T rich regions melt at lower temperatures than G-C rich areas), all factor into the specific melting point (T_m) for a probe-target hybrid.

Currently, quantitation of a signal is often performed by comparing the signal from an experimental target sample with an arbitrary biological reference (or

“control”) sample that competitively hybridizes to the same feature. (See, e.g., Figure 2.) Thus, comparisons are made within the same feature. For many biologically derived samples, it is often difficult to either obtain or determine the proper control for an experiment. Biologically derived controls also have problems with reproducibility. Also, such controls typically do not include a target sequence corresponding to every different probe sequence for each feature on an array. Thus, some features may never have control target sequences that can provide a signal.

With current methods, an absence of any control signal for any given feature may be caused by a lack of a sufficient amount of that specific target in the control, or by a lack of sufficient probe bound to the feature. These two causes, which have profoundly different biological implications, typically cannot be distinguished.

The shape of each feature can also affect the apparent amount of signal read at a given locus. A smaller feature creates a smaller area lit by the signal, and consequently an imaging system may read less signal. Also, it is often important to differentiate between background and actual signal at any given feature. Thus it is important to determine pixels at a given feature that are properly attributed to a positive signal rather than to background.

It can be difficult to quantitate accurately the number of probe molecules attached to a feature. This makes it difficult to accurately quantitate signal obtained from hybridization. Consequently, it is difficult, if not impossible, to compare signals obtained from hybridization to different probes at different features within an array, as well as to compare signals obtained from the same probe on the same array or signals obtained from the same probe on different arrays.

Summary of the Invention

In certain embodiments of the invention, synthetic control elements are provided. In certain embodiments, the controls comprise nucleic acid sequence control probes that are included in each feature. In certain embodiments, such bound control probes hybridize to labeled control target sequences, but do not hybridize to target sequences in the experimental sample. In certain embodiments, the control probe sequence may be attached to or contiguous with an experimental probe. In

certain embodiments, the control probe may be bound to the substrate separate from the experimental probe.

In certain embodiments, a control label is included in each feature of the array, which provides a control signal that indicates the amount of probe attached. In certain embodiments, a control label may be attached directly to at least some of the probes of each feature. In certain embodiments, the control label is provided by a nucleic acid molecule with a label attached, and the nucleic acid molecule is designed such that it does not hybridize to nucleic acids in an experimental sample.

In certain embodiments, the label that is attached to the feature may provide any number of signals, in addition to the signals provided by an experimental target signal and a control target signal.

In certain embodiments, labeled control sequences are provided such that there is a synthetic control sequence corresponding to all of the probes on all of the features.

In any of these embodiments, the experimental probes on any given feature may have the same sequence for hybridization to a specific target sequence in the sample. Also, in any of these embodiments, any given feature may contain more than one different experimental probe sequence for hybridization to different target sequences in the sample.

Kits for quantitating the signal and the attached probe are also provided. In certain embodiments, these kits comprise an array that has been created using the inventive methods. Kits further comprising an array and a synthetic control target are also provided.

Brief Description of the Figures

Figure 1 shows a close-up view of 342 features of a 55K (1in X 0.6 in) array with a fluorescent label.

Figure 2 shows competitive hybridization between experimental target and control target labeled with two different colors.

Figure 3 shows a third signal approach, where a third label is attached to each feature during manufacture of the array. The third signal can provide for feature

identification and quantitation. A competitive hybridization between experimental and control targets using two other signals may be performed with such embodiments. Also, the competitive hybridization may be performed with two experimental targets from two different samples using two other signals.

Figure 4 shows a non-competitive hybridization using synthetic control targets that are complementary to a control probe. In this embodiment, the control probe is attached to the experimental probe.

Figure 5 shows hybridization using labeled probes.

Detailed Description of the Invention

The section headings used herein are for organizational purposes only, and are not to be construed as limiting the subject matter described. All documents cited in this application, including, but not limited to patents, patent applications, articles, books, and treatises, are expressly incorporated by reference in their entirety for any purpose.

Definitions

The term “complement” refers to the compatibility of two sequences and their ability to match and form a hybrid. Thus, a target can be said to be a complement of the probe.

A “probe” is an oligonucleotide that is attached to the substrate of the array, and that is capable of binding to a complementary target sequence. The probe may include Watson-Crick bases or modified bases. Modified bases include, but are not limited to, the AEGIS bases (from Eragen Biosciences), which have been described, e.g., in U.S. Patent Nos. 5,432,272; 5,965,364; and 6,001,983. Additionally, bases may be joined by a natural phosphodiester bond or a different chemical linkage. Different chemical linkages include, but are not limited to, a peptide bond or a Locked Nucleic Acid (LNA) linkage, which is described, e.g., in published PCT applications WO 00/56748; and WO 00/66604.

In certain embodiments, control probes bind to control target sequences, and experimental probes bind to experimental target sequences. In certain embodiments, the control probes and experimental probes are contiguous (e.g., they may be made as one oligonucleotide or they may be attached to one another in any suitable manner).

“Target” nucleic acids are those nucleic acids capable of hybridizing to probes.

An “experimental target sample” is one or more nucleic acid molecules from a test sample that are used in an assay. The experimental target sequences are those that may be identified or quantified after they hybridize to experimental probes on the array. In certain embodiments, the experimental target sample may be extracted directly from the sample, e.g., as RNA or DNA. In certain embodiments, the sample may be replicated from the sample nucleic acid, e.g., such as cDNA which has been reverse transcribed from sample mRNA. Finally, the experimental target sample may also be amplified from the sample nucleic acid.

A “control target” is a nucleic acid that can hybridize to probes to give a reference signal. In certain embodiments, such control targets can be used as a reference against which the signal from the experimental target may be compared. The control target may compete, or may be non-competitive with, the experimental target. Thus, a single control target molecule may or may not compete with an experimental target molecule for binding to a single probe molecule.

The term “feature” refers to a single location deposited on an array which is read as a single data point. In certain embodiments, a single feature may have probes that all have the same sequence. In certain embodiments, a single feature may include probes with different sequences.

The arrays in certain embodiments of the invention can be used for any identification or separation of nucleic acid sequences. Examples include, but are not limited to, nucleic acid expression analysis, identification of tagged molecules in a sample, and separation of tagged molecules out of a sample.

