HU000026666T2

(19) **HU** (11) Lajstromszám: **E 026 666** (13) **T2**

MAGYARORSZÁG
Szellemi Tulajdon Nemzeti Hivatala

EURÓPAI SZABADALOM
SZÖVEGÉNEK FORDÍTÁSA

(21) Magyar ügyszám: **E 11 766613**
(22) A bejelentés napja: **2011. 04. 05.**

(96) Az európai bejelentés bejelentési száma:
EP 20110766613
(97) Az európai bejelentés közzétételi adatai:
EP 2556171 A1 **2011. 10. 13.**
(97) Az európai szabadalom megadásának meghirdetési adatai:
EP 2556171 B1 **2015. 09. 02.**

(51) Int. Cl.: ***C12Q 1/68*** *(2006.01)*

(86) A nemzetközi (PCT) bejelentési szám:
PCT/US 11/031308
(87) A nemzetközi közzétételi szám:
WO 11127099

(30) Elsőbbségi adatok:

201113080616	**2011. 04. 05.**	**US**
321124 P	**2010. 04. 05.**	**US**

(72) Feltaláló(k):
CHEE, Mark, S., La Jolla, CA 92037 (US)

(73) Jogosult(ak):
Prognosys Biosciences, Inc., La Jolla, CA 92037 (US)

(74) Képviselő:
Mester Tamás, SWORKS Nemzetközi Szabadalmi Ügyvivői Iroda Kft., Budapest

(54) **Térbelileg kódolt biológiai assay-ek**

Az európai szabadalom ellen, megadásának az Európai Szabadalmi Közlönyben való meghirdetésétől számított kilenc hónapon belül, felszólalást lehet benyújtani az Európai Szabadalmi Hivatalnál. (Európai Szabadalmi Egyezmény 99. cikk(1))

A fordítást a szabadalmas az 1995. évi XXXIII. törvény 84/H. §-a szerint nyújtotta be. A fordítás tartalmi helyességét a Szellemi Tulajdon Nemzeti Hivatala nem vizsgálta.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 2 556 171 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention
of the grant of the patent:
02.09.2015 Bulletin 2015/36

(21) Application number: **11766613.1**

(22) Date of filing: **05.04.2011**

(51) Int Cl.:
C12Q 1/68 $^{(2006.01)}$

(86) International application number:
PCT/US2011/031308

(87) International publication number:
WO 2011/127099 (13.10.2011 Gazette 2011/41)

(54) SPATIALLY ENCODED BIOLOGICAL ASSAYS

RÄUMLICH KODIERTE BIOLOGISCHE ASSAYS

TESTS BIOLOGIQUES À CODAGE SPATIAL

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **05.04.2011 US 201113080616**
05.04.2010 US 321124 P

(43) Date of publication of application:
13.02.2013 Bulletin 2013/07

(73) Proprietor: **Prognosys Biosciences, Inc.**
La Jolla, CA 92037 (US)

(72) Inventor: **CHEE, Mark, S.**
La Jolla, CA 92037 (US)

(74) Representative: **Green, Katherine et al**
Mewburn Ellis LLP
City Tower
40 Basinghall Street
London EC2V 5DE (GB)

(56) References cited:
WO-A1-2010/019826 US-A1- 2006 046 313
US-B1- 6 344 316

- **MIR KALIM U ET AL: "Sequencing by Cyclic Ligation and Cleavage (CycLiC) directly on a microarray captured template", NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, GB, vol. 37, no. 1, 1 January 2009 (2009-01-01), pages e5-1, XP002542688, ISSN: 1362-4962, DOI: 10.1093/NAR/GKN906 [retrieved on 2008-11-16]**

Printed by Jouve, 75001 PARIS (FR)

Description

FIELD OF THE INVENTION

[0001] This invention relates to assays of biological molecules, and more particularly to assays for determining spatial distributions of a large number of biological molecules in a solid sample simultaneously.

BACKGROUND OF THE INVENTION

[0002] In the following discussion certain articles and methods will be described for background and introductory purposes. Nothing contained herein is to be construed as an "admission" of,prior art. Applicant expressly reserves the right to demonstrate, were appropriate, that the articles and methods referenced herein do not constitute prior art under the applicable statutory provisions.

[0003] Comprehensive gene expression analysis and protein analysis have been useful tools in understanding mechanisms of biology. Use of these tools has allowed the identification of genes and proteins involved in development and in various diseases such as cancer and autoimmune disease. Conventional methods such as *in situ* hybridization and other multiplexed detection of different transcripts have revealed spatial patterns of gene expression and have helped shed light on the molecular basis of development and disease. Other technologies that have enabled the quantitative analysis of many RNA sequences per sample include microarrays (see Shi, et al., Nature Biotechnology, 24(9):1151-61 (2006); and Slonim and Yanai, Plos Computational Biology, 5(10):e1000543 (2009)); serial analysis of gene expression (SAGE) (see Velculescu, et al., Science, 270(5235):484-87 (1995)), high-throughput implementation of qPCR (see Spurgeon, et al., Plos ONE, 3(2):e1662 (2008), and *in situ* PCR (see Nuovo, Genome Res:, 4:151-67 (1995)). As useful as these methods are, however, they do not enable simultaneous measurement of the expression of many genes or the presence and/or activity of multiple proteins at many spatial locations in a sample. Laser capture microdissection has permitted the analysis of many genes at a small number of locations, but it is very extensive, laborious, and does not scale well. Certain PCR essays in a 2D format preserve spatial information (see Armani, et al., Lab on a Chip, 9(24): 3526-34 (2009)), but these methods have low spatial resolution because they rely on physical transference of tissue into wells, which also prevents random access to tissue samples and high levels of multiplexing.

[0004] At present, no practical method exists to analyze at high resolution the spatial expression patterns of large numbers of genes, proteins, .or other biologically active molecules simultaneously. There is thus a need for reproducible, high-resolution spatial maps of biological molecules in tissues. The present invention addresses this need.

[0005] WO2010/019826 describes detection and quantification of individual target molecules in biomolecular samples using stable nanoreporter probes with uniquely detectable signal.

[0006] MIR KALIM et al., NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, GB, vol. 37, no. 1, 1 January 2009 (2009-01-01), pages e5-1, describes Sequencing by Cyclic Ligation and Cleavage (CycLiC) directly on a microarray captured template.

SUMMARY OF THE INVENTION

[0007] This Summary is provided to introduce a selection of concepts in a simplified form that are further described below in the Detailed Description. This Summary is not intended to identify key or essential features of the claimed subject matter, nor is it intended to be used to limit the scope of the claimed subject matter. Other features, details; utilities, and advantages of the claimed subject matter will be apparent from the following written Detailed Description including those aspects illustrated in the accompanying drawings"and defined in the appended claim.

[0008] The disclosure relates to assay systems that provide high-resolution spatial maps of biological activity in tissues. The assay system comprises an assay capable of high levels of multiplexing where encoded probes are provided to a biological sample in defined spatial patterns; instrumentation capable of controlled delivery of reagent according to the spatial patterns; and a decoding scheme providing a readout that is digital in nature. In short, the present, invention provides the ability to look at many biological targets in many locations, providing the resolution of in situ hybridization with the highly-parallel data analysis of sequencing.

[0009] Thus, the invention provides a method according to claim 1 and a method according to claim 2.

[0010] in particular aspects of the invention the biological targets comprise nucleic acids and the encoded probes are oligonucleotides, and in some aspects, there are two encoded probes for each of the multiples, nucleic acid targets. In some aspects, the multiple biological targets comprise proteins, the probe regions of the encoding probes are proteins and the coding tags comprise oligonucleotides. In some aspects the multiple biological targets comprise enzymes. In some, aspects the probe regions of the encoded probes comprise antibodies, aptamers or small molecules.

[0011] Some aspects of the assay system further comprise an amplification step between the separating step and the determining step. In some aspects, the determining step is performed by nucleic acid sequencing, and in preferred aspects, the sequencing is high-throughput digital nucleic acid sequencing.

[0012] In some aspects, of the invention; the product of the multiple biological targets being assayed and the multiple sites in the sample is greater than 20, in some aspects product of the multiple biological targets being assayed and the multiple sites in the sample is greater than 50, in.some aspects the product of the multiple biological targets being assayed and the multiple sites in the sample is greater than 75, 100, 150, 500, 750, 1,000, 5,000, 10,000, 25,000, 50,000, 100,000, 500,000, or 1,000,000 or more. In other aspects, the sequence of at least fifty thousand encoding probes are determined in parallel, in other aspects the sequences of at least one hundred thousand encoding probes are determined in parallel, in some aspects the sequence of at least five hundred thousand encoding probes are determined in parallel, and in some aspects the sequence of at least one million, ten million, one hundred million, one billion, ten billion, one hundred billion or more encoding probes are determined in parallel.

[0013] In Some aspects, the known spatial pattern is determined by histological features of the sample. Also in some aspects; software programmed hardware performs at least two steps of the delivering step, the separation step, the determining step and the associating step.

[0014] In some aspects, the probe regions of the encoded probes are proteins and the separating step is accomplished by encoded probes that interact with the biological targets being captured by an affinity capture agent. In some aspects the probe regions of the encoding probes are nucleic acids and the separating step is accomplished by a washing of the sample.

[0015] The disclosure also relates to an assay system to determine spatial patterns of abundance or activity or both of multiple protein targets at multiple sites in a sample, where the assay system performs the following steps: providing a sample affixed to a support; delivering encoded probes for the multiple protein targets to the multiple sites in the sample in a known spatial pattern, where each encoded probe comprises a protein probe region that may interact with the protein targets and a coding tag that identifies a location of the site to which the encoded probe was delivered and the protein probe region of the encoding probe of which the coding tag is part; allowing the encoded probes to interact with the protein targets; separating encoded probes that interact with the protein targets from encoded probes that do not interact with the protein targets; determining all or a portion of a sequence of the encoded probes by high throughput sequencing, and associating the abundance or activity or both of the multiple protein targets to the locations of the multiple sites in the sample.

[0016] The disclosure also relates to an assay system to determine spatial patterns of abundance or activity of both of multiple biological targets at multiple sites in a sample, where the assay system performs the following steps: providing a sample affixed to a support; delivering encoded probes for the multiple biological targets to the multiple sites in the sample in a known spatial pattern, where each encoded probe comprise a probe region that may interact with the biological targets and a coding tag that identifies a location of the site to which the encoded probe was delivered and identifies the biological target; allowing the encoded probes to interact with the biological targets; determining all or a portion of a sequence of the encoded probes, and associating the abundance or activity or both of the multiple biological targets to the locations of the sites in the sample.

[0017] The invention can utilise various detection mechanisms, based on the molecules to be detected and the reagents needed for such detection system. Exemplary methods that can be used with the invention are described in more detail below.

DESCRIPTION OF THE FIGURES

[0018]

Figure 1 provides a simplified overview of the assay system using the present.

Figure 2 provides a simplified overview of one embodiment of the assay systems using the present invention for detecting nucleic acids.

Figure 3 is a representational depiction of one embodiment of the assay overviewed in Figure 2.

Figure 4 illustrates a general mechanism for one embodiment of a combinatorial encoding scheme of the assay systems using the invention.

Figure 5 provides a simplified, specific example of the embodiment of a combinatorial encoding scheme shown in Figure 4.

DEFINITIONS

[0019] The terms used herein are intended to have the plain and ordinary meaning as understood by those of ordinary skill in the art. The following definitions are intended to aid the reader in understanding the present invention, but are not intended to vary or otherwise limit the meaning of such terms unless specifically indicated.

[0020] The term "antibody" as used herein is intended to refer to an entire immunoglobulin or antibody or any functional fragment of an immunoglobulin molecule which is capable of specific binding to an antigen (antibodies and antigens are "binding partners" as defined herein). "Antibody" as used herein is meant to include the entire antibody as well as any antibody fragments capable of binding the antigen or antigenic fragment of interest. Examples of such peptides include complete antibody molecules, antibody fragments, such as Fab, F(ab')2, CDRS, VL, VH, and any other portion of an antibody which is capable of specifically binding to an antigen. Antibodies for assays of the invention are immunoreactive or immunospecific for, and therefore specifically and selectively bind to, proteins either detected (*i.e.,* biological targets) or used for detection (*i.e.,* probes) in the assays of the invention.

[0021] The term "binding agent" as used herein refers to any agent that specifically binds to a biological molecule of interest.

[0022] "Complementary" or "substantially complementary" refers to the hybridization or base pairing or the formation of a duplex between nucleotides or nucleic acids, such as, for instance, between the two strands of a double-stranded DNA molecule or between an oligonucleotide primer and a primer binding site on a single-stranded nucleic acid. Complementary nucleotides are, generally, A and T (or A and U), or C and G. Two single-stranded RNA or DNA molecules are said to be substantially complementary when the nucleotides of one strand, optimally aligned and compared and with appropriate nucleotide insertions or deletions, pair with at least about 80% of the other strand, usually at least about 90% to about 95%, and even about 98% to about 100%.

