Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2707958 A1 2009/06/18

(21) 2 707 958

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2008/11/24
(87) Date publication PCT/PCT Publication Date: 2009/06/18
(85) Entrée phase nationale/National Entry: 2010/06/03
(86) N° demande PCT/PCT Application No.: US 2008/084551
(87) N° publication PCT/PCT Publication No.: 2009/076057
(30) Priorités/Priorities: 2007/12/05 (US60/992,489); 2008/03/26 (US12/055,919); 2008/11/21 (US12/275,895)

(51) Cl.Int./Int.Cl. *C12Q 1/68* (2006.01)
(71) Demandeur/Applicant: PERKINELMER LAS, INC., US
(72) Inventeur/Inventor: ADLER, KARL EDWIN, US
(74) Agent: RIDOUT & MAYBEE LLP

(54) Titre : PROCEDES ET COMPOSITIONS CONCERNANT DES ANALYSES DE GAIN ET PERTE GENOMIQUES MULTIPLEXEES
(54) Title: METHODS AND COMPOSITIONS RELATING TO MULTIPLEX GENOMIC GAIN AND LOSS ASSAYS

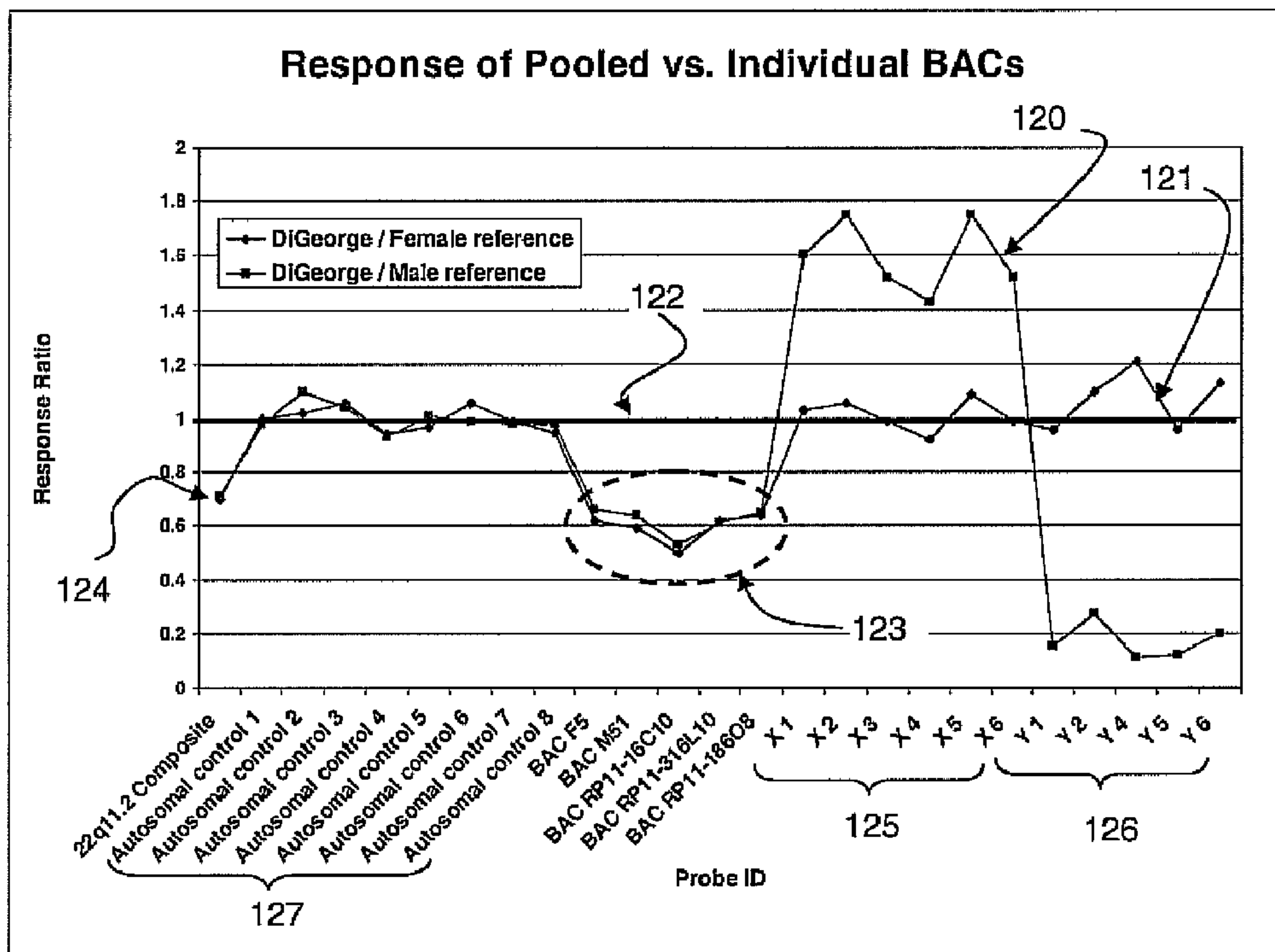

Figure 12

(57) Abrégé/Abstract:
Compositions and methods are provided for detecting genomic DNA gain and loss. Embodiments of inventive assays include using a substrate-attached composite nucleic acid probe which specifically hybridizes to two or more genomic loci in a genomic

(57) Abrégé(suite)/Abstract(continued):