Reliable Controls

Certain embodiments of the present inventions are directed to methods, compositions, and kits which provide quality control for the manufacture of the arrays. In certain embodiments, the invention can provide quantitation of signal to facilitate comparison between features within an array, comparison with features on other arrays, and comparison of multiple features on multiple arrays. Such

embodiments provide reliable and consistent control signals against which the experimental signal can be compared.

The manufacture of high density nucleic acid arrays, and methods of their use in diagnostic assays have been described in U.S. Patent Nos. 5,445,934; 5,552,270; 5,837,832; 6,040,138; and 6,045,996; and published PCT applications WO 00/39345 and WO 00/47767.

During manufacture, errors may occur during the depositing of probes on features within the array. Thus, there may be no probe deposited at a particular feature, or there may be variation in the amount of probe deposited at different features within an array or variation at counterpart features in different arrays. Such variation in probe deposition (including failure of probe attachment) may occur due to many factors. For example, the reactivity of the surfaces may vary from feature to feature and/or from array to array, different elements that are used for depositing probe at different features or on different arrays may create variation in the amount and distribution of probe, and differing environmental conditions, such as humidity, can impact probe deposition. Since the attachment typically involves a chemical binding reaction with the probe or with linkers, variations in such reactions, or the absence of active reagents, can create variations in probe deposition. Variation may result during any method for attaching nucleic acid on a substrate, e.g., during in situ synthesis of probe on a substrate. Variation may often result when arrays are made in different facilities, but also may result within the same facility.

An aim of certain embodiments of the invention is to account for such variation and to provide more accurate determination of amounts of sequences within an experimental sample. Thus, according to certain embodiments, the present invention allows one to determine whether sufficient probe is actually attached to an individual feature. In such embodiments, if no signal is detected for a given feature, the user can typically conclude that insufficient probe was bound to the substrate. Without such appropriate controls, when one uses the array, such a lack of signal may indicate that there was not sufficient complementary control target or experimental target in the sample, but it could also indicate that there was insufficient probe attached to the feature. The user would not be able to make a conclusion one way or

the other. Thus, such embodiments of the present invention provide an important control for the user.

Also, certain embodiments provide appropriate controls for manufacturing arrays. In certain embodiments, one can determine whether probe has attached to each feature without running a hybridization reaction. In certain embodiments, the controls can be used in hybridization reactions to test batches of arrays to provide quality control for appropriate probe deposition.

Also, according to certain embodiments, the invention provides more consistent controls for comparing experimental signals, which allows one to obtain an accurate ratio of experimental signal to control signal. That ratio allows one to more accurately compare that feature to other features on that array, or to features on other arrays when one employs the same controls on such other arrays.

Also, with different features on an array or on features of different arrays, variation in intensity of signal may be due to variation in the amount of experimental target in a sample, or may be due to variation in the amount of probe in counterpart features. In certain embodiments, the present invention allows one to determine whether such variation is due to variation in the amount of probe, since the control signal is not dependant on the amount of experimental or control target in the sample. Such embodiments include more consistent controls, which allows one to obtain a more accurate ratio of experimental signal to control signal. That ratio allows one to more accurately compare that feature to other features on that array, or to features on other arrays when one employs the same controls on such other arrays.

Also, when microarrays are placed in an optical reader, small misalignments, e.g., caused by misplacement of the array in the reader or by misalignment during array manufacture, may result in errors in identifying the source of a signal. In some known arrays, some features in the array are used to spot a label without probe in order to serve as a "landmark" in aligning the array. This results in fewer features available for experiments, and does not provide controls for the amount of probe deposited on the other features.

According to certain embodiments of the invention, a "landmark" is provided in features along with experimental probe. According to certain embodiments,

deposition of control label in easily identifiable patterns allows the features to provide a "landmark" function for aligning the array without sacrificing the number of features used for the assay. The landmark function also allows for easier identification of specific features within the array.

In some methods of array manufacture, features are deposited in an array on a polymeric substrate. According to certain embodiments, one can use polymeric film and methods of affixing nucleic acids such as those disclosed in published PCT application WO 99/53319. After attachment of the probes, the polymeric film is heated. In certain embodiments, it shrinks about twenty-five fold, to about four percent of its original size. After shrinking, the plastic substrate typically has folds on the surface approximately 10 microns across. The features typically are approximately 40 microns across. Such folding can create an irregular and uneven focal plane, and an indefinite depth of field.

Having a control signal at each feature typically allows one to correct for irregularities in the shape, size, and intensity of the feature, as well as in the focal plane and the depth of field. According to certain embodiments, placing the array on a larger surface which is shrunk allows for finer detail in depositing desired shapes for the feature, and greater regularity in probe density. Because one is initially depositing a larger feature that is later reduced in size, feature landmarks typically become easier to shape to provide useful landmarks for aligning arrays.

Also, according to certain embodiments, features are outlined with control signal. This allows one to scan features that are defined by the control signal, and disregard areas with no control signal. Areas without control signal would have no attached experimental probe. Thus, one can determine what part of the feature is background and that background should not be included in the quantitation of the features. This would make the reading of any experimental signal accurate.

In certain embodiments the identifiable pattern formed is a pattern of pixels, such as those read by an optical scanner. Detection of a control signal within a pixel indicates that probe is present within that pixel. Those pixels with no control signal are then known to be background. One can then easily distinguish the pixels which are part of the feature from the background, and scan for experimental signal only in those pixels which are part of the feature.

Certain embodiments of the invention are directed to software that is used for analysis of the controls. For example, software can be used for quantitation and comparison of the various signals from different features and/or from different arrays.

Examples of Controls

In the following, methods of providing a control system for certain embodiments are described. In all of the embodiments, the control signal can be used to confirm that experimental probe is present in a feature and can be used to help quantitate the experimental signal. In all the embodiments in which a control signal is present at a feature, that control signal can also be used to define the shape of the feature. In such embodiments where a control signal is present in a feature, the outline of the feature may be used to positively identify where experimental probe is bound, and distinguish that area from background where no experimental probe is bound.