[0023] "Hybridization" refers to the process in which two single-stranded polynucleotides bind non-covalently to form a stable double-stranded polynucleotide. The resulting (usually) double-stranded polynucleotide is a "hybrid" or "duplex." "Hybridization conditions" will typically include salt concentrations of approximately less than 1M, often less than about 500 mM and may be less than about 200 mM. A "hybridization buffer" is a buffered salt solution such as 5% SSPE, or other such buffers known in the art. Hybridization temperatures can be as low as 5°C, but are typically greater than 22°C, and more typically greater than about 30°C, and typically in excess of 37°C. Hybridizations are often performed under stringent conditions, *i.e.,* conditions under which a primer will hybridize to its target subsequence but will not hybridize to the other, non-complementary sequences. Stringent conditions are sequence-dependent and are different in different circumstances. For example, longer fragments may require higher hybridization temperatures for specific hybridization than short fragments. As other factors may affect the stringency of hybridization, including base composition and length of the complementary strands, presence of organic solvents, and the extent of base mismatching, the combination of parameters is more important than the absolute measure of any one parameter alone. Generally stringent conditions are selected to be about 5°C lower than the T_m for the specific sequence at a defined ionic strength and pH. Exemplary stringent conditions include a salt concentration of at least 0.01M to no more than 1M sodium ion concentration (or other salt) at a pH of about 7.0 to about 8.3 and a temperature of at least 25°C. For example, conditions of 5xSSPE (750 mM NaCl, 50 mM sodium phosphate, 5 mM EDTA at pH 7.4) and a temperature of approximately 30°C are suitable for allele-specific hybridizations, though a suitable temperature depends on the length and/or GC content of the region hybridized.

[0024] "Ligation" means to form a covalent bond or linkage between the termini of two or more nucleic acids, *e.g.,* oligonucleotides and/or polynucleotides, in a template-driven reaction. The nature of the bond or linkage may vary widely and the ligation may be carried out enzymatically or chemically. As used herein, ligations are usually carried out enzymatically to form a phosphodiester linkage between a 5' carbon terminal nucleotide of one oligonucleotide with a 3' carbon of another nucleotide.

[0025] "Nucleic acid", "oligonucleotide", "oligo" or grammatical equivalents used herein refers generally to at least two nucleotides covalently linked together. A nucleic acid generally will contain phosphodiester bonds, although in some cases nucleic acid analogs may be included that have alternative backbones such as phosphoramidite, phosphorodithioate, or methylphophoroamidite linkages; or peptide nucleic acid backbones and linkages. Other analog nucleic acids include those with bicyclic structures including locked nucleic acids, positive backbones, non-ionic backbones and non-ribose backbones. Modifications of the ribose-phosphate backbone may be done to increase the stability of the molecules; for example, PNA:DNA hybrids can exhibit higher stability in some environments.

[0026] "Primer" means an oligonucleotide, either natural or synthetic, that is capable, upon forming a duplex with a polynucleotide template, of acting as a point of initiation of nucleic acid synthesis and being extended from its 3' end along the template so that an extended duplex is formed. The sequence of nucleotides added during the extension process is determined by the sequence of the template polynucleotide. Primers usually are extended by a DNA polymerase.

[0027] The term "SNP" or "single nucleotide polymorphism" refers to a genetic variation between individuals; *e.g.*, a single nitrogenous base position in the DNA of organisms that is variable. SNPs are found across the genome; much of the genetic variation between individuals is due to variation at SNP loci, and often this genetic variation results in phenotypic variation between individuals. SNPs for use in the present invention and their respective alleles may be derived from any number of sources, such as public databases (U.C. Santa Cruz Human Genome Browser Gateway (http://genome.ucsc.edu/cgi-bin/hgGateway) or the NCBI dbSNP website (http://www.ncbi.nlm.nih.gov/SNP/), or may be experimentally determined as described in U.S. Pat. No. 6,969,589; and US Pub. No. 2006/0188875 entitled "Human Genomic Polymorphisms." Although the use of SNPs is described in some of the embodiments presented herein, it will be understood that other biallelic or multi-allelic genetic markers may also be used. A biallelic genetic marker is one that has two polymorphic forms, or alleles. As mentioned above, for a biallelic genetic marker that is associated with a trait, the allele that is more abundant in the genetic composition of a case group as compared to a control group is termed the "associated allele," and the other allele may be referred to as the "unassociated allele." Thus, for each biallelic polymorphism that is associated with a given trait (*e.g.,* a disease or drug response), there is a corresponding associated allele. Other biallelic polymorphisms that may be used with the methods presented herein include, but are not limited to multinucleotide changes, insertions, deletions, and translocations. It will be further appreciated that references to DNA herein may include genomic DNA, mitochondrial DNA, episomal DNA, and/or derivatives of DNA such as amplicons, RNA transcripts, cDNA, DNA analogs, etc. The polymorphic loci that are screened in an association study may be in a diploid or a haploid state and, ideally, would be from sites across the genome.

[0028] The term "selectively binds", "selective binding" and the like as used herein, when referring to a binding partner (*e.g.,* protein, nucleic acid, antibody or other affinity capture agent, etc.), refers to a binding reaction of two or more binding partners with high affinity and/or complementarity to ensure selective hybridization under designated assay conditions. Typically, specific binding will be at least three times the standard deviation of the background signal. Thus, under designated conditions the binding partner binds to its particular "target" molecule and does not bind in a significant amount to other molecules present in the sample.

[0029] "Sequencing", "sequence determination" and the like means determination of information relating to the nucleotide base sequence of a nucleic acid. Such information may include the identification or determination of partial as well as full sequence information of the nucleic acid. Sequence information may be determined with varying degrees of statistical reliability or confidence. In one aspect, the term includes the determination of the identity and ordering of a plurality of contiguous nucleotides in a nucleic acid. "High throughput digital sequencing" or "next generation sequencing" means sequence determination using methods that determine many (typically thousands to billions) of nucleic acid sequences in an intrinsically parallel manner, *i.e.* where DNA templates are prepared for sequencing not one at a time, but in a bulk process, and where many sequences are read out preferably in parallel, or alternatively using an ultra-high throughput serial process that itself may be parallelized. Such methods include but are not limited to pyrosequencing (for example, as commercialized by 454 Life Sciences, Inc., Branford, CT); sequencing by ligation (for example, as commercialized in the SOLiD™ technology, Life Technology, Inc., Carlsbad, CA); sequencing by synthesis using modified nucleotides (such as commercialized in TruSeq™ and HiSeq™ technology by Illumina, Inc., San Diego, CA, HeliScope™ by Helicos Biosciences Corporation, Cambridge, MA, and PacBio RS by Pacific Biosciences of California, Inc., Menlo Park, CA), sequencing by ion detection technologies (Ion Torrent, Inc., South San Francisco, CA); sequencing of DNA nanoballs (Complete Genomics, Inc., Mountain View, CA); nanopore-based sequencing technologies (for example, as developed by Oxford Nanopore Technologies, LTD, Oxford, UK), and like highly parallelized sequencing methods.

[0030] The term "T_m" is used in reference to the "melting temperature." The melting temperature is the temperature at which a population of double-stranded nucleic acid molecules becomes half dissociated into single strands. Several equations for calculating the T_m of nucleic acids are well known in the art. As indicated by standard references, a simple estimate of the T_m value may be calculated by the equation, T_m=81.5+0.41 (% G+C), when a nucleic acid is in aqueous solution at 1M NaCl (see *e.g.,* Anderson and Young, Quantitative Filter Hybridization, in Nucleic Acid Hybridization (1985)). Other references (*e.g.,* Allawi and SantaLucia, Jr., Biochemistry, 36:10581-94 (1997)) include alternative methods of imputation which take structural and environmental, as well as sequence characteristics into account for the calculation of T_m.

DETAILED DESCRIPTION OF THE INVENTION

[0031] The practice of the techniques described herein may employ, unless otherwise indicated, conventional techniques and descriptions of organic chemistry, polymer technology, molecular biology (including recombinant techniques), cell biology, biochemistry, and sequencing technology, which are within the skill of those who practice in the art. Such conventional techniques include polymer array synthesis, hybridization and ligation of polynucleotides, and detection of hybridization using a label. Specific illustrations of suitable techniques can be had by reference to the examples herein. However, other equivalents conventional procedures can, of course, also be used. Such conventional techniques and descriptions can be found in standard laboratory manuals such as Green, et al., Eds., Genome Analysis: A Laboratory manual Series (Vols. I-IV) (1999); Weiner, Gabriel, Stephens, Eds., Generic Variation: A Laboratory Manual (2007); Dieffenbach, Dveksler, Eds., PGR Primer: A Laboratory Manual (2003); Bowtell and Sambrook, DNA Microarrays: A Molecular Cloning Manual (2003); Mount, Bioinformatics: Sequence and Genome Analysis (2004); Sambrook and Russell, Condensed Protocols from Molecular Cloning: A Laboratory Manual (2006); and Sambrook and Russell, Molecular Cloning: Laboratory Manual (2002) (all from Cold Spring Harbour Laboratory Press); Stryer, Biochemistry (4th Ed.) (1995) W.H. Freeman; New York N.Y.; Gait, "Oligonucleotide Synthesis: A Practical Approach" (2002) IRL Press, London; Nelson and Cox, Lehninger. Principles of Biochemistry (2000) 3rd Ed., W. H. Freeman Pub., New York, N.Y.; and Berg, et al., Biochemistry (2002) 5th Ed., W.H. Freeman Pub., New York, N.Y.

[0032] Note that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a nucleic acid" refers, to one or more nucleic acids, and reference to "the assay" includes reference to equivalent steps and methods known to those skilled in the art, and so forth.

[0033] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs.

[0034] Where a range of values is provided, it is understood that each intervening value, between the upper and lower limit of that range and any other stated or intervening value in that stated range is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges, and are also encompassed within the invention, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either both of those included limits are, also included in the invention.

[0035] In the following description, numerous specific details are set forth to provide a more thorough understanding of the present invention. However, it will be apparent to one of skill in the art that the present invention may be practiced without one or more of these specific details. In other instances, well-known features and procedures well known to those skilled in the art have not been described in order to avoid obscuring the invention.

The Invention in General (The invention is defined in the claims.)

[0036] The invention relates to spatially-encoded, multiplexed assays comprising 1) an assay capable of high levels of multiplexing with an efficient spatial encoding scheme; 2) instrumentation capable of delivering reagents according to a spatial pattern; and 3) decoding determined by a readout that is digital in nature. The assay systems detect the presence or absence and relative amount of a biological target or biological activity indicative of a biological target, as well as the location of the biological target or activity in a biological sample, *e.g.*, a tissue section or other biological structure disposed upon a support such as a microscope slide or culture dish.

[0037] The assay system further provides instrumentation with an ability to deliver reagents in a spatially-defined pattern. This instrumentation, together, with software, reagents and protocols, provides a key component of the highly innovative assay system, allowing for measurement of numerous biological targets or activities in a meaningful spatial environment, including gene expression and peptide localization. An encoding scheme used in these assay systems allows one to determine the location of biological targets or activity (or lack thereof in the biological samples after the products of the multiplexed assay are removed from the biological sample and pooled for analysis. Decoding of the encoding scheme can be performed by, *e.g.*, next-generation sequencing, which easily provides millions to trillions of datapoints at low cost. The assay results such as the amount or activity of biological targets can then be mapped back to specific location in the biological sample. The assay systems open a new analytical window into the complex spatial patterns of cellular function and regulation in biological samples.

[0038] A simplified overview of the assay system 100 using the present invention is provided at Figure 1. At step 110, a biological sample affixed to a support is provided. The biological sample contains biological targets of interest. Biological targets can include any molecule of interest, such as nucleic acids (including, *e.g.*, RNA transcripts, genomic DNA sequences, cDNAs, amplicons, or other nucleic acid sequences) and proteins, enzymes and the like. At step 120, encoded probes are delivered to the biological sample according to a known spatial pattern. Encoded probes comprise probes, which can interact with biological targets of interest, and coding tags, which identify the positions in the sample of the biological targets being assayed, and thus can be used to link assay results back to locations in the sample. Coding tags in most embodiments are oligonucleotides. However, coding tags may also be mass tags, fluorescent labels, or other moieties.

[0039] In some embodiments, the probe and coding tag portions of the encoded probe are pre-coupled before being delivered to the biological sample, For example, in the case where the encoded probes are oligonucleotides, both the probe and coding tag sequence can be synthesized as a single oligonucleotide. Alternatively, the probe and coding tag portions of the encoding probes can be synthesized or obtained separately and combined before delivery to the biological sample (*e.g.*,) two separate oligonucleotides can be synthesized and coupled by, *e.g.*, ligation; or an antibody and an oligonucleotide can be prepared separately and conjugated before delivery to the biological sample). Also, as is described in Figures 2-5, the probes and the coding tags (in encoding oligonucleotides) are synthesized separately, and are delivered to the biological sample at different steps (*e.g.*, probes first and coding tags thereafter, or vice versa) in the assay.

[0040] At step 130, the encoded probes are allowed to react or interact with the biological targets, *i.e.*, conditions are provided to allow *e.g.*, oligonucleotides to hybridize to nucleic acid targets, enzymes to catalyze reactions with protein targets, antibodies to bind epitopes, etc. In the case where the biological targets are nucleic acids, the encoded probes are typically oligonucleotides and hybridize to the target nucleic acids. In the case that the biological targets are proteins, the encoded probes typically are aptamers, small molecules, or oligonucleotide-conjugated proteins that interact with target proteins by binding to them or by reacting with them (that is, one of the proteins is a substrate for the other). Encoding oligonucleotides may be coupled to the probes (proteins) by conjugation, chemical or photo-crosslinking via suitable groups and the like.

[0041] Once encoded probes interact with the biological targets, the encoded probes that interacted with the biological targets must be separated from the encoded probes that did not interact with the biological targets at step 140. In the case where the biological targets are nucleic acids and the encoded probes are oligonucleotides, the separation can be accomplished by, *e.g.*, washing the unhybridized encoded probes from the sample. Similarly, for other assays that are based on affinity binding, including those using aptamer, small molecule, and protein probes, washing steps can be used to remove low affinity binders. In the case where the probe is transformed via interaction with the target, *e.g.*, in the case of a peptide, *e.g.*, via cleavage by a protease or phosphorylation by a kinase, it is convenient to collect, all encoded probes--both encoded probes that interacted with the biological targets and were transformed and encoded probes that were not transformed. After collection or pooling, an antibody or other affinity capture agent can be used to capture probes that were transformed by addition of a moiety (*e.g.*, a phosphate group). In cases where probes have been transformed via cleavage, the transformed probes can be separated, *e.g.*, by capturing the non-transformed probes via a tag that is removed from the transformed probes during the transformation (*e.g.*, by cleavage), or by adding a new tag at the site of cleavage.