region of a reference genome. The genomic region is characterized by a first terminus and a second terminus and has an intermediate region disposed between the first terminus and second terminus of at least 400 kilobases. The composite nucleic acid probe includes nucleic acid sequences which specifically hybridize to substantially an entire first genomic locus including the first terminus and to substantially an entire second genomic locus including the second terminus. Methods and compositions are provided which include assessment of two or more genomic DNA references.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 June 2009 (18.06.2009)

PCT

(10) International Publication Number
WO 2009/076057 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2008/084551

(22) International Filing Date:
24 November 2008 (24.11.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/992,489	5 December 2007 (05.12.2007)	US
12/055,919	26 March 2008 (26.03.2008)	US
12/275,895	21 November 2008 (21.11.2008)	US

(71) Applicant *(for all designated States except US)*: **PERKINELMER LAS, INC.** [US/US]; 940 Winter Street, Waltham, MA 02451 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **ADLER, Karl, Edwin** [US/US]; 2 Shandel Drive, Newburyport, MA 01950-3250 (US).

(74) Agents: STAPLE, Julie, K. et al.; Gifford, Krass, Sprinkle, Anderson & Citkowski, P.C., 2701 Troy Center Drive, Suite 330, Post Office Box 7021, Troy, MI 48007-7021 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

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

[Continued on next page]