In any of the embodiments of this invention, the probes and the labeled molecules can be attached to the substrate in any manner. In certain embodiments, the probes and labeled molecules are attached covalently to the substrate. In certain embodiments, the probes and labeled molecules are attached by other methods, including but not limited to UV cross-linking, electrostatic attachment, polylysine coating of the substrate, hybridization to other nucleic acids on the substrate, and in situ synthesis of the nucleic acid on the substrate. In certain embodiments, the probes and labeled molecules are attached with a linker molecule. Certain embodiments include, but are not limited to, polyethylene glycol linker molecules, peptide linker molecules, or C6 linker molecules.

In certain embodiments, the control label is attached to the feature along with the probe. The label may be conjugated onto the experimental probe, or onto a different molecule. In certain of these embodiments, when the label is attached to a different molecule, the labeled molecule can be attached to the feature using the same chemical reaction or linker that is used to attach the probes. In certain embodiments, label is conjugated to a subset of the experimental probes, such that a specific percent of the probe molecules bound to the feature are labeled. One can control the ratios of the unlabeled probe and the labeled probe in the sample, thus the signal from the

labeled probe can be used to calculate the amount of total experimental probe bound to a feature.

In certain embodiments, one may want to have a control for determining whether hybridization occurs. In certain embodiments, a feature is deposited on an array with a control probe that hybridizes to a control target known to be in the sample. Detection of the complementary control target oligonucleotide at the control probe feature indicates that hybridization is occurring.

In certain embodiments, each feature of the array contains a nucleic acid control probe which is not complementary to any experimental target sequence. In such embodiments, one uses a control target that is complementary to the control probe. The control target does not compete for binding with the experimental target. In certain embodiments, one uses control probes that all have the same nucleic acid sequence and control targets that all have the same sequence that is complementary to the control probe sequence. Thus, in these embodiments, a different control sequence need not be included for each different array feature. This reduces the number of synthesis reactions used to make the control targets for the assay.

To decrease the chance of cross hybridization with experimental targets, in certain embodiments, the control probes and control targets may include non-Watson-Crick bases. Such bases typically would not be included in experimental targets from a biological sample, and typically would not hybridize with the naturally occurring Watson-Crick bases in the experimental probes and experimental targets. Synthetic non-Watson-Crick bases, such as the AEGIS bases, are described, e.g., in U.S. Patent Nos. 5,432,272; 5,965,364; and 6,001,983, available from Eragen Biosciences, Inc.

In certain embodiments, the non-competitive control probe is attached to the experimental probe (e.g., the control probe and target probe can be synthesized as one oligonucleotide). The two contiguous probes are then attached to the substrate. The attachment to the substrate may occur through the control probe, the experimental probe, or a linker that is attached to one of the probes. This method has an added advantage that the stoichiometry of experimental and control probes within a feature is the same. Thus, in these embodiments, by determining signal from the control target that binds to the control probe one can easily determine not only whether experimental probe has bound to the substrate, but also the amount of such binding.

In other embodiments, the non-competitive control probe is not contiguous with the experimental probe, and is bound to the substrate with the same substrate binding reaction as the experimental probe. These embodiments provide a control for determining the amount of experimental probe that is attached to the feature. Because the control and experimental probes bind to the substrate by the same chemical reaction, the amount of control probe that is detected in a feature should be representative of the amount of experimental probe that is also attached to that feature. Thus, failure to attach experimental probe or variations in the amount of experimental probe bound to the substrate can be detected by observing the amount of control signal obtained after hybridization of the control target to the control probe.

In certain of these embodiments, one prepares a predetermined ratio of control probe to experimental probe. Thus, a reliable and consistent control signal can be added to the array quickly.

In certain embodiments, one employs labeled experimental targets (a first signal), labeled control targets (second signal), and control labels that are attached to the features during array manufacture (a third signal). (See Figure 3.) In these embodiments, the third signal provides a control for accurately determining the amount and placement of experimental probe that is bound to the feature. The amount of a given control target in traditional biological derived control samples is typically not known. Thus, it is possible that a given control target sequence is not present in the control sample. Thus, without the third signal, absence of control signal at a feature cannot be positively attributed to lack of control target sequence in the sample or to the lack of experimental probe on the feature. Also, since the amount of control target typically is not known, one typically cannot accurately determine that differences in signal intensity from feature to feature is due to different levels of target in the sample or is due to different levels of probe bound to the substrate.

In these embodiments, the competitive hybridization is carried out with the experimental target (which provides a first signal if there is binding) and control target (which provides a second signal if there is binding). If one observes second signal at a feature, the control target signal can affirm that hybridization can occur with a given probe.

In certain embodiments, one can use three signals but replace the control target (with the second label) with a different experimental target that has the second label. Thus, within a single feature, one can determine whether there is complementary experimental target in two different samples. In certain embodiments, one could obviously increase the number of different experimental target samples with different labels. In certain embodiments, one could employ more than one experimental target sample along with the control target and the control attached to the feature. In such embodiments, one could use different labels for each experimental target sample as well as for the control target and the control attached to the feature.

In some instances, more than one experimental probe sequence may be employed for determining the amount of a target in a sample. For example, this may be useful in instances in which variations in splicing result in different mRNA transcripts for the same gene, which transcripts have different overall sequences. In such cases, it may be desirable to have different experimental probe sequences complementary to different portions of the gene located on the same feature. In certain embodiments of the invention, several experimental probes are created that are complementary to different regions of a transcript. These different experimental probes are deposited on the same feature. This would allow experimental target molecules representing a given gene a greater opportunity to hybridize to the feature. Any number of different experimental probes may be included in the same feature. Consequently, the experimental signal could be a more accurate indication of the levels of expression of a given gene in the sample, and only one feature is used to accomplish this result rather than multiple features. Of course these multiple experimental probes can be directed to other nucleic acids, such as intergenic regions, introns, etc.

Synthetic Control Targets

Reliable and consistent controls may also be provided by creating synthetic control targets. Such targets typically would not require any special array design, and could be used with arrays from different sources or from different manufacturers.

Under current methods, control targets are typically derived from biological samples. In an attempt to obtain a control target molecule which hybridizes to each

feature, researchers have used pools of nucleic acid from several different tissues, or several different samples. This typically does not give reliable or consistent results.

According to certain embodiments of the invention, a synthetic control target sample or pool can be produced that includes known target sequences complementary to any number of features of an array. The amounts of such target sequences in such synthetic control target samples can also be known. Such pools of oligonucleotide targets provide a consistent control signal for the entire array and for different arrays.