[0042] Once the reacted (transformed) or interacted encoded probes are separated from the unreacted or uninteracted encoded probes, the sequence of the reacted and/or interacted encoded probes is determined by, preferably, sequencing. The sequence of the encoded probes allows the mapping of the assay results back to locations in the biological sample.

[0043] Figure 2 provides a simplified overview of an assay system using the present invention embodying an efficient implementation of a combinatorial coding scheme for the encoding of spatial information. For purposes of this overview, the probes are oligonucleotides, but as explained elsewhere, other types of probes can also be used. In step 210, a biological sample affixed to a support, *e.g.*, a tissue sample or other biological structure, is provided. In step 220 one or more oligonucleotide probes are delivered to the biological sample, where the oligonucleotide probes are capable of hybridizing with biological targets in the biological sample. In step 230, the oligonucleotide probes are allowed to interact with (hybridize to) the nucleic acid targets; that is, appropriate conditions are provided where oligonucleotide probes can hybridize to the target nucleic acids.

[0044] In step 240, the Oligonucleotide probes that did not hybridize to target nucleic acids are removed, and thereby separated from oligonucleotide probes that did hybridize to target nucleic acids. In this embodiment, separation can be accomplished by, *e.g.*, washing the sample to remove unhybridized oligonucleotide probes. Next, in step 250, encoding oligonucleotides (the encoding agents) are delivered to the biological sample according to a chosen spatial pattern, where the encoding oligonucleotides comprise coding tags that are used to encoded the location of biological targets in the biological sample. Note that in contrast to the assay system of Figure 1, here the probes and encoding agents (encoding oligonucleotides) are delivered in separate steps. In step 260, the encoding oligonucleotides are coupled to the oligonucleotide probes to create encoded probes. In this case where the probes are oligonucleotides, the encoding oligonucleotides may be coupled to the oligonucleotides probes by, *e.g.*, ligation. Alternatively, the information in the encoding oligonucleotides can be transferred by using a DNA polymerase to extend a probe oligonucleotide that acts as a primer, and thereby copy and incorporate the sequence of the encoding oligonucleotides.

[0045] In step 270, the sequence of the coding tags in the encoded probes as well as the sequence or a portion of the sequence of the probe itself is determined, and in step 280, the target nucleic acids are mapped back to the biological sample. In some embodiments, the abundance of sequences reveals the relative quantity of biological targets at the location. Although this embodiment shows the individual steps in a particular order, so as to better explain the invention, the precise order of the steps can be varied. For example, steps 220 and 250 can be combined, so that a mixture of the probes and encoding oligonucleotides is delivered according to a chosen spatial pattern. Coupling step 260 can then be carried out immediately after the combined steps 220 and 250, or concomitantly with them. In this case, step 240 would then occur after step 260. It can therefore be appreciated that the two key results of this series of steps, *i.e.*, the location-specific encoding of probe molecules and the separation of probe molecules based on their ability to interact with corresponding target molecules, can be accomplished with some flexibility in the implementation of the particular steps. Similarly, there is considerable flexibility in the design of the coding scheme. As described *infra*, the assays of the invention are particularly amenable to combinatorial methods.

[0046] Thus, the present invention provides an ability to look at many different biological targets in many locations, providing the resolution of in situ hybridization with the highly-parallel data analysis of sequencing. In some embodiments, the sum of the multiple biological targets being assayed and the multiple sites in the biological sample is greater than 20, in other embodiments, the sum of the multiple biological targets being assayed and the multiple sites in the biological sample is greater than 50, in other embodiments, the sum of the multiple biological targets being assayed and the multiple sites in the biological sample is greater than 100, greater than 500, 1,000, 10,000, 25,000, 100,000, 500,000, 1,000,000. It will be appreciated that, due to the spatial encoding dimension of the invention, even much larger numbers can be contemplated. For example, assaying 10,000 targets per location x 10,000 locations would generate 10^8 different assays, and even larger numbers than these can easily be contemplated, particularly if spatial locations with resolution on the order of that of single cells are utilized. Further, in embodiments where high-throughput digital sequencing is employed, the sequences of at least 1,000 encoding probes are typically determined in parallel. More typically, using a digital readout, it is desirable to obtain multiple sequence reads for each assay (defined by a probe and a spatial location code). It is desirable to obtain an average of at least 3 copies per assay, and more typically at least 10 or at least 30 copies per assay, depending on the design of the experiment and requirements of the assay. For a quantitative readout with suitable dynamic range, it may be desirable to obtain at lest 1,000 reads per assay. Therefore, if 1,000,000 assays are carried out, the number of sequence reads may be 1 billion or more. With high-throughput digital sequencing, and allowing for redundnacy, the sequence of at least 10,000 encoding probes are determined in parallel or the sequence of at least 100,000, 500,000, 1,000,000, 10,000,000, 100,000,000, 1,000,000,000 or more encoding probes are-determined in parallel.

Assays

[0047] The assay portion of the assay systems comprise the following general steps: delivering probes and encoding agents where the encoding agents, (in some embodiments pre-coupled to the probes) are delivered to the sample according to a known spatial pattern, allowing the probes to interact or react with biological targets in the sample, and, if the probes and encoding agents have not been pre-coupled, coupling the encoding agents to probes.

[0048] The samples include virtually any biological sample or samples that can be affixed to a support or provided essentially in a two-dimensional manner, where the ability to tie an assayed biological target or activity back to the location within the biological sample is important. Exemplary biological samples include tissue sections (*e.g.*, including whole animal sectioning and tissue biopsies), cell populations on slides or culture dishes, and the like. The assay systems are particularly advantageous in that they are compatible with numerous biological sample types, including fresh samples, such as primary tissue sections, and preserved samples (including but not limited to frozen samples and paraformalin-fixed, paraffin-embedded (FFPE) samples. An important aspect of the assay systems is that the biological samples are immobilized on a substrate surface having discrete, independently measurable areas.

[0049] The biological targets to be detected can be any biological molecules including but not limited to proteins, nucleic acids, lipids, carbohydrates, ions, or multicomponent complexes containing any of the above. Examples of subcellular targets include organelles, e.g., mitochondria, Golgi apparatus, endoplasmic reticulum, chlordplasts, endocytic vesicles, exocytic vesicles, vacuoles, lysosomes, etc.

[0050] In some particular embodiments, the assay system is used to analyze nucleic acids, *e.g.*, by genotyping, quantitation of DNA copy number or RNA transcripts, localization of particular transcripts within samples, and the like. Figure 3 illustrates an overall scheme for an exemplary assay for, *e.g.*, detecting single nucleotide polymorphisms (SNPs) that can be used with the assay system. In Figure 3, two oligonucleotide probes are provided. Each oligonucleotide probe comprises a target-specific region (located on either side of the SNP to be analyzed) seen at 305 and 307, and ligation regions, seen at 301 and 303. The oligonucleotide probes are allowed to hybridize to a target nucleic acid (not shown) in the biological sample. At step 302, one of the oligonucleotide probes is extended to incorporate the SNP sequence and ligated to the other probe to form an extended probe comprising target nucleic acid region 309 and ligation regions 301 and 303.

[0051] Two encoding agents, both comprising a coding tag (seen at 315 and 317), a ligation region (seen at 311 and 313), and a primer region (seen at 319 and 321) are combined with and ligated to the extended probe at step 304 to form an encoded target-specific oligonucleotide. Again, in contrast with Figure 1, the probes and encoding agents are delivered at separate steps. Doing so allows use of the combinatorial embodiments described *infra.* In preferred embodiments, the encoding oligonucleotides within a pair of encoding oligonucleotides ligate specifically to one side of the target sequence or the other (*i.e.*, 5' or 3' of the target sequence). Also, typically, the ligation and primer regions of the encoding oligonucleotides and probes are universal; that is, the set of ligation and primer regions used in constructing the probes and encoding oligonucleotides are constant, and only the target-specific regions of the probes and the coding tags of the encoding oligonucleotides differ. However, again in alternative embodiments, the ligation and primer regions are not universal and differ between probes and encoding agents.

[0052] Following ligation, the encoded probes are eluted, pooled, and, optionally, sequencing adapters are added to the encoded probes via PCR. In alternative embodiments, sequencing primers may be ligated to the encoding oligonucleotides, or sequencing primer sequences can be included as part of the encoding oligonucleotide. As seen in Figure 3, each sequencing adapter comprises primer region 319 or 321, compatible with the primer regions 319 and 321 on the encoded probes. The final construct comprising first adapter 327, first primer region 319, first coding tag 315, ligation regions 311 and 301, target region 309, ligation regions 313 and 303, second coding tag 317, second primer region 325 and second adapter 329 is now ready for input into a digital high-throughput sequencing process.

[0053] A combination of extension and ligation reactions are exemplified in Figure 3, but it should be appreciated that a variety of reactions may be used to couple the encoding oligonucleotides to the target-specific oligonucleotides, including ligation only (*e.g.*, for oligonucleotides that hybridize to contiguous portions of the target nucleic acid sequence). Alternatively, an assay utilizing an additional oligonucleotide, such as in the GOLDENGATE® assay (see Fan, et al., Cold Spring Symp. Quant. Biol., 68:69-78 (2003); (Illumina, Inc., San Diego, CA)), may be employed.

[0054] In other embodiments, the assay system also can be used to analyze peptides or proteins, the presence of antibodies, enzymatic and other protein activities, posttranslational modifications, active and non-active forms of peptides, as well as peptide isoforms in a biological sample. Accordingly, the probes may comprise an active region of an enzyme, a binding domain of an immunoglobulin, defined domains of proteins, whole proteins, synthetic peptides, peptides with introduced mutations, aptamers and the like.

[0055] In certain aspects, the probes are substrates for enzymes or proenzymes, *e.g.*, kinases, phosphatases, zymogens, proteases, or fragments thereof. In certain aspects, the probes are phosphorylation substrates used to detect proteins involved in one or more signal transduction pathways, *e.g.*, a kinase or a phosphatase. In another specific approach the probes are specific protease substrates that associate only with individual proteases or classes of proteases. In other approaches, the probes are different processed forms, isoforms and/or domains of an enzyme. Protein-based probes are typically conjugated or otherwise linked to oligonucleotide encoding agents. The oligonucleotide encoding agents in this case would also include a nucleotide sequence component that allows for identification of the protein probe.

[0056] In certain aspects, the present invention provides assays for evaluating differences in the amount and/or activity of biological targets between different locations in a sample and/or between samples. The method includes determining a plurality of encoded results from the biological sample and evaluating the differences in quantity of the biological targets at each location in the biological sample.

Combinatorial Embodiments

[0057] To maximize the efficiency of encoding, a combinatorial approach using pairs of coding tags in the encoding oligonucleotides can be used. By de-coupling the target-specific information and the coding tags, the number of oligonucleotides required is dramatically reduced, with a concomitant decrease in cost.

[0058] Figure 4 illustrates a general mechanism for one embodiment of a combinatorial encoding scheme the invention, where nucleic acids in a representative tissue section (shown at 416) are assayed. Figure 4 at A shows two target-specific/encoding oligonucleotide constructs 420 and 422 (*e.g.*, formed between steps 302 and 304 of Figure 3) specifically bound to a target nucleic acid 402 of interest. The first encoded probe 420 comprises coding tag 408, associated with, *e.g.*, a universal priming site for amplification of the assay products or an adapter to enable identification of the coding identifiers using sequencing technologies 404. The second encoded probe 422 comprises coding tag 406, associated with, *e.g.*, a universal priming site for amplification of the assay products or an adapter to enable identification of the coding identifiers using sequencing technologies 410.

[0059] Figure 4 at B shows the spatial pattern that may be used for twenty different coding tags, a1 through a10 (coding tag 406 on encoded probe 420) and b1 through b10 (coding tag 408 encoded probe 422). Coding tag a1, for example, is deposited on the biological sample in ten discrete areas or spots (shown as the first horizontal line of spots in 412). Coding.tag a2 is deposited on the biological sample in ten spots on the second horizontal line in 412. Coding tag a3 is deposited on the biological sample in ten spots on the third horizontal line in 412, and so on. Whereas the "a" tags are deposited in ten horizontal rows, the "b" tags are deposited in ten vertical rows as shown in 414. For example, coding tag b1 is deposited on the biological sample in ten discrete spots in the first vertical row of 414, coding tag b2 is deposited on the biological sample in ten discrete spots in the second vertical row of 414, and so on. Using such a configuration allows for twenty coding tags to uniquely define 100 different locations on the biological sample.

[0060] Figure 4 at C shows a representative tissue section 416 coincident with coding tag grid 418. The arrows show how the "a" coding tags and the "b" coding tags are deposited on grid 418 that is coincident with tissue section 416. If, once sequenced, coding tags a1 and b4, *e.g.*, are associated with a target nucleic acid sequence, then that target nucleic acid sequence (*i.e.*, biological target) was present in the tissue section at location a1, b4.