(54) Title: METHODS AND COMPOSITIONS RELATING TO MULTIPLEXED GENOMIC GAIN AND LOSS ASSAYS

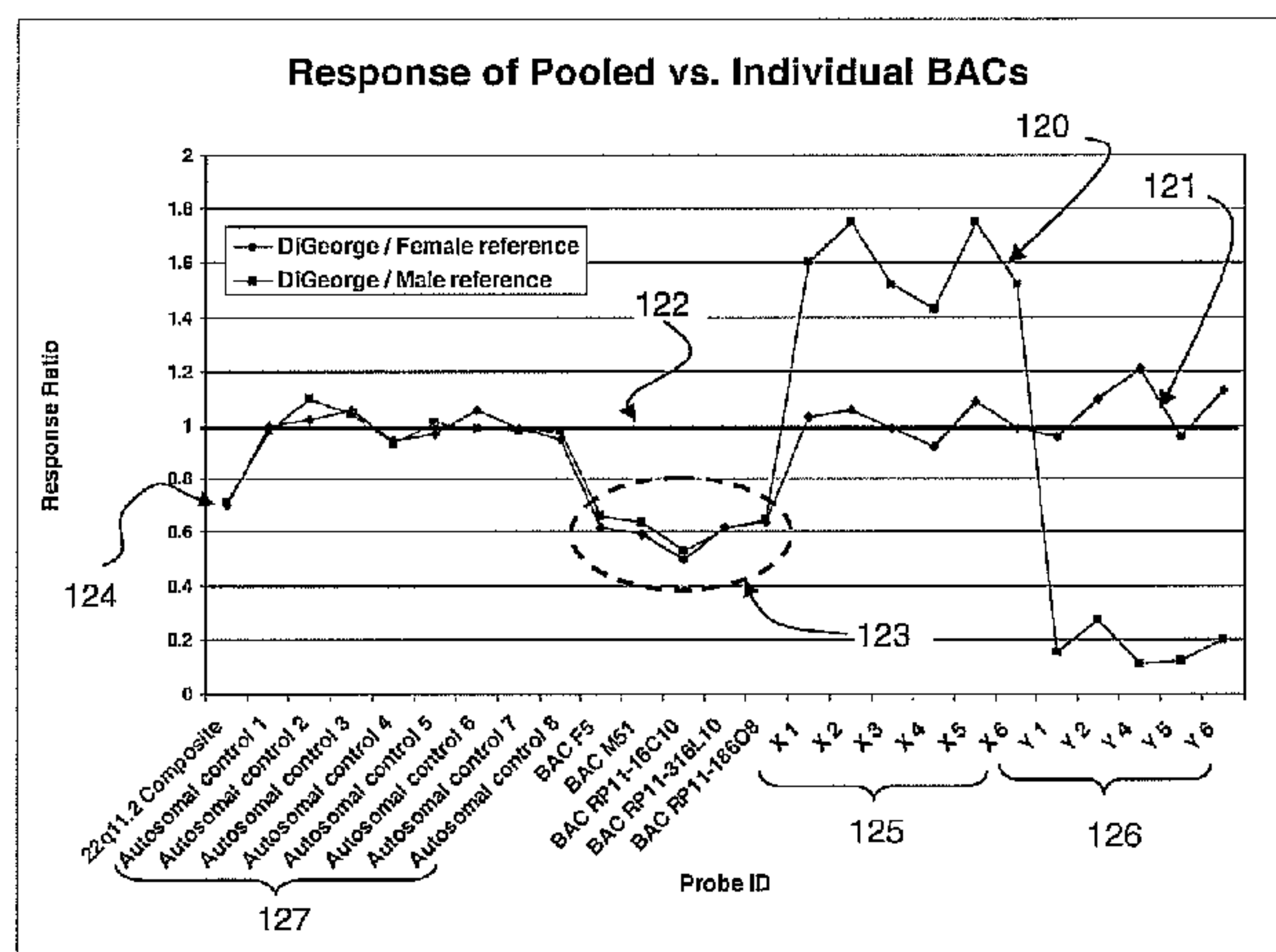

Figure 12

(57) Abstract: Compositions and methods are provided for detecting genomic DNA gain and loss. Embodiments of inventive assays include using a substrate-attached composite nucleic acid probe which specifically hybridizes to two or more genomic loci in a genomic region of a reference genome. The genomic region is characterized by a first terminus and a second terminus and has an intermediate region disposed between the first terminus and second terminus of at least 400 kilobases. The composite nucleic acid probe includes nucleic acid sequences which specifically hybridize to substantially an entire first genomic locus including the first terminus and to substantially an entire second genomic locus including the second terminus. Methods and compositions are provided which include assessment of two or more genomic DNA references.

WO 2009/076057 A2

— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

METHODS AND COMPOSITIONS RELATING TO MULTIPLEX GENOMIC GAIN AND LOSS ASSAYS

5 REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority from U.S. Patent Application Serial No. 12/275,895, filed November 21, 2008, which is a continuation-in-part of U.S. Patent Application Serial No. 12/055,919, filed March 26, 2008, which claims priority from U.S. Provisional Patent Application Serial No. 60/992,489, filed December 5, 2007. The entire content of each 10 application is incorporated herein by reference.

FIELD OF THE INVENTION

[0002] Technology described herein relates generally to methods and compositions for detection of nucleic acids. More specifically described are methods and compositions for 15 genomic gain and loss assays.

BACKGROUND OF THE INVENTION

[0003] Assays for detection of genomic gain and loss allow for detection and diagnosis of genetic abnormalities which can underlie disease, behavioral and cognitive conditions, and other 20 genetic-based pathologies.