According to certain embodiments, the synthetic control target sample can be used to quantify the amount of probe in the features by comparing the signal generated from the control targets to the signal generated from the experimental targets. In such embodiments, since one knows that there is control target that will hybridize to each probe that is supposed to be included on the array, one will always obtain control target signal if sufficient probe is attached to the feature. Thus, a lack of such control target signal indicates that insufficient probe is attached to the feature. Also, since one knows the quantity of control target, one can accurately quantify the amount of experimental target in a sample.

In certain embodiments of the invention, premade sets of synthetic control targets can be made to correspond to different sample types. For example, synthetic control targets can be compiled for known expressed transcripts from whole blood. This set of control targets could be used for any array designed to be used with experimental samples from whole blood. Similarly, synthetic control targets could be composed to correspond to an entire genome. Such a control target pool could be applicable to any diagnostic array designed for human samples.

In certain embodiments, the control targets are made from a set of oligonucleotides which contain a number of random nucleotide bases. In certain embodiments, the control targets are from about 6 to 30 bases, 8 to 15 bases, 10 to 12, or 10 bases in length. By using a number of random nucleotides, one increases the statistical likelihood that there would be some control target molecule which would bind to any given probe. This statistical likelihood may also be increased by using fewer than the typical 25 to 50 base length target oligonucleotides along with the random nucleotides. One may also accomplish an increased likelihood that control target molecule will bind to any given probe by increasing the concentration of the

control target molecules. In these embodiments, the pool of control target oligonucleotides may be designed without knowing the particular probe sequences.

With shorter oligonucleotides, the melting temperature (T_m) is lower, which causes such short oligonucleotides to bind to their complements less efficiently. This possible problem typically can be solved by altering the chemical bonds between the bases. Nucleotide bases joined by peptide bonds, instead of the natural phosphodiester bonds, have higher melting temperatures. In certain embodiments, the control target is composed of short oligonucleotides of nine or ten bases in length, joined by peptide bonds or LNA linkages, such as those described in published PCT applications WO 00/56748; and WO 00/66604. Such targets have the T_m of much longer oligonucleotides, allowing them to withstand the stringent conditions typically used for hybridization assays.

Examples

The following prophetic examples are offered to more fully illustrate the invention, but do not limit the scope thereof.

Example 1

In this example, a standard nucleic acid array is created with a predetermined ratio of labeled control probe to unlabeled experimental probe. (See Figure 5.) The total amount of probe bound to each feature is determined by examining the signals from the labels on the control probes, and calculating the total amount of probe based on the predetermined ratio. One can use this information to more confidently determine the relative amount of experimental target in a sample. Specifically, knowing the relative amount of experimental probe at two different features, one can more confidently determine whether variations in experimental signal intensity at different features is resulting from an actual difference of the level of experimental targets in the sample or is resulting from variations in the amount of experimental probe deposited at different features. This comparison may be accomplished by dividing the experimental signal by the control signal at each feature.

For example, assume that the experimental signal at a feature corresponding to Gene A in a first array is measured at 1000 arbitrary units. Also, assume that the control signal at the same feature is 250 units. This provides a ratio of 4:1 of

experimental over control signal. Next, for a different experimental sample, assume that the signal at a feature for Gene A for a different experimental sample in a second array is measured at 400 units, and that the control signal at that feature on the second array is measured at 100 units. This also provides the same 4:1 ratio. Although the experimental signal at the Gene A feature on the second array is markedly less than the signal on the first array, the correction for the control signal indicates that the transcript in both samples appears to be expressed at the same level. This method may also be used to compare two different target sequences at different features in the same array.

Example 2

In this example, at least some of the experimental probes for each feature include an attached or contiguous control oligonucleotide probe. (See Figure 4.) The control probe is designed to minimize cross-hybridization with any of the intended experimental target molecules. The control oligonucleotide sequence can be identical in each feature.

During hybridization, one uses a control target, which complements the control probe oligonucleotide sequence. The control target hybridizes to the control probe. The control target does not compete with the experimental target hybridization, and thus does not interfere with the experimental target binding to the experimental probe portion of the same molecule. Because the control target has an identical number of molecules to bind to as the experimental target, the signal ratio is very accurate among all the features on the array. Also, because the control probe oligonucleotide sequence is identical in every feature, one can employ control targets all having the same sequence. Consequently, one can manufacture control targets with only one synthesis reaction.

Example 3

In this example, an array and a set of corresponding control targets are combined together as a kit. An array is designed as a specific diagnostic assay to search for a certain set of diseases in a human patient, such as a set of diseases known to affect various transcription levels in blood cells. A set of control target oligonucleotides is synthesized to correspond to each feature in the array. The

features include different experimental probes that allow one to determine the amount of each transcript in the sample.

Example 4

A set of synthetic control target oligonucleotides are provided. The oligonucleotides represent every known transcript expressed in the liver. Included in the pool of control target oligonucleotides are those transcripts known to be expressed in liver in diseased or cancerous states. The entire pool of oligonucleotides is packaged as a synthetic liver control target kit. This synthetic control target kit can be employed in any nucleic acid array used as an assay on liver tissue, regardless of manufacturer, or the identity of the transcripts that are sought.

Example 5

This example refers to synthetic control targets which can be used on any array. In certain embodiments, control target oligonucleotides are generated using random bases in a number of places. Because the control target oligonucleotides are short in this embodiment (8 to 12 bases in length), and because they include several random bases, the probability that any probe would be complementary to at least one of the control target molecules is very high. Consequently, a pool of control target oligonucleotides is generated which is complementary to any probe on any array.

Because the control target oligonucleotides are short, they possess a low T_m (which varies based on their sequence composition). While a ten base-pair hybrid typically would not survive the stringent conditions necessary for a standard microarray hybridization assay, the oligonucleotides can be chemically modified to increase the T_m of their hybrids. For example, locked nucleic acids (LNA's) have inverted stereochemistries at C-3' and C-4' to provide an L-ribo confirmation. Bases joined by LNA's hybridize to their complements with high affinity, and have much higher melting temperatures than their naturally occurring counterparts with phosphodiester bonds.

The short control target oligonucleotides include a label attached. These synthetic control targets comprise a universal control target sample which provides a reliable and consistent control signal for the amount of probe sequence bound to each feature of the array, and provides a control signal that can be compared with the

experimental signal for analysis. This synthetic control target sample can be used in any microarray assay, regardless of probe sequences or array manufacture.