[0061] Figure 5 provides a simplified, specific example of the encoding scheme of the assay systems using the invention. Figure 5 shows encoding oligonucleotides 510, comprising a1, a2, a3, a4 and b1, b3, b3 and b4. Target-specific oligonucleotides (TSOs) (probes) 1 and 2 are shown at 520. A deposit or dispensing scheme is shown at 530. Like the grid exemplified in Figure 4, encoding oligonucleotides a1 through a4 are deposited in spots in a pattern (here, in a vertical pattern), and encoding oligonucleotides b1 through b4 are deposited in spots in a pattern (here, a horizontal pattern). The grid though shown as a square with spots is actually a deposition pattern on a biological sample (not shown) such as tissue section 416 shown in Figure 4.

[0062] The target-specific oligonucleotides are delivered to the biological sample, where the target-specific oligonucleotides hybridize to target nucleic acids in the biological sample if target nucleic acids are present. Unhybridized target-specific oligonucleotides are then removed, *e.g.*, by washing. The encoding oligonucleotides are then delivered to the biological sample according to the spatial pattern shown at 530. The encoding oligonucleotides are ligated (or, *e.g.*, extended and ligated) to any target-specific oligonucleotides that hybridized to the target nucleic acid in the biological sample, the ligated constructs are then eluted from the biological sample, pooled, and sequencing adapters are added through, *e.g.,* PCR or ligation, if the sequences were not previously included in the encoding oligonucleotides. The ligated constructs are sequenced by, *e.g.,* high throughput or "next generation" sequencing.

[0063] The pool of resulting sequences is shown at 540. A sequence readout was obtained for target-specific oligonucleotide 1 only at a4b1, a4b2, a1b3, a2b3, a3b3, a4b3 and a4b4 (positions shown with horizontal lines). A sequence readout was obtained for target-specific oligonucleotide 2 only at a1b1 (position shown with vertical lines). A sequence readout was obtained for both target-specific oligonucleotides 1 and 2 at positions a2b1, a3b1, a1b2, a2b2, and a3b2 (positions shown with cross-hatching). No sequence readout was obtained for either target-specific oligonucleotides at a1b4, a2b4 or a3b4 (positions shown without shading). Thus, in the biological sample on which the assay took place the first target nucleic acid was detected in a large portion of the left side and at the bottom of the biological sample, the second target nucleic acid was detected only in the upper left portion of the biological sample, and neither target nucleic acid was detected in the upper right portion of the biological sample. The differential expression of the two target nucleic acids now can be mapped back to the biological sample and to the biological structures or cell types in these locations in the biological sample.

[0064] In addition to location information, information relating to relative abundance of the encoded tags can be obtained. For example, if it is found that there are ten times as many a4T1b1 sequences occurring in the data set as compared to a4T1b2 sequences, this would indicate that target nucleic acid sequence 1 is ten times more abundant at the a4T1b1 location than at the a4T1b2 location.

[0065] In the case of nucleotide analysis as shown in Figure 3, by ligating the coding tags directly to target-specific oligonucleotides, only 2n target-specific oligonucleotides are needed fof h targets. For example, using the combinatorial approach outlined in Figure 2, assaying 100 different targets at 10,000 spatial locations would require 2 x 100 target-specific oligonucleotides and 2 x 100 encoding oligonucleotides: The total count of assay oligonucleotides would be only 400 (200 target-specific and 200 encoding), not counting universal primers. In contrast, if the coding oligonucleotides were hot decoupled from the target-specific oligonucleotides, (n x X positional codes) + (n x Y positional codes) would be needed, or in the above example, 20,000 oligonucleotides, not counting universal primer sequences. Moreover, though the embodiments shown in Figures 2- 5 depict a combinatorial scheme using two encoding agents (coding tags), three, four or more encoding agents and coding tags may be used, and attached to the probe or one another by varying means and in varying combinations of steps.

[0066] Due to the spatial encoding aspect of the assay system, a large amount of information can be generated with even a modest number of assays. For example, five or more biological targets assayed at five or more positions in the sample generates 25 or more combinations. Using digital sequencing as a readout, the optimum number of sequence reads per combination depends on the sensitivity and dynamic range required, and can be adjusted. For example, if for each combination on average 100 reads are sampled, the total for 25 combination is 25,000 reads. If 1,000 targets are assayed at 1,000 locations with an average sampling depth of 1,000, then 10^9 reads are required. These numbers, although large, are within the capacity of intrinsically parallel digital sequencing methods, which can generate datasets of billions or even trillions of reads in a reasonable timeframe and at a very low cost per read. Therefore, by varying the numbers of positions interrogated or biological targets assayed, or both, and using digital sequencing, large amounts of information can be obtained. In specific aspects; multiple locations are interrogated for two or more biological molecules.

Reagent Delivery Systems

[0067] The reagent delivery system includes instrumentation that allows the delivery of reagents to discrete portions of the biological sample, maintaining the integrity of the spatial patterns of the encoding scheme. Reagent delivery systems of the assay systems comprise optional imaging means; reagent delivery hardware and control software. Reagent delivery can be achieved in a number of different ways. It should be noted that reagent delivery may be to many different biological samples at one time. A single tissue section has been exemplified herein; however, multiple biological samples may be affixed and analyzed simultaneously. For example, serial sections of a tissue sample can be analyzed in parallel and the data combined to build a 3D map.

[0068] Integral to the assay system is instrumentation that allows for spatial patterning of reagents onto the biological sample. Technologies for formulating and delivering both biological molecules (*e.g.* oligonucleotides or antibodies) and chemical reagents (*e.g.*, small molecules or dNTPs) are known in the art, and uses of these instrument systems are known to one skilled in the art and easily adaptable to the assay systems. One example of a suitable reagent delivery system is the Labcyte™ Echo acoustic liquid handler, which can be used to deliver nanoliter scale droplets containing biological molecules with high precision and reproducibility. One skilled in the art could incorporate this reagent delivery device into the overall system, using software to specify the locations to which reagents should be delivered.

[0069] Other instruments that can be used for the deposition of agents and/or coding identifiers onto biological samples include, but are not limited to, ink jet spotting; mechanical spotting by means of pin, pen or capillary; micro contact printing; photochemical or photolithographic methods; and the like: For several applications, it may be preferred to segment or sequester certain areas of the biological samples into one or more assay areas for different reagent distributions and/or biological target determination. The assay areas may be physically separated using barriers or channels.

[0070] In one exemplary aspect, the reagent delivery system may be a flow-based system. The flow-based systems for reagent delivery in the present invention can include instrumentation such as one or more pumps, valves, fluid reservoirs, channels, and/or reagent storage cells. Reagent delivery systems are configured to move fluid to contact a discrete section of the biological sample. Movement of the reagents can be driven by a pump disposed, for example, downstream of the fluid reagents. The pump can drive each fluid reagent to (and past) the reaction compartment. Alternatively, reagents may be driven through the fluid by gravity. US Pub. Nos. 20070166725 and 20050239192 disclose certain general-purpose fluidics tools that can be used with the assay systems, allowing for the precise manipulation of gases, liquids and solids to accomplish very complex analytical manipulations with relatively simple hardware.

[0071] In a more specific example, one or more flow-cells can be attached to the substrate-affixed biological sample from above. The flow-cell can include inlet and outlet tubes connected thereto and optionally an external pump is used to deliver reagents to the flow-cell and across the biological sample. The flow cells are configured to deliver reagents only to certain portions of the biological sample, restricting the amount and type of reagent delivered to any specific section of the biological sample.

[0072] In another aspect, a microfluidic system can be integrated into the substrate upon which the biological sample is disposed or externally attached on top of the substrate. Microfluidic passages for holding and carrying fluid may be formed on and/or above the planar substrate by a fluidics layer abutted to the substrate. Fluid reagents can be selected and delivered according to selective opening and closing of valves disposed between reagent reservoirs.

[0073] Pumps generally include any mechanism for moving fluid and/or reagents disposed in fluid. In some examples, the pump can be configured to move fluid and/or reagents through passages with small volumes (*i.e.,* microfluidic structures). The pump can operate mechanically by exerting a positive or negative pressure on fluid and/or on a structure carrying fluid, electrically by appropriate application of an electric field(s), or both, among other means. Exemplary mechanical pumps may include syringe pumps, peristaltic pumps, rotary pumps, pressurized gas, pipettors, etc. Mechanical pumps may be micromachined, molded, etc. Exemplary electrical pumps may include electrodes and may operate by electrophoresis, electroendoosmosis, electrocapillarity, dielectrophoresis (including traveling wave forms thereof), and/or the like.

[0074] Valves generally include any mechanism for regulating the passage of fluid through a channel. Valves can include, for example, deformable members that can be selectively deformed to partially or completely close a channel, a movable projection that can be selectively extended into a channel to partially or completely block a channel, an electrocapillary structure, and/or the like.

[0075] An open gasket can be attached to the top of the biological sample and the sample and reagents can be injected into the gasket. Suitable gasket materials include, but are not limited to, neoprene, nitrile, and silicone rubber. Alternatively, a watertight reaction chamber may be formed by a gasket sandwiched between the biological sample on the substrate and a chemically inert, water resistant material such as, but not limited to, black-anodized aluminum, thermoplastics (*e.g.,* polystyrene, polycarbonate, etc), glass, etc.

[0076] In an optional embodiment, the assay system comprises imaging means to determine features and organization of the biological sample of interest. The images obtained, *e.g.*, may be used to design the deposition pattern of the reagents. Imaging means are optional, as an individual can instead view the biological sample using, *e.g.,* a microscope, analyze the organization of the biological sample, and specify a spatial pattern for delivery assay reagents. If included, the delivery system can comprise a microcircuit arrangement including an image, such as a CCD or IGFET-based (*e.g.*, CMOS-based) imager and an ultrasonic sprayer for reagent delivery such as described in US Pub. No. 20090197326. Also, it should be noted that although Figures 4 and 5 illustrate using a x,y grid configuration, other configurations can be used, such as, *e.g.*, following the topology of a tissue sample; targeting certain groups of cells, cell layers and/or cell types in a tissue, and the like.

[0077] In yet another alternative, the reagent delivery system controls the delivery of reagents to specific patterns on a biological sample surface using semiconductor techniques such as masking and spraying. Specific areas of a biological sample can be protected from exposure to reagents through use of a mask to protect specific areas from exposure. The reagents may be introduced to the biological sample using conventional techniques such as spraying or fluid flow. The use of masked delivery results in a patterned delivery scheme on the substrate surface.

[0078] In a preferred approach, the reagent delivery instrumentation is based on inkjet printing technology. There are a variety of different ink-jetting mechanisms (*e.g.,* thermal, piezoelectric) and compatibility has been shown with aqueous. and organic ink formulations. Sets of independently actuated nozzles can be used to deliver multiple reagents at the same time, and very high resolutions are be achieved.

[0079] In order to target specific sites of interest, an informative image of the biological sample to be assayed may be used to assist in the reagent delivery methods and associated encoding scheme. Sample regions of the biological sample can be identified using image processing (*e.g.*, images of cell types differentiated by immunohistochemistry or other staining chemistries) integrated with other features of the assay system. In some aspects, software is used to automatically translate image information into a reagent delivery pattern. A mechanism to register and align very piecisely the biological sample for reagent delivery is thus an important component of the assay systems using invention. Mechanisms such as the use of fiducial markers on slides and/or other very accurate physical positioning systems can be adapted to this purpose.

[0080] The invention preferably is used with a complete suite of software tailored to the assay system. Optionally, oligonucleotide design software is used to design the encoding nucleotides (and in embodiments where nucleic acids are assayed, the target-specific oligonucleotides) for the specific assay to be run, and may be integrated as a part of the system. Also optionally, algorithms and software for reagent delivery and data analysis (*i.e.*, sequence analysis) may be integrated to determine assay results. Integrated data analysis is particularly useful, as the type of dataset that is generated may be massive as a consequence of scale. Algorithms and software tools that are specifically designed for analysis of the spatially-associated data generated by the assay systems, including pattern-analysis software arid visualization tools, enhance the value of the data generated by the assay systems.

[0081] In certain aspects, the assay system comprises processes for making arid carrying out the quality control of reagents, *e.g.,* the integrity and sequence fidelity of oligonucleotide pools: In particular, reagents are formulated according to factors such as volatility, stability at key temperatures, and chemical compatibility for compatibility with the reagent delivery instrumentation and may be analyzed by instrumentation integrated within the assay system.

Sequencing

[0082] Numerous methods can be used to identify the coding tags and probe sequences in the encoded probes of the assay systems. The coding tags can be detected using techniques such as mass spectroscopy (*e.g.*, Maldi-T of, LC-MS/MS), nuclear magnetic resonance imaging, or, preferably, nucleic acid sequencing. Examples of techniques for decoding the coding tags can be found, for example, in US Pub. No. 20080220434. For example, the coding tags may be oligonucleotide mass tags (OMTs or massTags). Such tags are described, *e.g.,* in US Pub. No. 20090305237, . In yet another alternative, the encoded probes can be amplified and hybridized to a microarray. This would require separate amplification reactions to be carried out, in which each amplification is specific to a particular spatial code or subset of codes, accomplished by using code-specific primers. Each amplification would also incorporate a different resolvable label (e.g. fluorophor). Following hybridization, the relative amounts or a particular target mapping to different-spatial locations in the sample can be determined by the relative abundances of the resolvable labels.