[0004] Array-CGH (aCGH) is a multiplex assay method where immobilized DNA probes capture labeled complementary genomic DNA sequences from a test sample and a reference sample. Each probe generates a ratio of the test/reference amounts of one of the constituent sequences defined by the probe sequence. In aCGH it is common to use bacterial artificial 25 chromosome (BAC) DNA as the probe material immobilized onto the array's solid support. Typical BAC sequence length is between about 100 kilobases and about 250 kilobases, with 175 kilobases length typical. Probes of this length are well-suited to aCGH in that they are long enough to generate large hybridization signals and to not respond to small mutations such as SNPs, yet short enough to provide sufficient genomic resolution to detect the break points of a 30 genomic gain or loss with some precision. For example, a well-designed BAC array can detect different sizes of deletion regions in microdeletion syndromes such as 1p36 deletion syndrome or DiGeorge syndrome, where deletions spanning different portions of the region associated with the disorder may correspond to different phenotypes.

[0005] The trend in the design of so-called constitutional aCGH microarrays has been 35 toward using multiple separate probes, such as BAC probes in each of the genomic regions

associated with inborn DNA disorders to detect the extent of genomic gains and losses with higher resolution. The higher resolution reveals more information on the exact size and boundaries of a gain or loss region. However, there are cases in which high resolution detection of the exact size and boundaries of a gain or loss region may not be necessary.

5 **[0006]** Aneuploidies are examples of genomic gain-loss anomalies where detecting breakpoints at high resolution is not relevant. Aneuploidies are a gain (most typically, e.g. trisomy) or a loss of an entire chromosome (average size over 100 megabases), as opposed to a microdeletion disorder where the deleted region may be only from typically 1 to 10 or so megabases in extent. Detection of aneuploidy (also known as chromosome enumeration) can be 10 done with as few as one BAC probe in an aCGH assay, however noise on the sample/normal ratio from that single probe can lead to ambiguity about the result. The use of multiple probes targeted to the subject chromosome where the majority of the probes all show the same gain or loss provides more confidence in an aneuploidy detection result from an aCGH assay, and that is the routine construction of CGH arrays. Most array analyses require concordant ratio signals 15 from at least two or three or more BAC probes in a region before a gain or loss is definitively "called" or determined.

[0007] Additional individual probes in the gain or loss region increases the confidence of the detection of the gain or loss which is valuable in cases where the assay signal ratios are noisy. Noise is present in assays for a variety of reasons. Sometimes the assay is simply noisy 20 due to non-uniform conditions of either probe immobilization or sample incubation or due to non-optimum conditions in labeling, hybridization, washing, or drying. In other cases, noise may be induced by the sample having been amplified using whole-genome amplification (WGA) such as the phi-29 or DOP PCR methods, where the amplified sample shows sequence-specific amplification bias. WGA is used when the initial sample amount is limited, such as the DNA 25 from a single cell or from just a few cells. Such amplified samples have varying DNA product yields in different parts of the genome (amplification bias), superimposing genomic gain-loss noise onto the assay that generally results in significant variation of sample-to-reference ratio response between multiple individually arrayed or assayed BAC probes. The standard method for compensating for this noise is to use a plurality of probes to span the genomic sequence 30 region of interest and in some manner, such as averaging the ratio responses across the genomic region, utilizing the composite results of the plurality of probes to make a gain-loss call across the region. Averaging of ratios across multiple individually arrayed or assayed probes is particularly common in oligonucleotide CGH arrays, which typically show more probe-to-probe ratio variation than BAC arrays.

[0008] Unfortunately, use of multiple individually arrayed or assayed probes directed to the same region in order to increase confidence can be limiting. For example, there are configurations of immobilized probe arrays where the use of multiple individually arrayed or assayed probes to detect each aneuploidy is problematic. Some array CGH formats are restricted 5 in the number of probes that they can accommodate. An example of such an array is one printed on the bottom of a microplate well, or one printed on just a small segment of a microscope slide substrate that is configured to accommodate multiple separate samples on a single substrate. These arrays commonly utilize between 9 (3x3 spot array matrix) and 100 (10x10) probes. In such a 100-plex case a maximum of 4 probes could be accommodated for enumerating each of 10 the 24 chromosomes (1 – 22, X and Y), and effectively the entire capacity of the array would have been used up for chromosome enumeration.