While the present invention has been described in terms of these exemplary embodiments, the skilled artisan will readily understand that numerous variations and modifications of these embodiments are possible without undue experimentation. All such variations and modifications are within the intended scope of the claimed invention.

What is claimed is:

1. A method for detecting the presence or absence or amount of an experimental target-specific probe on a substrate comprising:
 - exposing a substrate containing a first feature comprising an experimental target-specific probe and a control-specific probe to a labeled control target, such that the labeled control target binds specifically to control-specific probe bound to the substrate, and
 - measuring a signal from labeled control target bound to control-specific probe to determine the presence or absence or amount of experimental probe.
2. The method of claim 1, wherein the experimental target-specific probe and control-specific probe are polynucleotides.
3. The method of claim 2, wherein the control-specific probes that are polynucleotides contain synthetic non-Watson-Crick bases.
4. The method of claim 1, wherein the control-specific probe is attached to the experimental target-specific probe.
5. The method of claim 1, wherein the control-specific probe is not attached to the experimental target-specific probe.
6. The method of claim 1, wherein the labeled control target comprises a fluorophore.
7. The method of claim 1, wherein the substrate is exposed to a labeled experimental target molecule in a sample, such that labeled experimental target molecule is bound to experimental target-specific probe.
8. The method of claim 7, wherein the signal from labeled control target bound to control-specific probe is used to quantitate the amount of labeled experimental target molecule in a sample.
9. The method of claim 1, wherein the substrate further contains a second feature comprising a second experimental target-specific probe and the control specific probe.

10. A method for detecting the presence or absence or amount of an experimental target probe on a substrate comprising:
 - exposing a substrate containing a feature comprising an experimental target probe to a labeled control target, such that the labeled control target binds to experimental target probe bound to the substrate, and
 - measuring a signal from labeled control target bound to experimental target probe to determine the presence or absence or amount of experimental probe.
11. A kit comprising:
 - a substrate comprising a feature containing an attached experimental target-specific probe and an attached control-specific probe, and
 - a labeled control target, wherein the labeled control target binds specifically to the control-specific probe, and does not bind specifically to the experimental target-specific probe.
12. The kit of claim 11, wherein the control-specific probe, experimental target-specific probe, and the labeled control target are polynucleotides.
13. The kit of claim 12, wherein at least the control-specific probe and the labeled control target contain non-Watson-Crick base pairs.
14. The kit of claim 11, wherein the control-specific probe is attached to the experimental target-specific probe.
15. The kit of claim 11, wherein the control-specific probe is not attached to the experimental target-specific probe.
16. A synthetic control for use in a nucleic acid array, comprising a pool of different control polynucleotides attached to labels, wherein each control polynucleotide is designed to complement a polynucleotide that is attached to a feature in the array.
17. The synthetic control of claim 16, wherein the different control polynucleotides complement different expressed transcripts or alleles of a specific species of organism.

18. The synthetic control of claim 17, wherein the species is selected from a human, rabbit, mouse, and rat.
19. The synthetic control of claim 16, wherein the different control polynucleotides complement a broad range of expressed transcripts of a specific sample type.
20. The synthetic control of claim 19, wherein the sample type is selected from whole blood, peripheral blood lymphocytes, hair, semen, lymph, amniotic fluid, cerebrospinal fluid, placental tissue, vaginal tissue, sputum, and skin.
21. The synthetic control of claim 16, wherein the different control polynucleotides are short sequences comprising random nucleotides.
22. The synthetic control of claim 21, wherein the different control polynucleotides are designed such that the pool of different control polynucleotides will complement any probe attached to a nucleic acid array.
23. The synthetic control of claim 22, wherein the different control polynucleotides comprise bases or linkages that are chemically modified such that the different control polynucleotides have a T_M above that of naturally occurring nucleotides.
24. The synthetic control of claim 23, wherein the bases are linked by peptide bonds.
25. The synthetic control of claim 21, wherein the different control polynucleotides are 9 or 10 bases in length.
26. A method for determining the relative amount of two different experimental target nucleic acid sequences in a sample, comprising:
 - providing a nucleic acid array, wherein the array comprises a first feature that comprises first experimental target-specific probes that are complementary to a first experimental target sequence and control-specific probes that do not hybridize to experimental target sequences, and a second feature that comprises second experimental target-specific probes that are complementary to a second experimental target sequence and control-specific probes that do not hybridize to experimental target sequences;

contacting the array with:

- (1) a sample that includes experimental target sequences that are labeled with a first label; and
- (2) synthetic control target sequences that are labeled with a second label and that hybridize to the control-specific probes on each of the first and second features, but that do not hybridize with the experimental target-specific probes;

determining the intensity of any signal from the first and second labels in the first and second features;

determining the ratio of the intensity of the signal from the first label to the intensity of the signal from the second label for each of the first and second features; and

comparing the ratios of the intensity of the signal for the first and second features to calculate the relative amount of first and second experimental nucleic acid target sequences in the sample.

CLOSE UP VIEW SHOWING
342 FEATURES OF A 55K
(1X0.6 in) FEATURE
ARRAY OF SPOTTED DYE.