[0083] In one particularly preferred aspect, the resulting coding tags according to the assay system are substrates for high-throughput, next-generation sequencing, and highly parallel next-generation sequencing methods are used to conform the sequence of the coding tags, for example, with SOLiD™ technology (Life Technologies, Inc.) or Genome Analyzer (Illumina, Inc.). Such next-generation sequencing methods can be carried out, for example, using a one pass sequencing method or using paired-end sequencing. Next generation sequencing methods include, but are not limited to, hybridization-based methods, such as disclosed in *e.g.,* Drmanac, U.S. Pat. Nos. 6,864,052; 6,309,824; and 6,401,267; and Drmanac et al, U.S. patent publication 2005/0191656; sequencing-by-synthesis methods, *e.g.*, U.S. Pat. Nos. 6,210,891; 6,828,100; 6,969,488; 6,897,023; 6,833,246; 6,911,345; 6,787,308; 7,297,518; 7,462,449 and 7,501,245; US Publication Application Nos. 20110059436; 20040106110; 20030064398; and 20030022207; Ronaghi, et al, Science, 281: 363-365 (1998); and Li, et al, Proc. Natl. Acad. Sci., 100: 414-419 (2003); ligation-based methods, *e.g.*, U.S. Patient Nos. 5,912,148 and 6,130,073; and U.S. Pat. Appln Nos. 20100105052, 20070207482 and 20090018024; nanopore sequencing *e.g.*, U.S. Pat. Appln Nos. 20070036511; 20080032301; 20080128627; 20090082212; and Soni and Meller, Clin Chem 53: 1996-2001 (2007)), as well as other methods, *e.g.*, U.S. Pat. Appln Nos. 20110033854; 20090264299; 20090155781; and 20090005252; also, see, McKernan, et al., Genome Res., 19:1527-41 (2009) and Bentley, et al., Nature 456:53-59 (2008).

Applications of Assay System

[0084] It will be apparent to one skilled in the art upon reading the present disclosure that there are numerous important areas of biological research, diagnostics, and drug development that will benefit from a high throughput multiplexed assay system that can measure simultaneously the amount and spatial location of a biological target in a biological sample. For example, combining the agility to estiinate the relative abundance of different RNA transcripts with the ability to reconstruct an image of spatial patterns of abundance across many locations, which may be as small as or seven smaller than individual cells, in a tissue enables many different areas of basic research. The following are Exemplary uses and are by no means meant to be limiting in scope.

[0085] In one example, 3-dimensional patterns of gene expression are determined by analyzing a series of tissue sections, in a manner analogous to image reconstruction in CT scanning. Such a method can be used to measure changes in gene expression in disease pathology, *e.g.*, in cancerous tissue and/or a tissue upon injury, inflammation or inflection. With the assay systems using the invention, more detailed information on gene expression and protein localization in complex tissues is obtained, leading to new insights into the function and regulation both in normal and diseased states, and provides new hypotheses that can be tested. For example, an assay system using the invention may enable some of the insights gained from many individual studies and larger programs like ENCODE (Birney, et al., Nature, 447:799-816 (2007)) and modENCODE to be integrated at the tissue level. The assay systems also aid computational efforts to model interacting networks of gene expression in the field of systems biology.

[0086] The assay systems also provide a novel approach to analysis of somatic variation, *e.g.*; somatic mutations in cancer of variability in response to infections organisms. For example, tumors are typically highly heterogeneous, containing cancer cells as well as genetically normal cells in an abnormal local environment. Cancer cells undergo mutation and selection, and in this process it is not unusual for local clones to develop. Identifying relatively rare somatic mutations in the context of tumors may enable the study of the role of key mutations in the selection of clonal variants. Transcriptional patterns associated with angiogenesis, inflammation, or other cancer-related processes in both cancer and genetically normal cells can be analyzed for insights into cancer biology and assist in the development of new therapeutic agents for the treatment of cancers. In another example, individuals have varying susceptibility to infectious organisms, and the assay systems using the invention can be used to study the interaction between microbes and tissues or the various cell types within the tissue.

[0087] Importantly, in addition to providing spatially-associated information, the invention allows a great increase in the sensitivity of detecting rare mutations, as signal to noise can be dramatically increased since only a small location is assayed in any given reaction. In a typical assay for rare mutations in a mixed sample, the sample is treated in bulk, *i.e.,* nucleic acids are extracted from many cells into a single pool. Thus, if a mutation is present in one cell in 10,000, it must be detected against a background of normal DNA from ~10,000 cells. In contrast, with the assay systems using the invention many cells can be analyzed, but individual cells or small groups of cells would be identified by the spatial coding system. Therefore, in the assay systems using the present invention, background is reduced by orders of magnitude, greatly increasing sensitivity. Furthermore, the spatial organization of mutant cells can be observed, which may be particularly important in detecting key mutations in tissue sections in cancer. Already molecular histological analyses are yielding insights into cancer biology and may have potential for use in diagnostics. The technology of the invention promises to greatly increase the power of such approaches.

EXAMPLES

[0088] The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the present invention, and are not intended to limit the scope of what the inventor regards as his invention, nor are they intended to represent or imply that the experiments below are all of or the only experiments performed. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

[0089] Efforts have been made to ensure accuracy with respect to numbers used (*e.g.*, amounts, temperature, etc.) but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, molecular weight is weight average molecular weight, temperature is in degrees centigrade, and pressure is at or near atmospheric.

EXAMPLE 1: Initial Proof of concept of Encoding Scheme

[0090] As an initial proof of concept, a model system is developed using a microarray to demonstrate a working single-plex assay. The basic design validates the concept of the assay, and establishes a working assay prior to addressing issues related to the analysis of a more complicated biological sample. Conventional sequencing is used as a readout for this proof of concept.

[0091] A microarray is used as a proxy for a tissue section. The target sequences of the microarray are fully specified, so that the composition of the targets are known and can be varied systematically. Synthetic oligonucleotide templates are attached to a glass slide via a 5' amino modification. Each slide has a single oligonucleotide template sequence, and the assays that are carried out may employ either ligation, or extension followed by ligation as this may be useful in determining certain polymorphisms.

[0092] Once the *in situ* part of the assay is complete, the reaction products are eluted and analyzed by qPCR to determine presence or absence of a product and estimate yield, and by conventional sequencing to determine the structure of the assay products. The single-plex assays that are tested include appropriate positive and negative controls, and a single nucleotide variant (SNV) to check ability to discriminate single base variants.

EXAMPLE 2: Scalability

[0093] The complexity of the assay system is increased to establish scalability of the assay for use in high throughput studies. Scalability of both the spatial encoding and assay systems is demonstrated by carrying out a 24-plex x 24-site assay using a microarray model system.

[0094] The amount of biological target, here a DNA target sequence, at each assay location is systematically varied on microarray substrate. For example, in a microarray with 50 micron spot size (center to center), a 1mm^2 area contains ~400 spots. The region around each site is optionally occupied by a region that is devoid of these spots to allow individual resolvability of the target sequences. Alternatively, the spots may be clustered, with two or more directly adjacent spots surrounded by or adjacent to a region that is devoid of target sequences.

[0095] In order to demonstrate that spatial encoding is accurate, the sites comprise different target compositions to show that the assay readout matches the expected composition of each site. With 24 target sequences, a simple digital pattern is made with each site having a different set of 12 targets present and 12 targets absent, to make a binary code (0 = absent, 1 = present). The assay readout is then determined to show that the detected regions match the expected signal after spatial decoding. In this particular example, the code space is large enough (2^{24}) so that even a few errors would not result in different codes being mixed up. Moreover, this design allows identification of errors and allows an estimation not only of accuracy of spatial encoding but also of accuracy calling the presence or absence of target sequences.

[0096] The ability to detect quantitative differences is evaluated by generating dose-response curves for each of the 24 assays that are carried out at each site in a 24-site assay. This allows estimation of the limit of detection, dynamic range, and power to detect a given fold-change across the range.

[0097] In one aspect, a latin square design is used to represent individual targets at different ratios by varying the number of features for each target. In other words, with multiple spots in a site, the number of spots allocated to each of the 24 target sequences can be varied and each of the 24 sites can have a different composition. A 1 x 3 inch microarray is sufficiently large to permit multiple replicates. This larger set of 24 sequences will require deconvolution, and this is accomplished using high throughput techniques such as next-generation sequencing technologies (*e.g.*, SOLiD™ technology (Life Technologies, Inc., Carlsbad, CA) or Genome Analyzer (Illumina, Inc., San Diego, CA)). The use of the 24-plex assay demonstrates both the accuracy of spatial encoding and decoding, and the quantitative response of the assay system.

EXAMPLE 3: Adaptation of the assay to preserved samples.

[0098] Genomic DNA is assayed as a proof of concept for assaying RNA, as it provides a way to establish a single-copy reference signal. Once a working assay is developed for FFPE samples, it is adapted to an RNA assay. To this end, assay oligonucleotide concentrations are assayed to ensure compatibility with high multiplexing. Assuming a cell diameter of 10 microns, and delivery of a 10 micron diameter reagent droplet to an individual cell, the volume of the droplet will be ~500 μl and can contain ~3 x 10^{11} molecules at a 1 μM concentration. To assay 1,000 target sequences in 10,000 cells, ~2,000 targeting oligonucleotides would be required in a droplet. Therefore, each droplet could contain ~160 million copies of each assay oligo, a vast excess over the few thousand target sequences in a cell.

[0099] The handling of small absolute numbers of product molecules generated from very small or compromised samples are enhanced to counter the issue of low recovery efficiency; that is, elution efficient and losses resulting from absorption of molecules to surfaces are prevented. An approach to addressing the latter issue is to include a carrier material, such as glycogen or carrier nucleic acids.

Example 4: Adapting the Assay to a Biological Sample.

[0100] A control RNA template is immobilized to a solid support in order to create an artificial systems. The assay is performed using T4 DNA ligase, which can repair nicks in DNA/RNA hybrids. Assays are carried out on matched slides, or different sections of the same slide, where in one case gDNA is assayed and in the other RNA is assayed. When assaying gDNA the slide can be pretreated with RNase, and when assaying RNA the slide is pretreated with DNase. Results of the assay are confirmed by extracting dDNA or RNA and quantitating the relative amounts by qPCR or RT-qPCR respectively.

[0101] In order make the tissue section RNA assays as informative as possible, pre-existing information on expression levels in specific tissues to target transcripts across a range of abundances are used in the assay design. Both high abundance transcripts, as well as some medium and low abundance transcripts, are targeted to enable an initial assessment of the quantitative performance characteristics of the assay.

[0102] The preceding merely illustrates the principles of the invention. Furthermore, all examples and conditional language recited herein are principally intended to aid the reader in understanding the principles of the invention and the concepts contributed by the inventors to furthering the art, and are to be construed as being without limitation to such specifically recited examples and conditions. Moreover, all statements herein reciting principles, aspects, and embodiments of the invention as well as specific examples thereof, are intended to encompass both structural and functional equivalent thereof. Additionally, it is intended that such equivalents include both currently known equivalents and equivalents developed in the future, i.e., any elements developed that performs the same function, regardless of structure. The scope of the present invention, therefore, is not intended to be limited to the exemplary embodiments shown any described herein.

Claims

1. A method to determine spatial patterns of abundance or activity or both of multiple nucleic acid targets at multiple sites in a sample, comprising the following steps:

providing a sample affixed to a support;
delivering oligonucleotide probes for multiple nucleic acid targets to the multiple sites in the sample;
allowing the oligonucleotide probes to hybridize with the nucleic acid targets;
washing the unhybridized oligonucleotide probes from the sample;
delivering encoding agents to locations of the multiple sites in the sample according to a known spatial pattern, wherein at least two encoding agents are delivered to each of the multiple sites and the combination of the encoding agents delivered to each site is different;
coupling the encoding agents to the oligonucleotide probes hybridized to the nucleic acid targets to form encoded probes;
determining all or a portion of a sequence of the encoded probes; and
associating the abundance or activity or both of the multiple biological targets to the locations of multiple sites in the sample.

2. A method to determine spatial patterns of abundance or activity or both of multiple biological targets at multiple sites in a sample, comprising the following steps:

providing a sample affixed to a support;
delivering encoded probes for multiple biological targets in a known spatial pattern to the multiple sites in the sample, wherein each encoded probe comprises a probe region capable of interaction with a biological target and a coding tag comprising an oligonucleotide that identifies the location of the site to which the encoded probe is delivered;
allowing the encoded probes to interact with the biological targets;
determining a sequence of the encoded probes, wherein the sequence comprises the coding tag oligonucleotide sequence, thereby identifying the location of the site to which the encoded probe is delivered; and
associating the abundance or activity or both of the multiple biological targets to the locations of the sites in the sample.

3. The method of claim 2, further comprising after the allowing step and before the determining step a step of separating the encoded probes that interact with the biological targets from the encoded probes that do not interact with the biological targets.

4. The method of claim 3, wherein the biological targets are nucleic acids and the encoded probes are oligonucleotide probes.

5. The method of claim 1 or 4, wherein there are two oligonucleotide probes for each of the nucleic acid biological targets.

6. The method of claim 3, wherein the multiple biological targets are proteins, the probe regions of the encoding probes are proteins and the coding tags comprise oligonucleotides.

7. The method of claim 3, wherein the multiple biological targets comprise enzymes.

8. The method of claim 3, wherein the probe regions of the encoded probes comprise antibodies, aptamers, small molecules, enzyme substrates, putative enzyme substrates, or affinity capture agents.

9. The method of any one of claims 1, 3, and 6, further comprising an amplification step between the separating step or the coupling step and the determining step.

10. The method of any one of claims-1-3, wherein the determining step is performed by nucleic acid sequencing or high-throughput sequencing.

11. The method of claim 1, wherein the coupling step is performed by ligation or by extension followed by ligation.

12. The method of any one of claims 1-3, wherein the product of the multiple biological targets being assayed and the multiple sites in the sample is greater than 1,000,000.

13. The method of any one of claims 1-3 and 6, wherein the sequences of at least one million encoding probes are determined in parallel.

14. The method of any one of claims 1-3, wherein the known spatial pattern is determined by histological features of the sample.