[0009] Besides planar microarrays another method for measuring genomic gains and losses in a sample using multiple probes simultaneously is the use of a set of encoded particles or microspheres with immobilized probes. The most widely used encoded particle platform is the 15 Luminex xMAP system that uses fluorescent color-coding to distinguish 100 different microsphere or bead types. This system can support a maximum of 100 different probes in a genomic gain-loss assay. This encoded microsphere platform does not support the two-dye two-color readout of ratio results from competitive hybridization as microarrays do, but the same ratios can be generated by separate side-by-side assays of test and reference samples, 20 normalizing appropriately, and calculating the ratios.

[0010] In these multiplex assay formats where the maximum numbers of probes is limited there is a continuing need to reliably assay for aneuploidies of up to 24 chromosomes (1 through 22, plus X and Y) without utilizing the entire or even the majority of the possible set of probes just for that purpose. For example, in constitutional disorder assay panels it is most often desired 25 to probe for several microdeletion disorders in addition to aneuploidies. Also, assaying for microdeletion disorders (such as Wolf-Hirschhorn, Willams-Beuren, Cri-du-Chat and the like) can be done robustly using one or two lower resolution probes rather than a larger number of conventional probes, allowing more disorders to be assayed on a limited-probe platform Similarly in cancer cytogenetics there are generally several regions where fairly high-resolution 30 gain-loss data is required in addition to other areas where the loss of an entire chromosome or chromosome arm is the desired resolution. Such panels would be more useful if the genomic regions requiring only low resolution did not utilize more than a few probes in the panel, freeing up assay platform capacity for more probes for the regions needing higher resolution.

[0011] Therefore, methods and compositions are required for reliably assaying for chromosome-scale and chromosome-arm scale gains and losses while allowing for the option of assaying other genomic regions simultaneously with one or more higher resolution probes.

5

SUMMARY OF THE INVENTION

[0012] A method of assaying a DNA sample is described which includes providing a substrate-attached composite nucleic acid probe. The composite nucleic acid probe includes nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome. The genomic region is characterized by a first terminus and a
10 second terminus with an intermediate region of at least 400 kilobases disposed between the termini. The composite nucleic acid probe includes nucleic acid sequences which specifically hybridize to substantially an entire first genomic locus which includes the first terminus of the genomic region and nucleic acid sequences which specifically hybridize to substantially an entire second genomic locus which includes the second terminus. The first genomic locus and second
15 genomic locus each typically include at least about 100 kilobases. Optionally, nucleic acid sequences of the composite probe specifically hybridize to additional loci within the genomic region.

[0013] The substrate-attached composite nucleic acid probe is hybridized with sample genomic DNA at a stringency sufficient to achieve specific hybridization. The substrate-
20 attached composite nucleic acid probe is also hybridized with reference genomic DNA.

[0014] A first signal indicating specific hybridization of the substrate-attached composite nucleic acid probe with the sample genomic DNA is detected along with a second signal indicating specific hybridization of the substrate-attached composite nucleic acid probe with the reference genomic DNA. The first signal and the second signal are compared to detect
25 differences between the first and second signals, indicative of differences between the sample DNA and the reference DNA.

[0015] The nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome can be derived from two or more large-insert DNA vectors. For instance, the nucleic acid sequences are derived from two or more large-insert DNA
30 vectors such as bacterial artificial chromosomes, yeast artificial chromosomes, human artificial chromosomes, cosmids, plasmids, phagemids, phage DNA and fosmids. The nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome can also be derived from isolated chromosomes and/or isolated chromosome fragments.

[0016] In a particular option, the nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome are amplicons derived from two or more large-insert DNA vectors, isolated chromosomes, isolated chromosome fragments or a combination of these or other sources of nucleic acids.

[0017] In a particular option, the substrate is a plurality of particles, such as a plurality of encoded particles. In a further option, the substrate is a planar substrate.

[0018] The nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome individually have a length in the range of about 20 – 250,000 nucleotides, inclusive.

[0019] A method of assaying sample nucleic acid is provided which includes providing a multiplex reagent including a mixture of two or more encoded particle sets encoded such that each particle of each encoded particle set is detectably distinguishable from each particle of each other encoded particle set. The encoded particles include attached nucleic acid sequences which specifically hybridize to at least one genomic locus of a reference genome, and at least one encoded particle set includes an attached composite nucleic acid probe.

[0020] The multiplex reagent is hybridized with sample genomic nucleic acid and with reference nucleic acid, together or in parallel.