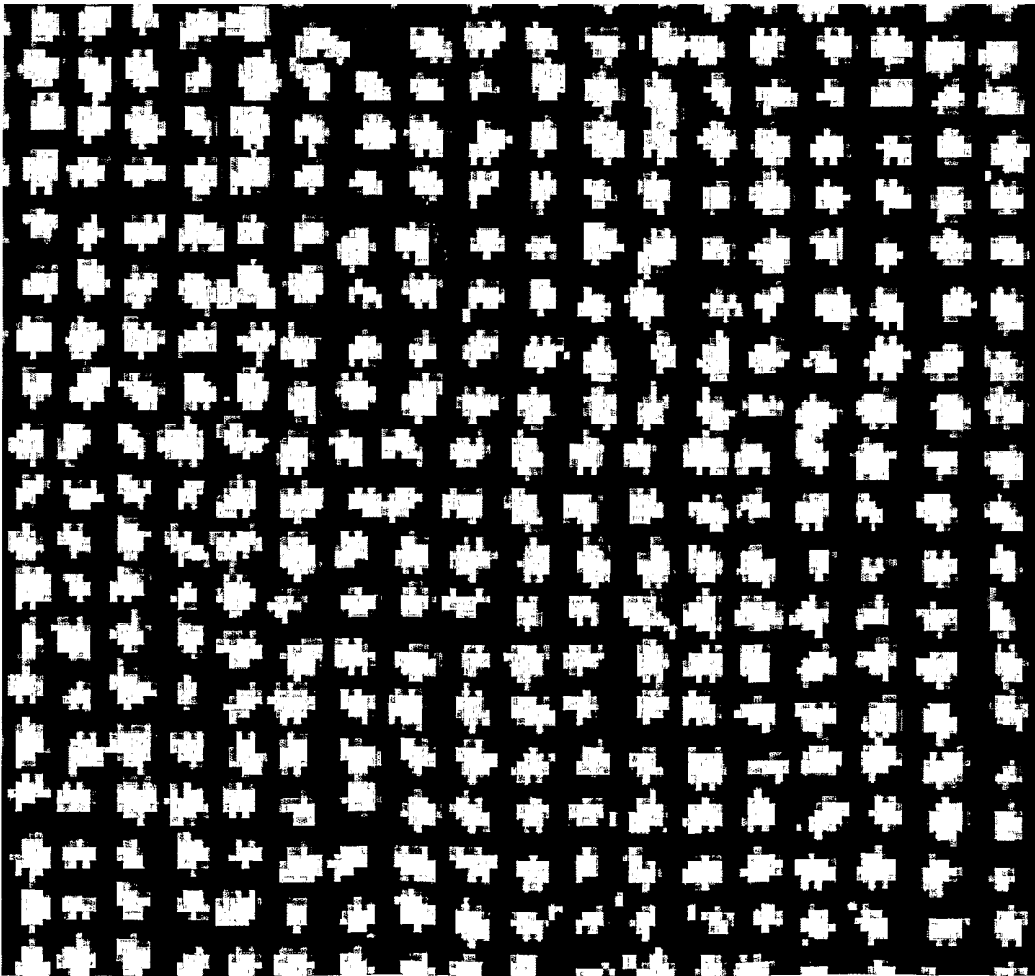


FIG. 1

A COMPETITIVE HYBRIDIZATION BETWEEN EXPERIMENTAL TARGET AND CONTROL TARGET LABELED WITH TWO DIFFERENT SIGNALS. NOTE IN THIS APPROACH FEATURES ARE NOT DETECTABLE UNLESS A HYBRIDIZATION EVENT OCCURS.