15. The method of any one of claims 1-3, wherein software programmed hardware performs at least two steps of the delivering step, the separation step, the determining step and the associating step.

Patentansprüche

1. Verfahren zur Bestimmung räumlicher Häufigkeits- und/oder Aktivitätsmuster von mehreren Nucleinsäurezielen an mehreren Stellen in einer Probe, wobei das Verfahren folgende Schritte umfasst:

das Bereitstellen einer an einem Träger befestigten Probe;

das Zuführen von Oligonucleotidsonden für mehrere Nucleinsäureziele zu den mehreren Stellen in der Probe;
das Zulassen des Hybridisierens der Oligonucleotidsonden mit den Nucleinsäurezielen;
das Auswaschen der nicht hybridisierten Oligonucleotidsonden aus der Probe;
das Zuführen von Kodierungsmitteln an die Orte der mehreren Stellen in der Probe gemäß einem bekannten räumlichen Muster, wobei zumindest zwei Kodierungsmittel jeder der mehreren Stellen zugeführt werden und die Kombination der jeder Stelle zugeführten Kodierungsmittel unterschiedlich ist;
das Kuppeln der Kodierungsmittel an die an die Nucleinsäureziele hybridisierten Oligonucleotidsonden, um kodierte Sonden zu bilden;
das Bestimmen der vollständigen Sequenz oder eines Teils der Sequenz der kodierten Sonden und
das Zuordnen von Häufigkeit und/oder Aktivität der mehreren biologischen Ziele zu den Orten der mehreren Stellen in der Probe.

2. Verfahren zur Bestimmung räumlicher Häufigkeits- und/oder Aktivitätsmuster von mehreren Nucleinsäurezielen an mehreren Stellen in einer Probe, wobei das Verfahren folgende Schritte umfasst:

das Bereitstellen einer an einem Träger befestigten Probe;
das Zuführen kodierter Sonden für mehrere biologische Ziele in einem bekannten räumlichen Muster an mehrere Stellen in der Probe, worin jede kodierte Sonde eine Sondenregion, die in der Lage ist, mit einem biologischen Ziel in Wechselwirkung zu treten, und eine Kodierungsmarkierung umfasst, die ein Oligonucleotid umfasst, das den Ort der Stelle identifiziert, welcher die kodierte Sonde zugeführt wird;
das Zulassen des Wechselwirkens der kodierten Sonden mit den biologischen Zielen;
das Bestimmen der Sequenz der kodierten Sonden, worin die Sequenz die Oligonucleotidsequenz der Kodierungsmarkierung umfasst, wodurch der Ort der Stelle identifiziert wird, der die kodierte Sonde zugeführt wird; und
das Zuordnen von Häufigkeit und/oder Aktivität der mehreren biologischen Ziele zu den Orten der mehreren Stellen in der Probe.

3. Verfahren nach Anspruch 2, das ferner nach dem Schritt des Zulassens und vor dem Schritt des Bestimmens einen Schritt des Trennens der kodierten Sonden, welche mit den biologischen Zielen in Wechselwirkung treten, von jenen kodierten Sonden umfasst, die nicht mit den biologischen Zielen in Wechselwirkung treten.

4. Verfahren nach Anspruch 3, worin die biologischen Ziele Nucleinsäuren sind und die kodierten Sonden Oligonucleotidsonden sind.

5. Verfahren nach Anspruch 1 oder 4, worin zwei Oligonucleotidsonden für jedes biologische Nucleinsäureziel vorliegen.

6. Verfahren nach Anspruch 3, worin die mehreren biologischen Ziele Proteine sind, die Sondenregionen der kodierten Sonden Proteine sind, und die Kodierungsmarkierungen Oligonucleotide umfassen.

7. Verfahren nach Anspruch 3, worin die mehreren biologischen Ziele Enzyme umfassen.

8. Verfahren nach Anspruch 3, worin die Sondenregionen der kodierten Sonden Antikörper, Aptamere, kleine Moleküle, Enzymsubstrate, mutmaßliche Enzymsubstrate oder Affinitätseinfangmittel umfassen.

9. Verfahren nach einem der Ansprüche 1, 3 und 6, das weiters einen Amplifizierungsschritt zwischen dem Trennschritt oder dem Kupplungsschritt und dem Bestimmungsschritt umfasst.

10. Verfahren nach einem der Ansprüche 1 bis 3, worin der Bestimmungsschritt durch Nucleinsäuresequenzierung oder Hochdurchsatzsequenzierung durchgeführt wird.

11. Verfahren nach Anspruch 1, worin der Kupplungsschritt durch Ligation oder Erweiterung gefolgt von Ligation durchgeführt wird.

12. Verfahren nach einem der Ansprüche 1 bis 3, worin das Produkt der mehreren biologischen Ziele getestet wird und die mehreren Stellen in der Probe mehr als 1.000.000 sind.

13. Verfahren nach einem der Ansprüche 1 bis 3 und 6, worin die Sequenzen von zumindest einer Million kodierenden Sonden gleichzeitig bestimmt werden.

14. Verfahren nach einem der Ansprüche 1 bis 3, worin das bekannte räumliche Muster anhand von histologischen Merkmalen der Probe bestimmt wird.

15. Verfahren nach einem der Schritte 1 bis 3, worin mittels Software programmierte Hardware zumindest zwei Schritte ausgewählt aus Zuführungsschritt, Trennschritt, Bestimmungsschritt und Zuordnungsschritt ausführt.

Revendications

1. Procédé de détermination de motifs spatiaux

d'abondance ou d'activité ou des deux de cibles d'acide nucléique multiples à des sites multiples dans un échantillon, comprenant les étapes suivantes :

réaliser un échantillon fixé à un support ;
délivrer des sonde d'oligonucléotides pour des cibles d'acide nucléique multiples aux sites multiples dans l'échantillon ;
permettre aux sondes d'oligonucléotide d'hybrider avec les cibles d'acide nucléique ;
laver les sondes d'oligonucléotide non hybridées de l'échantillon ;
délivrer des agents de codage aux emplacements des sites multiples dans l'échantillon en accord avec un motif spatial connu, où au moins deux agents de codage sont délivrés à chacun des sites multiples et la combinaison des agents de codage délivrés à chaque site est différente ;
coupler les agents de codage aux sondes d'oligonucléotides hybridées aux cibles d'acide nucléique pour former des sondes codées ;
déterminer l'ensemble ou une portion d'une séquence des sondes codées ; et
associer l'abondance ou l'activité ou les deux des cibles biologiques multiples aux emplacements de sites multiples dans l'échantillon.

2. Procédé de détermination de motifs spatiaux d'abondance ou d'activité ou des deux de cibles d'acide nucléique multiples à des sites multiples dans un échantillon, comprenant les étapes suivantes :

réaliser un échantillon fixé à un support ;
délivrer des sondes codées pour des cibles biologiques multiples en un motif spatial connu aux sites multiples dans l'échantillon, où chaque sonde codée comprend une région de sonde apte à interagir avec une cible biologique et un marqueur de codage comprenant un oligonucléotide qui identifie l'emplacement du site auquel le sonde codée est délivrée ;
permettre aux sondes codées d'interagir avec les cibles biologiques ;
déterminer une séquence de sondes codées, où la séquence comprend la séquence d'oligonucléotides à marqueur de codage, en identifiant ainsi l'emplacement du site auquel la sonde codée est délivrée ; et
associer l'abondance ou l'activité ou les deux des cibles biologiques multiples aux emplacements des sites dans l'échantillon.

3. Procédé selon la revendication 2, comprenant en outre, après l'étape de permission et avant l'étape de détermination, une étape de séparation des sondes codées qui interagissent avec les cibles biologiques des sondes codées qui n'interagissent pas avec les cibles biologiques.

4. Procédé selon la revendication 3, dans lequel les cibles biologiques sont des acides nucléiques et les sondes codées sont des sondes d'oligonucléotides.

5. Procédé selon la revendication 1 ou 4, dans lequel il y a deux sondes d'oligonucléotides pour chacune des cibles biologiques d'acide nucléique.

6. Procédé selon la revendication 3, dans lequel les cibles biologiques multiples sont des protéines, les régions de sonde des sondes de codage sont des protéines et les marqueurs de codage comprennent des oligonucléotides.

7. Procédé selon la revendication 3, dans lequel les cibles biologiques multiples comprennent des enzymes.

8. Procédé selon la revendication 3, dans lequel les régions de sonde des sondes codées comprennent des anticorps, aptamères, petites molécules, substrats d'enzymes, substrats d'enzymes putatives ou agents de capture d'affinité.

9. Procédé selon l'une quelconque des revendications 1, 3 et 6, comprenant en outre une étape d'amplification entre l'étape de séparation ou l'étape de couplage et l'étape de détermination.

10. Procédé selon l'une quelconque des revendications 1 à 3, dans lequel l'étape de détermination est exécutée par séquençage d'acide nucléique ou séquençage à haut débit.

11. Procédé selon la revendication 1, dans lequel l'étape de couplage st exécutée par ligature ou par extension suivie de ligature.

12. Procédé selon l'une quelconque des revendications 1 à 3, où le produit des cibles biologiques multiples testées et des sites multiples dans l'échantillon est plus grand que 1.000.000.

13. Procédé selon l'une quelconque des revendications 1 à 3 et 6, dans lequel les séquences d'au moins un million de sondes de codage sont déterminées en parallèle.

14. Procédé selon l'une quelconque des revendications 1 à 3, dans lequel le motif spatial connu est déterminé par des propriétés histologiques de l'échantillon.

15. Procédé selon l'une quelconque des revendications 1 à 3, dans lequel le matériel programmé par un logiciel exécute au moins deux étapes de l'étape de délivrance, l'étape de séparation, l'étape de détermination et l'étape d'association.

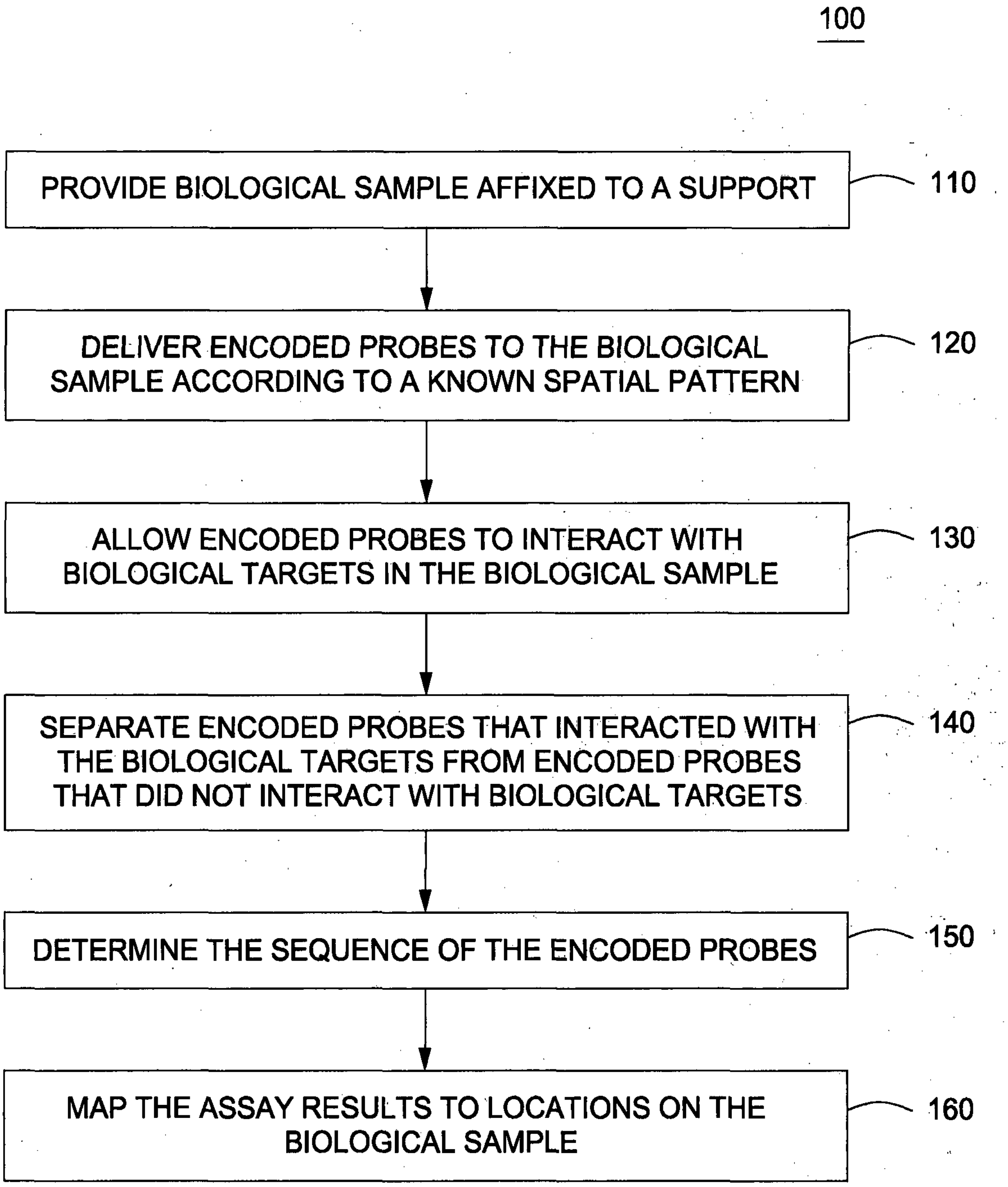
100
PROVIDE BIOLOGICAL SAMPLE AFFIXED TO A SUPPORT
110
DELIVER ENCODED PROBES TO THE BIOLOGICAL SAMPLE ACCORDING TO A KNOWN SPATIAL PATTERN
120
ALLOW ENCODED PROBES TO INTERACT WITH BIOLOGICAL TARGETS IN THE BIOLOGICAL SAMPLE
130
SEPARATE ENCODED PROBES THAT INTERACTED WITH THE BIOLOGICAL TARGETS FROM ENCODED PROBES THAT DID NOT INTERACT WITH BIOLOGICAL TARGETS
140
DETERMINE THE SEQUENCE OF THE ENCODED PROBES
150
MAP THE ASSAY RESULTS TO LOCATIONS ON THE BIOLOGICAL SAMPLE
160