[0021] A first signal is detected which indicates specific hybridization of the attached nucleic acid sequences with detectably labeled sample nucleic acid and a second signal is detected indicating specific hybridization of the attached nucleic acid sequences with detectably labeled reference nucleic acid. The encoded particles are then identified so as to associate particle encoding with the first signal or with the second signal. The first signal and the second signal for each encoded particle set are then compared and differences in the first and second signals are indicative of differences between the sample and reference nucleic acids.

[0022] A reagent for assay of nucleic acids is provided which includes a first composite nucleic acid probe attached to a solid substrate. The first composite nucleic acid probe includes nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome. The genomic region is characterized by a first terminus and a second terminus and has an intermediate region disposed between the first terminus and second terminus of at least 400 kilobases. The first composite nucleic acid probe includes nucleic acid sequences which specifically hybridize to substantially an entire first genomic locus including the first terminus of the genomic region and nucleic acid sequences which specifically hybridize to substantially an entire second genomic locus including the second terminus of the genomic region.

[0023] In a particular option, a reagent for assay of nucleic acids includes at least two, and may include more, composite probes. For example, a second composite nucleic acid probe attached to a solid substrate includes nucleic acid sequences which specifically hybridize to two or more genomic loci in a second genomic region of a reference genome. The second genomic region is characterized by a first terminus and a second terminus and has an intermediate region disposed between the first terminus and second terminus of at least 400 kilobases. The second composite nucleic acid probe includes nucleic acid sequences which specifically hybridize to substantially an entire first genomic locus comprising the first terminus of the second genomic region and to substantially an entire second genomic locus comprising the second terminus of the second genomic region.

[0024] The solid substrate is a planar substrate in one option. An array including one or more composite probes can be included on a planar substrate. In a further option, non-composite probes can also be included on the planar substrate to provide a panel of composite and other probes.

[0025] In a further aspect, the solid substrate is a first plurality of particles. A second composite probe or non-composite probe can be attached to a second plurality of particles.

[0026] The first plurality of particles and second plurality of particles are distinguishably encoded for use in certain assay types. The first plurality and second plurality of distinguishably encoded particles are optionally mixed to provide a multiplex assay reagent. Additional particles having additional attached probes can be used in nucleic acid assay described herein in multiplex or separate assay formats.

[0027] A method of preparing a substrate-attached composite nucleic acid probe reagent for assay of DNA is provided herein which includes isolating a first nucleic acid sequence which specifically hybridizes to substantially an entire first genomic locus which includes a first terminus of a genomic region of a reference genome. The method includes isolating a second nucleic acid sequence which specifically hybridizes to substantially an entire second genomic locus comprising a second terminus of the genomic region of the reference genome. The first and the second nucleic acid sequences are mixed to produce a composite probe and the composite probe is then bound to a solid substrate, producing a substrate-attached composite nucleic acid probe reagent for assay of nucleic acids, such as genomic DNA.

[0028] In a particular option, the first nucleic acid sequence is isolated from a first large-insert vector and the second nucleic acid sequence is isolated from a second large-insert vector. For example, the first nucleic acid sequence is isolated from a first BAC and the second nucleic acid sequence is isolated from a second BAC.

[0029] Optionally, the first and the second nucleic acid sequences are amplified prior to or after mixing.

[0030] Methods are provided according to particular embodiments which utilize two or more reference samples.

5 **[0031]** Methods of assaying a DNA sample are provided according to particular embodiments which include providing a substrate-attached nucleic acid probe; hybridizing the substrate-attached nucleic acid probe with sample genomic DNA obtained from a subject;

[0032] hybridizing the substrate-attached nucleic acid probe with first reference genomic DNA; and hybridizing the substrate-attached nucleic acid probe with second reference genomic

10 DNA. A first signal is then detected indicating specific hybridization of the substrate-attached nucleic acid probe with the sample genomic DNA, a second signal indicating specific hybridization of the substrate-attached nucleic acid probe with the first reference genomic DNA and a third signal indicating specific hybridization of the substrate-attached nucleic acid probe with the second reference genomic DNA. The first signal and the second signal are compared to

15 detect differences between the first and second signals, and the differences of the first and second signals are indicative of differences between the sample DNA and the first reference DNA. The first signal and the third signal are also compared to detect differences between the first and third signals, and the differences between the first and third signals are indicative of differences between the sample DNA and the second reference DNA.