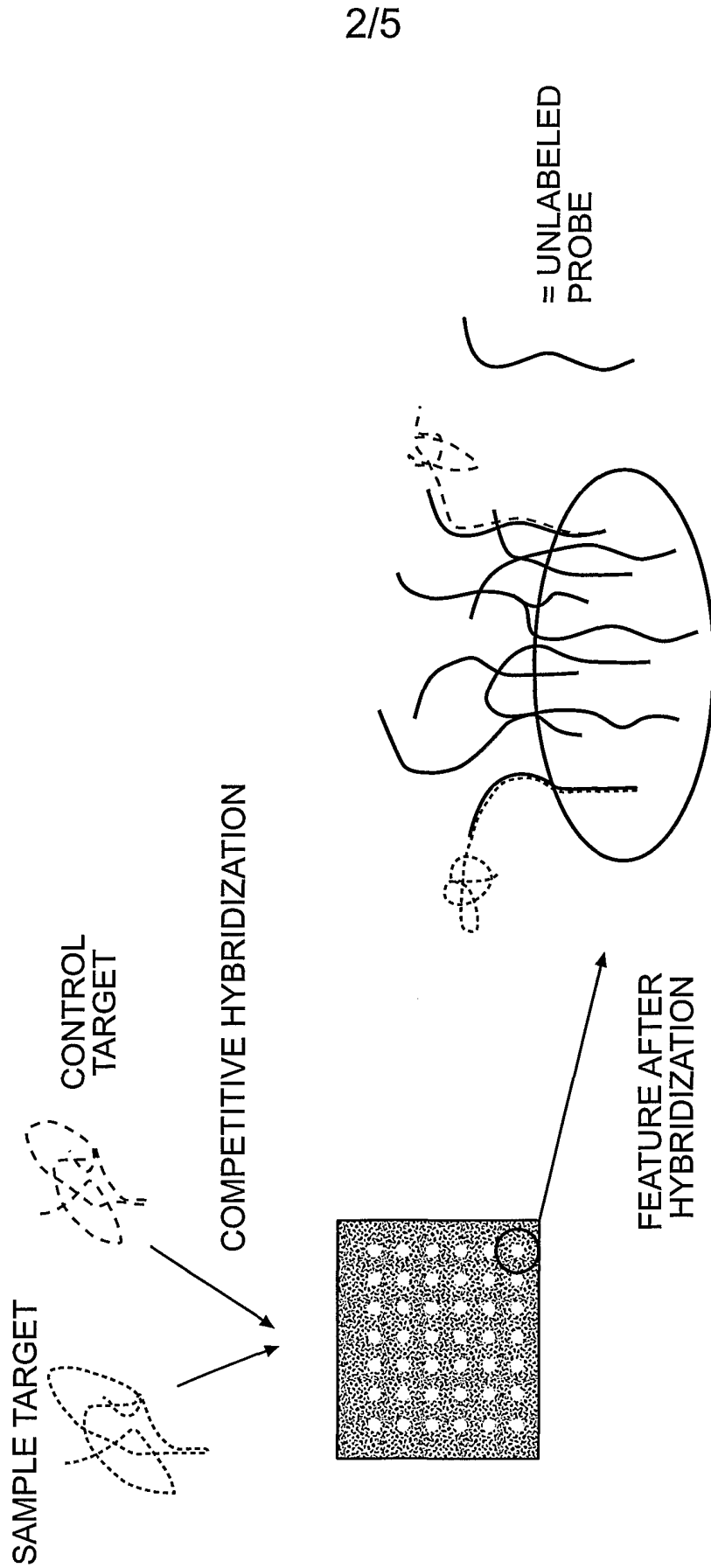
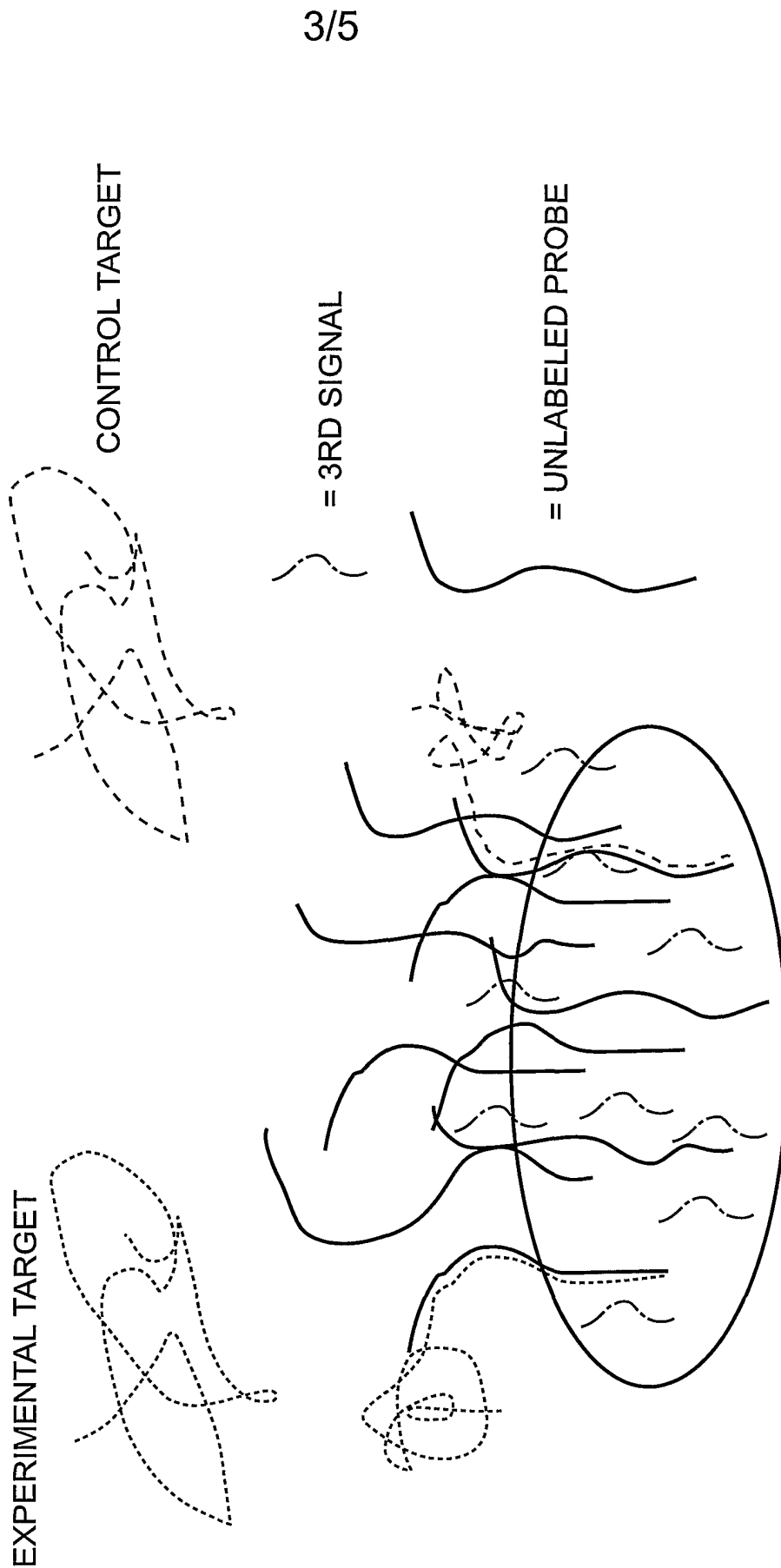


FIG. 2

PUT A THIRD SIGNAL INTO EACH FEATURE DURING MANUFACTURING OF THE ARRAY. USE THIRD SIGNAL FOR SPOT FINDING AND QUANTITATION. A COMPETITIVE HYBRIDIZATION BETWEEN BIOLOGICALLY DERIVED EXPERIMENTAL AND CONTROL TARGETS LABELED WITH TWO DIFFERENT SIGNALS IS THEN PERFORMED.



3/5

FIG. 3

FEATURE

DEPOSIT EACH FEATURE WITHOUT LABEL. PERFORM NON-COMPETITIVE HYBRIDIZATION TO ARRAY PROBES USING A SYNTHETIC FIRST-LABELED CONTROL OLIGO AND SECOND-LABELED cDNA SAMPLES. REQUIRES A HYBRIDIZATION REAGENT CONTAINING A SINGLE FIRST-LABELED CONTROL OLIGO COMPLEMENTARY TO EACH FEATURE ON THE ARRAY. USE THE FIRST SIGNAL ON THE CONTROL OLIGO FOR SPOT FINDING AND QUANTITATION.