FIG. 1

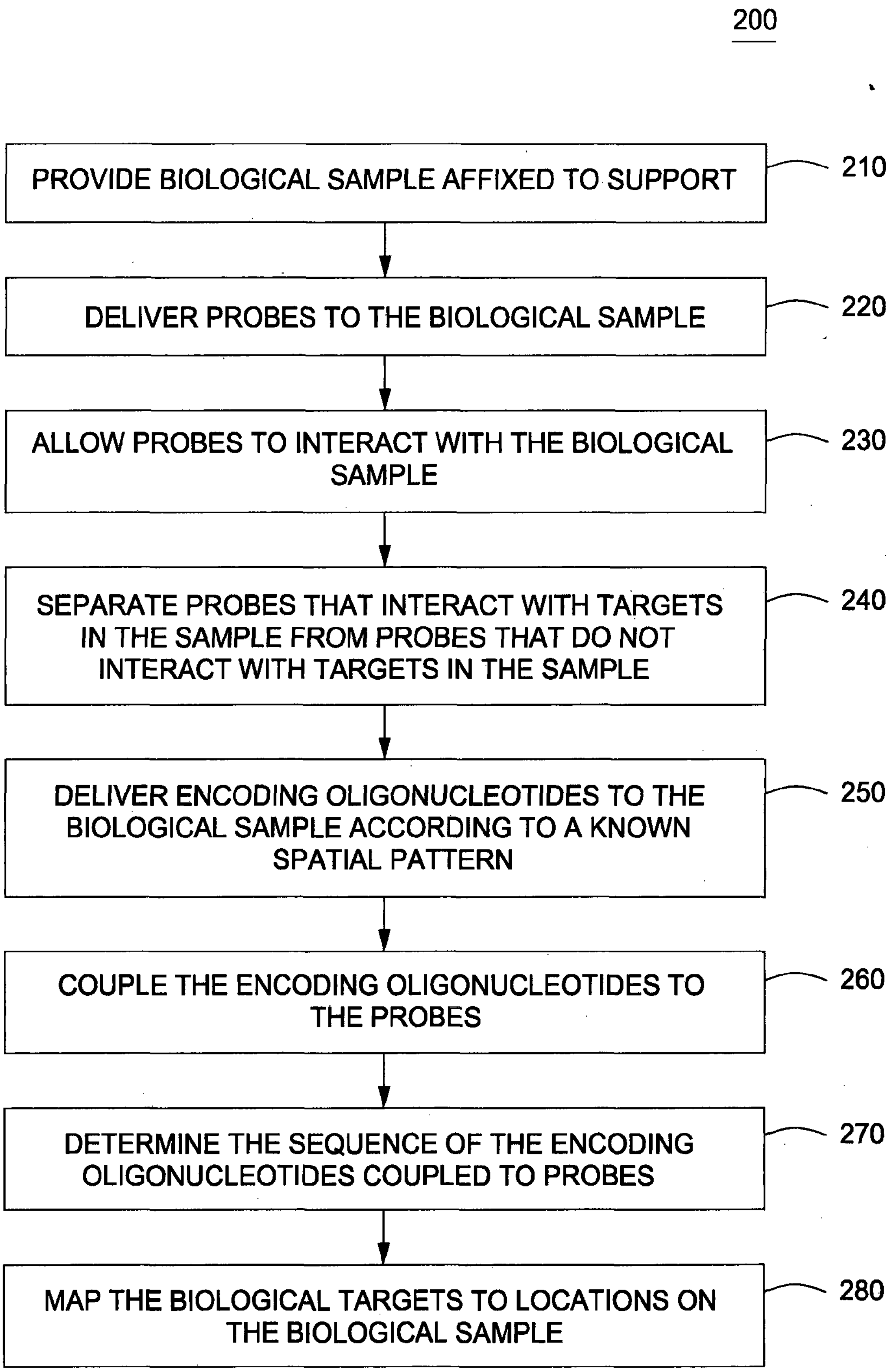
200
PROVIDE BIOLOGICAL SAMPLE AFFIXED TO SUPPORT
210
DELIVER PROBES TO THE BIOLOGICAL SAMPLE
220
ALLOW PROBES TO INTERACT WITH THE BIOLOGICAL SAMPLE
230
SEPARATE PROBES THAT INTERACT WITH TARGETS IN THE SAMPLE FROM PROBES THAT DO NOT INTERACT WITH TARGETS IN THE SAMPLE
240
DELIVER ENCODING OLIGONUCLEOTIDES TO THE BIOLOGICAL SAMPLE ACCORDING TO A KNOWN SPATIAL PATTERN
250
COUPLE THE ENCODING OLIGONUCLEOTIDES TO THE PROBES
260
DETERMINE THE SEQUENCE OF THE ENCODING OLIGONUCLEOTIDES COUPLED TO PROBES
270
MAP THE BIOLOGICAL TARGETS TO LOCATIONS ON THE BIOLOGICAL SAMPLE
280

FIG. 2

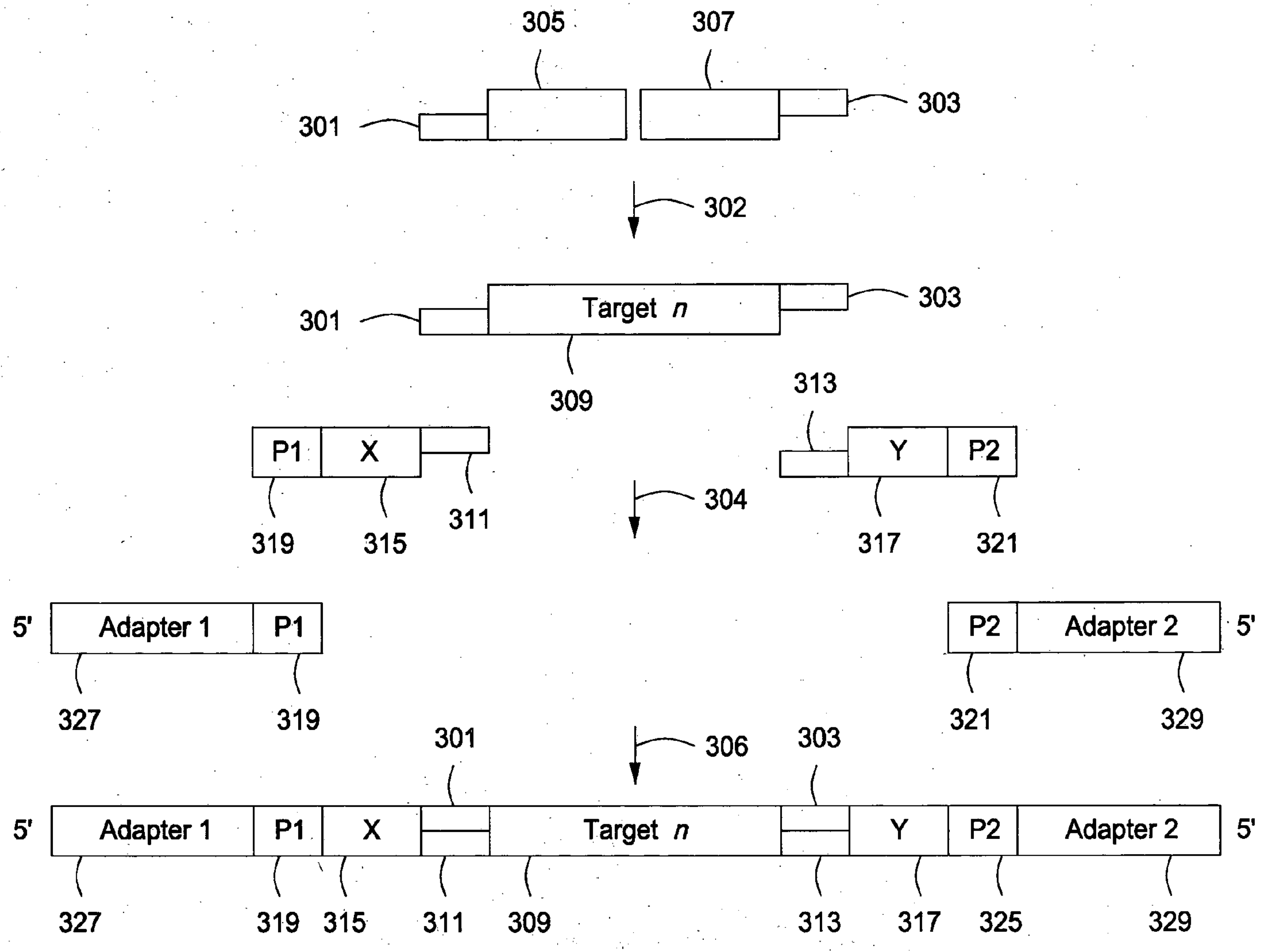
305
307
301
303
302
Target n
301
303
309
313
P1
X
Y
P2
319
315
311
304
317
321
5'
Adapter 1
P1
P2
Adapter 2
5'
327
319
321
329
301
306
303
5'
Adapter 1
P1
X
Target n
Y
P2
Adapter 2
5'
327
319
315
311
309
313
317
325
329

FIG. 3

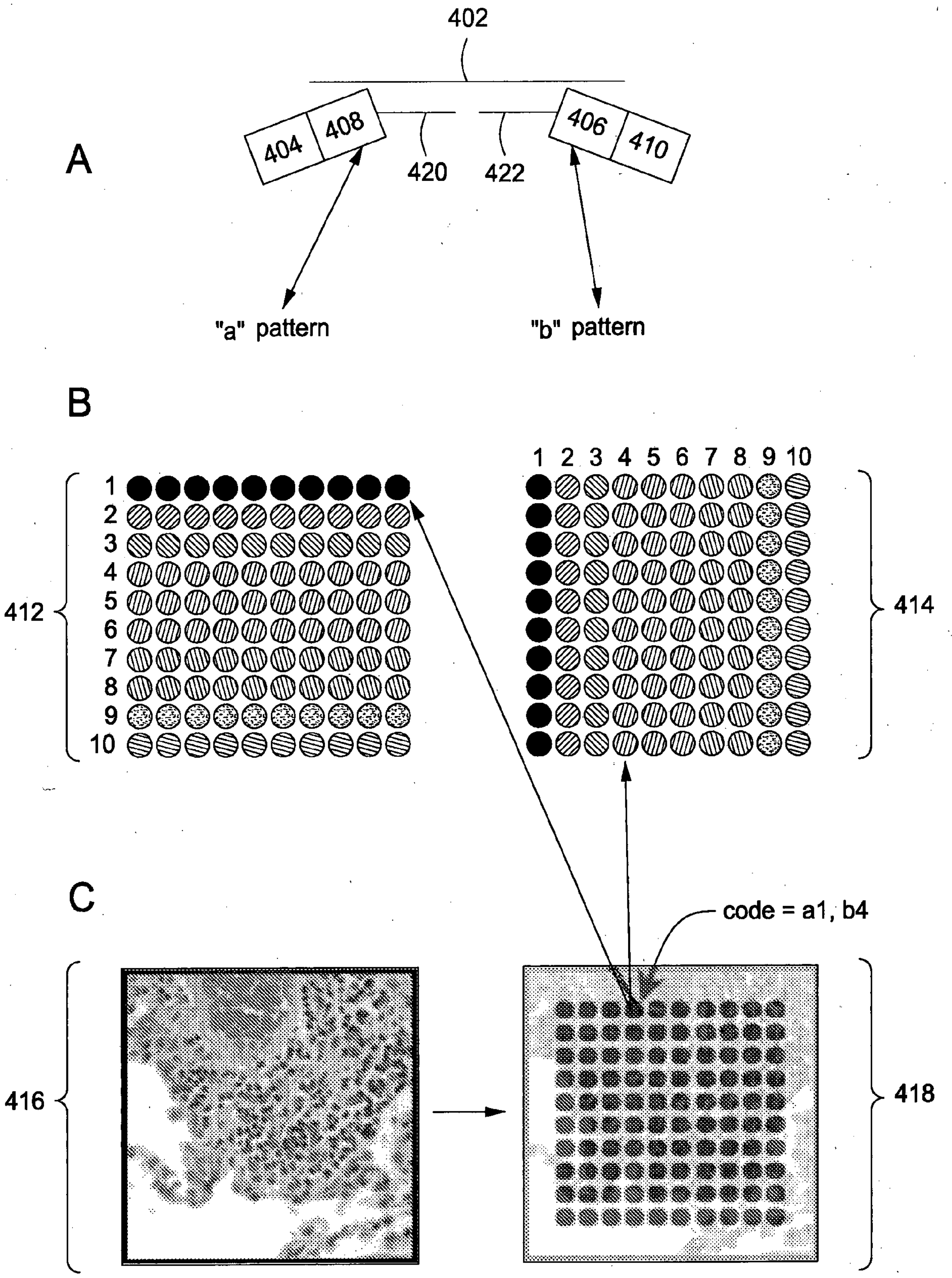
A
402
404
408
406
410
420
422
"a" pattern
"b" pattern
B
1
2
3
4
5
6
7
8
9
10
1 2 3 4 5 6 7 8 9 10
412
414
C
code = a1, b4
416
418