20 **[0033]** Optionally, differences between the sample DNA and the first reference DNA are compared with differences between the sample DNA and the second reference DNA, according to embodiments of methods of assaying the DNA sample.

[0034] In a further option, the first reference genomic DNA includes male-specific genomic DNA, and the second reference genomic DNA includes female-specific genomic DNA such that

25 comparison of the differences of the first and second signals and comparison of the differences of the first and third signals is indicative of gender of the subject.

[0035] In another option, the first reference genomic DNA includes first condition-specific genomic DNA, and the second reference genomic DNA comprises second condition-specific genomic DNA. Thus, comparison of the differences of the first and second signals and

30 comparison of the differences of the first and third signals is indicative of a disease state of the subject.

[0036] Methods are provided according to some embodiments where the first reference genomic DNA includes first condition-specific genomic DNA, and the second reference genomic DNA includes second condition-specific genomic DNA, and comparison of the

differences between the first and second signals and comparison of the differences between the first and third signals is indicative of metabolic age of the subject.

[0037] The substrate-attached nucleic acid probe can include a plurality of encoded particles and/or a planar substrate. Optionally, the substrate-attached nucleic acid probe includes 5 a plurality of oligonucleotides and/or a plurality of amplicons. The substrate-attached nucleic acid probe can include insert DNA isolated from a large-insert DNA vector and/or isolated chromosomal DNA.

[0038] Methods of assaying a DNA sample are provided according to embodiments of the present invention which include providing a first encoded particle set inlcuding encoded particles 10 having attached amplicons. The amplicons include random nucleic acid sequences which together represent substantially an entire first template DNA sequence. The amplicons of the first encoded particle set are hybridized with detectably labeled sample DNA, detectably labeled first reference DNA, and with detectably labeled second reference DNA. A first signal is detected which is indicative of specific hybridization of the amplicons of the first encoded 15 particle set with detectably labeled sample DNA. A second signal is detected which is indicative of specific hybridization of the amplicons of the first encoded particle set with detectably labeled first reference DNA. A third signal is detected which is indicative of specific hybridization of the amplicons of the first encoded particle set with detectably labeled second reference DNA. The first signal and the second signal are compared to detect differences between the first and second 20 signals, and the differences of the first and second signals are indicative of differences between the sample DNA and the reference DNA. The first signal and the third signal are compared to detect differences between the first and third signals, and the differences of the first and third signals are indicative of differences between the sample DNA and the second reference DNA.

[0039] In one option, the amplicons have a length in the range of about 500 – 1200 25 nucleotides, inclusive.

[0040] Methods according to particular embodiments further include providing a second encoded particle set including encoded particles having attached amplicons, where the amplicons include random nucleic acid sequences which together represent substantially an entire second template DNA sequence. The amplicons of the second encoded particle set are hybridized with 30 detectably labeled sample DNA, detectably labeled first reference DNA and detectably labeled second reference DNA. A first signal indicating specific hybridization of the amplicons of the second encoded particle set with detectably labeled sample DNA is detected. A second signal indicating specific hybridization of the amplicons of the second encoded particle set with detectably labeled reference DNA is detected. A third signal indicating specific hybridization of

the amplicons of the second encoded particle set with detectably labeled second reference DNA is detected. The first signal indicating specific hybridization of the amplicons of the second encoded particle set and the second signal indicating specific hybridization of the amplicons of the second encoded particle set are compared to detect differences between the first and second
5 signals, the differences of the first and second signals indicative of differences between the sample DNA and the first reference DNA. The first signal indicating specific hybridization of the amplicons of the second encoded particle set and the third signal indicating specific hybridization of the amplicons of the second encoded particle set are compared to detect differences between the first and third signals, the differences of the first and third signals
10 indicative of differences between the sample DNA and the second reference DNA.

[0041] Optionally, the first and second encoded particle sets are provided in a mixture. Particle encoding is detected to associate each particle set identity with specific signals obtained from hybridization with sample and reference DNA.