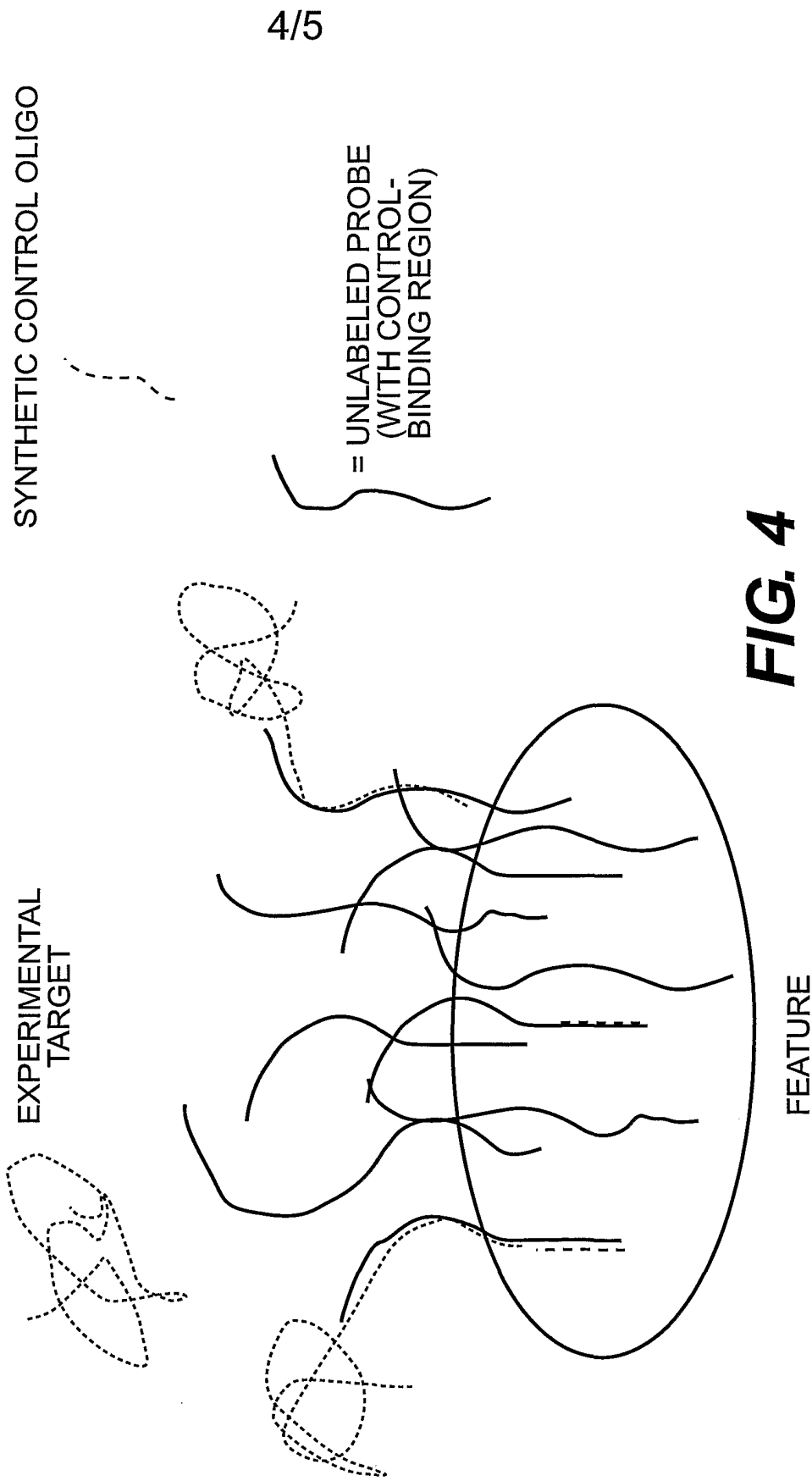


FIG. 4

PUT A LABEL INTO EACH FEATURE DURING MANUFACTURING OF THE ARRAY. LABEL PLACED EITHER DIRECTLY ON PROBE OLIGOS OR ONTO DIFFERENT CO-SPOTTED OLIGOS. USE A FIRST SIGNAL FOR FEATURE FINDING AND QUANTITATION. A GENE EXPRESSION ASSAY WOULD INCLUDE A 1-COLOR HYBRIDIZATION OF A SECOND LABELED EXPERIMENTAL TARGET ONTO ARRAY. USE THE FIRST SIGNAL TO HELP BETTER QUANTITATE THE SIGNAL OF THE SECOND LABEL.

SAMPLE TARGET

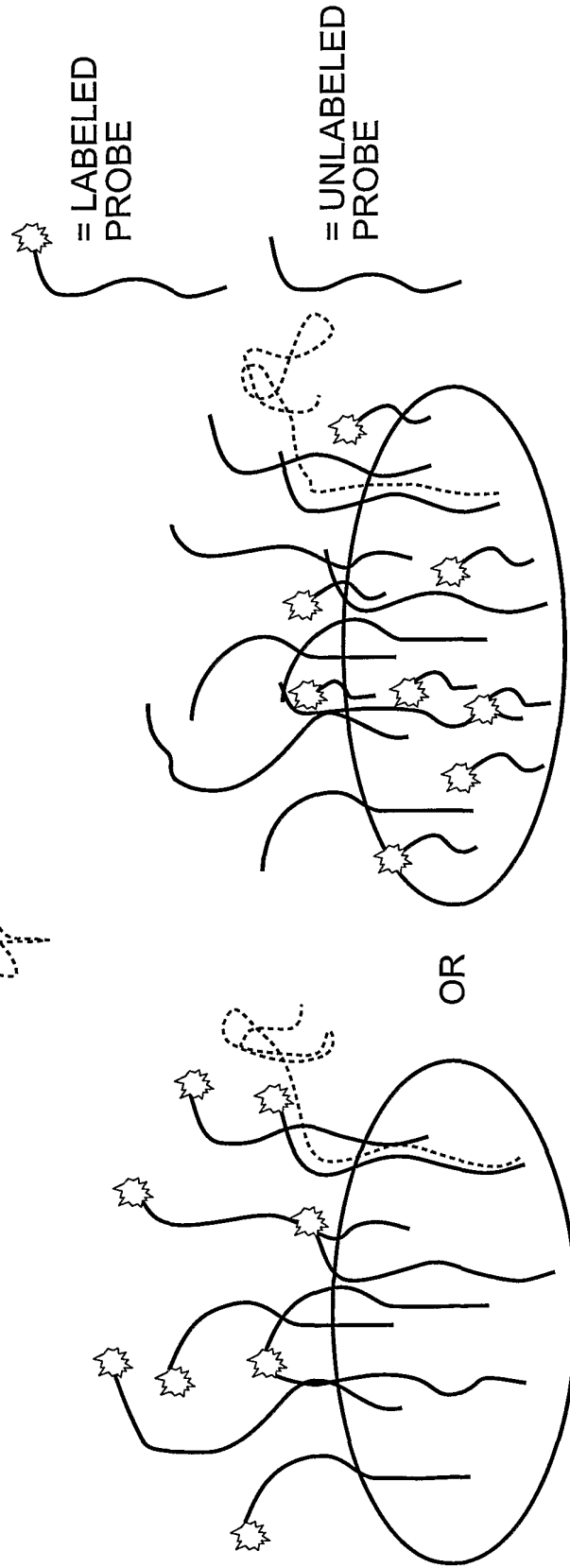
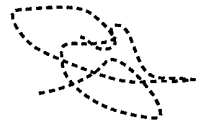


FIG. 5