FIG. 4

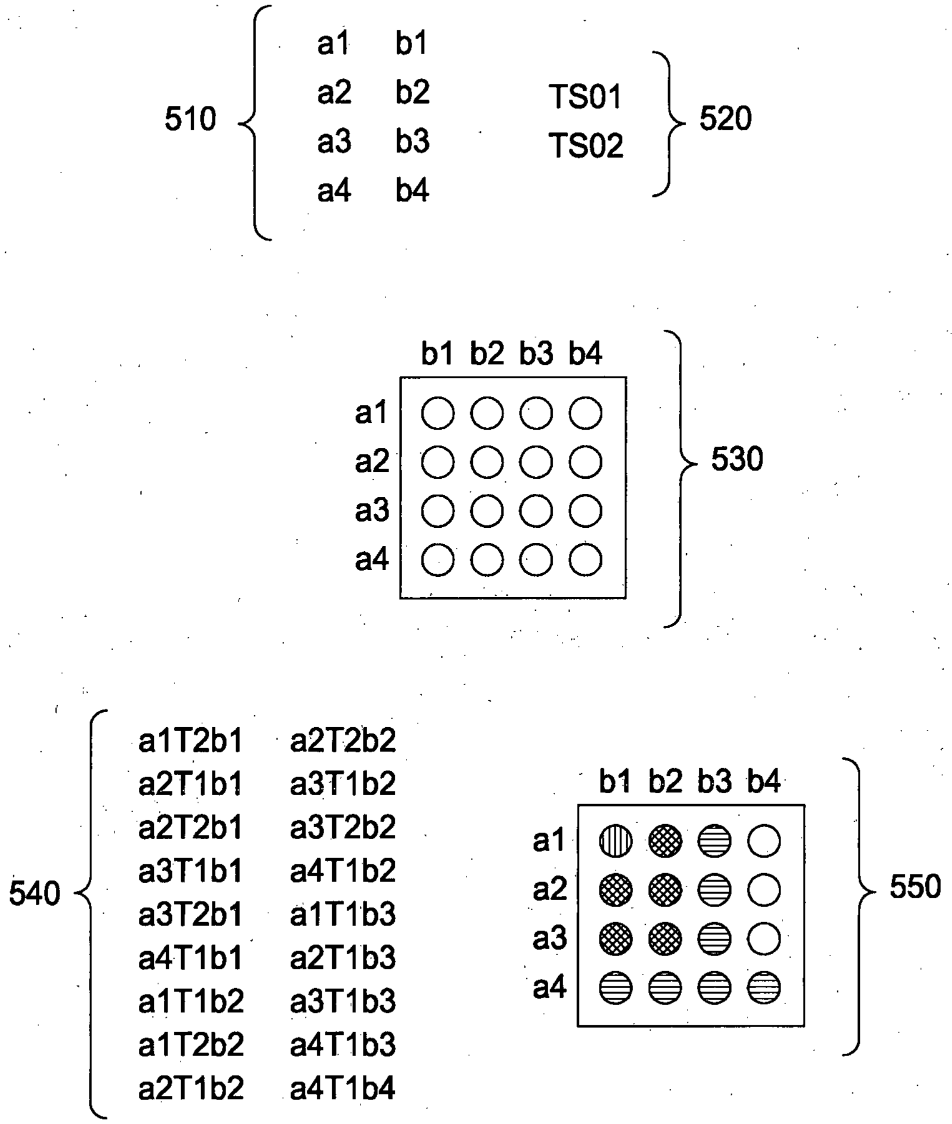
510
a1 b1
a2 b2
a3 b3
a4 b4
TS01
TS02
520
b1 b2 b3 b4
a1
a2
a3
a4
530
540
a1T2b1
a2T1b1
a2T2b1
a3T1b1
a3T2b1
a4T1b1
a1T1b2
a1T2b2
a2T1b2
a2T2b2
a3T1b2
a3T2b2
a4T1b2
a1T1b3
a2T1b3
a3T1b3
a4T1b3
a4T1b4
b1 b2 b3 b4
a1
a2
a3
a4
550

FIG. 5

REFERENCES CITED IN THE DESCRIPTION

Patent documents cited in the description

- WO 2010019826 A **[0005]**
- US 6969589 B **[0027]**
- US 20060188875 A **[0027]**
- US 20070166725 A **[0070]**
- US 20050239192 A **[0070]**
- US 20090197326 A **[0076]**
- US 20080220434 A **[0082]**
- US 20090305237 A **[0082]**
- US 6864052 B, Drmanac **[0083]**
- US 6309824 B **[0083]**
- US 6401267 B **[0083]**
- US 20050191656 A, Drmanac **[0083]**
- US 6210891 B **[0083]**
- US 6828100 B **[0083]**
- US 6969488 B **[0083]**
- US 6897023 B **[0083]**
- US 6833246 B **[0083]**
- US 6911345 B **[0083]**
- US 6787308 B **[0083]**
- US 7297518 B **[0083]**
- US 7462449 B **[0083]**
- US 7501245 B **[0083]**
- US 20110059436 A **[0083]**
- US 20040106110 A **[0083]**
- US 20030064398 A **[0083]**
- US 20030022207 A **[0083]**
- US 5912148 A **[0083]**
- US 6130073 A **[0083]**
- US 20100105052 A **[0083]**
- US 20070207482 A **[0083]**
- US 20090018024 A **[0083]**
- US 20070036511 A **[0083]**
- US 20080032301 A **[0083]**
- US 20080128627 A **[0083]**
- US 20090082212 A **[0083]**
- US 20110033854 A **[0083]**
- US 20090264299 A **[0083]**
- US 20090155781 A **[0083]**
- US 20090005252 A **[0083]**

Non-patent literature cited in the description

- **SHI et al.** *Nature Biotechnology,* 2006, vol. 24 (9), 1151-61 **[0003]**
- **SLONIM ; YANAI.** *Plos Computational Biology,* 2009, vol. 5 (10), 1000543 **[0003]**
- **VELCULESCU et al.** *Science,* 1995, vol. 270 (5235), 484-87 **[0003]**
- **SPURGEON et al.** *Plos ONE,* 2008, vol. 3 (2), 1662 **[0003]**
- **NUOVO.** *Genome Res,* 1995, vol. 4, 151-67 **[0003]**
- **ARMANI et al.** *Lab on a Chip,* 2009, vol. 9 (24), 3526-34 **[0003]**
- **MIR KALIM et al.** NUCLEIC ACIDS RESEARCH. OXFORD UNIVERSITY PRESS, GB, 01 January 2009, vol. 37, e5-1 **[0006]**
- **ANDERSON ; YOUNG.** Quantitative Filter Hybridization. *Nucleic Acid Hybridization,* 1985 **[0030]**
- **ALLAWI ; SANTALUCIA, JR.** *Biochemistry,* 1997, vol. 36, 10581-94 **[0030]**
- Genome Analysis: A Laboratory manual Series. 1999, vol. I-IV **[0031]**
- Generic Variation: A Laboratory Manual. 2007 **[0031]**
- PGR Primer: A Laboratory Manual. 2003 **[0031]**
- **BOWTELL ; SAMBROOK.** DNA Microarrays: A Molecular Cloning Manual. 2003 **[0031]**
- **MOUNT.** *Bioinformatics: Sequence and Genome Analysis,* 2004 **[0031]**
- **SAMBROOK ; RUSSELL.** Condensed Protocols from Molecular Cloning: A Laboratory Manual. 2006 **[0031]**
- **SAMBROOK ; RUSSELL.** Molecular Cloning: Laboratory Manual. Cold Spring Harbour Laboratory Press, 2002 **[0031]**
- **STRYER.** Biochemistry. W.H. Freeman, 1995 **[0031]**
- **GAIT.** Oligonucleotide Synthesis: A Practical Approach. IRL Press, 2002 **[0031]**
- **NELSON ; COX.** Lehninger. Principles of Biochemistry. W. H. Freeman Pub, 2000 **[0031]**
- **BERG et al.** Biochemistry. W.H. Freeman Pub, 2002 **[0031]**
- **FAN et al.** *Cold Spring Symp. Quant. Biol.,* 2003, vol. 68, 69-78 **[0053]**
- **RONAGHI et al.** *Science,* 1998, vol. 281, 363-365 **[0083]**
- **LI et al.** *Proc. Natl. Acad. Sci.,* 2003, vol. 100, 414-419 **[0083]**
- **SONI ; MELLER.** *Clin Chem,* 2007, vol. 53, 1996-2001 **[0083]**
- **MCKERNAN et al.** *Genome Res.,* 2009, vol. 19, 1527-41 **[0083]**

- **BENTLEY et al.** *Nature,* 2008, vol. 456, 53-59 **[0083]**
- **BIRNEY et al.** *Nature,* 2007, vol. 447, 799-816 **[0085]**

Szabadalmi igénypontok

1. Eljárás több nukleinsav cél térbeli gyakorisági és/vagy aktivitási mintájának meghatározására több helyen egy tesztmintában, ahol az eljárás a következő lépéseket tartalmazza:

minta rendelkezésre bocsátását, amely rögzítve van tartóhoz;

oligonukleotid próbák szállítását több nukleinsav cél számára több helyre a tesztmintában;

megengedjük, hogy a oligonukleotid minták hibridizáljanak a nukleinsav célokkal;

nem hibridizált oligonukleotid minták kimosását a tesztmintából;

kódoló szerek szállítását a több hely lokációihoz a tesztmintában az ismert térbeli minta szerint, ahol legalább kettő kódoló szer van szállítva a több hely mindegyikéhez és a kódoló szerek kombinációja, amelyek mindegyik helyhez vannak szállítva, különbözőek;

a kódoló szerek hozzárendelését az oligonukleotid próbákhoz, amelyek hibridizálva vannak nukleinsav célokhoz, hogy kódolt próbákat alkossunk;

a kódolt próbák szekvenciája egészének vagy egy részének meghatározását ; és

hozzárendeljük a biológiai célok gyakoriságát és/vagy aktivitását a több hely lokációihoz a mintában.

2. Eljárás több biológiai cél térbeli gyakorisági és/vagy aktivitási mintájának meghatározására több helyen egy tesztmintában, amely tartalmazza a következő lépéseket:

minta rendelkezésre bocsátását, amely rögzítve van tartóhoz;

kódolt minták szállítását több biológiai cél számára ismert térbeli mintában több hely számára a mintában, ahol mindegyik kódolt próba tartalmaz egy próba tartományt, amely képes kölcsönhatásra biológiai céllal és kódoló jelölést, amely tartalmaz egy oligonukleotidot, amely azonosítja a hely lokációját, amelyhez a kódolt próba szállítva van;

megengedjük a kódolt mintáknak, hogy kölcsönhassanak a biológiai célokkal;

a kódolt minták szekvenciájának meghatározását, ahol a szekvencia tartalmazza a kódoló jelölés oligonukleotid szekvenciáját, miáltal azonosítja a hely lokációját, amelyhez a kódolt próba szállítva van; és

hozzárendeljük a biológiai célok gyakoriságát és/vagy aktivitását a több hely lokációjához a mintában.

3.A 2. igénypont szerinti eljárás, amely továbbá tartalmazza a megengedő lépés után és a meghatározó lépés előtt a kódolt minták elkülönítésének lépését, amely kölcsönhat a biológiai célokkal, a kódolt mintáktól, amelyek nem kölcsönhatnak a biológiai célokkal.

4. A 3. igénypont szerinti eljárás, ahol a biológiai célok nukleinsavak és a kódolt minták oligonukleotid minták.

5. Az 1. vagy 4. igénypont szerinti eljárás, ahol van kettő oligonukleotid minta a nukleinsav biológiai célok mindegyike számára.

6. A 3. igénypont szerinti eljárás, ahol a több biológiai célok proteinek, a kódoló minták mintatartományai proteinek és a kódoló jelölések tartalmaznak oligonukleotidokat.

7. A 3. igénypont szerinti eljárás, ahol a több biológiai cél tartalmaz enzimeket.

8. A 3. igénypont szerinti eljárás, ahol a kódolt minták próba tartományai tartalmaznak antitesteket, aptamereket, kis molekulákat, enzim szubsztrátokat, vélelmezett enzim szubsztrátokat, vagy affinitásbefogó szereket.

9. A 1., 3. és 6. igénypont szerinti eljárás, amely továbbá tartalmaz amplifikációs lépést az elkülönítő lépés vagy hozzárendelő lépés és a meghatározó lépés között.

10. Az 1-3. igénypontok bármelyike szerinti eljárás, ahol a meghatározó lépés nukleinsav szekvenálással vagy high-throughput szekvenálással van megvalósítva.

11. Az 1. igénypont szerinti eljárás, ahol a hozzárendelő lépés ligálás vagy extenziót követő ligálás útján van meghatározva.

12. Az 1-3. igénypontok bármelyike szerinti eljárás, ahol a több biológiai célok terméke vizsgálva van és a több hely mintában nagyobb, mint 1.000.000.

13. Az 1-3. és 6. igénypontok bármelyike szerinti eljárás, ahol a legalább egymillió kódoló minták szekvenciái párhuzamosan vannak meghatározva.

14. Az 1-3. igénypontok bármelyike szerinti eljárás, ahol az ismert térbeli minta a minta szövettani jellemzői által van meghatározva .

15. Az 1-3. igénypontok bármelyike szerinti eljárás, ahol szoftverprogramozott hardver valósít meg a szállítási lépés, az elkülönítő lépés, meghatározó lépés és hozzárendelő lépés közül legalább kettőt.