15 BRIEF DESCRIPTION OF THE DRAWINGS

[0042] Figure 1 is a flowchart illustrating an embodiment including preparing amplicons from template DNA using two amplification reactions and immobilizing the amplicons as probes onto a set of encoded beads, where the beads in the set all have the same ID code;

[0043] Figure 1A is a flowchart illustrating an embodiment including preparing BAC
20 amplicons from a single BAC clone and immobilizing the amplicons as probes onto a set of encoded beads, generating a bead set where the beads in the set all have the same ID code;

[0044] Figure 2 is a flowchart illustrating an embodiment including mixing m different encoded bead sets, each with its respective immobilized BAC-amplicon probe DNA, together to make a multiplexed encoded bead set;

25 **[0045]** Figure 3 is a flowchart illustrating an embodiment including running a multiplexed genomic gain and loss assay on n samples using a multiplexed encoded bead set;

[0046] Figure 3A is a flowchart illustrating an embodiment including running a multiplexed genomic gain and loss assay on n samples using a multiplexed encoded bead set;

[0047] Figure 4 is a schematic diagram of a 96-well SBS-standard microplate, showing
30 example locations of duplicate references and duplicate samples for running an assay on 46 samples in parallel;

[0048] Figure 5 is an example of data generated using a Coriell DNA sample having a trisomy on chromosome 13, sex male;

[0049] Figure 6 is an example of data generated using a Coriell DNA sample having a trisomy on chromosome 18, sex male;

[0050] Figure 7 is an example of data generated using a Coriell DNA sample having a trisomy on chromosome 21, sex female;

5 **[0051]** Figure 8 is an example of data generated using a Coriell DNA sample having a 5-copy amplification of the X chromosome;

[0052] Figure 9 is a table displaying the BAC clones used to generate amplicons immobilized onto encoded beads in the example assays, their chromosome and cytoband locations, the sequence of the negative control oligonucleotide, and the bead ID (Luminex bead

10 region) for the bead set to which each amplicon probe is immobilized;

[0053] Figure 10A is a schematic flowchart showing a process for making a composite probe according to an aspect of a process described herein;

[0054] Figure 10B is a schematic flowchart showing a process for making a composite probe according to an aspect of a process described herein;

15 **[0055]** Figure 11 is a schematic flowchart showing a process for making a composite probe according to an aspect of a process described herein;

[0056] Figure 12 is a graph of plotted data from a test assay demonstrating the use of a composite probe with a DiGeorge syndrome reference DNA sample;

[0057] Figure 13 is a graph showing ratio data from a Luminex bead array gain – loss assay

20 using two reference genomic DNA samples; and

[0058] Figure 14 is a graph showing ratio data from a Luminex bead array gain – loss assay using two reference genomic DNA samples.

25

DETAILED DESCRIPTION OF THE INVENTION

[0059] Methods and compositions relating to assays for chromosomal gains and losses are provided herein. Broadly, methods and compositions are described herein which relate to assays of genomic DNA gain and loss using substrate-attached nucleic acid probes.

[0060] Scientific and technical terms used herein are intended to have the meanings

30 commonly understood by those of ordinary skill in the art. Such terms are found defined and used in context in various standard references illustratively including J. Sambrook and D.W. Russell, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press; 3rd Ed., 2001; F.M. Ausubel, Ed., Short Protocols in Molecular Biology, Current Protocols; 5th Ed., 2002; B. Alberts et al., Molecular Biology of the Cell, 4th Ed., Garland, 2002; D.L. Nelson and

M.M. Cox, Lehninger Principles of Biochemistry, 4th Ed., W.H. Freeman & Company, 2004; and Herdewijn, P. (Ed.), Oligonucleotide Synthesis: Methods and Applications, Methods in Molecular Biology, Humana Press, 2004.

[0061] The term "nucleic acid" as used herein refers to RNA or DNA molecules having more than one nucleotide in any form including single-stranded, double-stranded, oligonucleotide or polynucleotide.

[0062] Probes

[0063] The term "probe" is used herein to refer to a nucleic acid used to identify a target nucleic acid to which the probe specifically binds.

[0064] A nucleic acid probe for use in an assay described herein can encompass all or part of a genome of a cell or organism. The nucleic acid probe can encompass DNA representing one or more chromosomes, a portion of a chromosome, a genetic locus, a gene or a portion of a gene. The nucleic acid probe can be in any form, such as an insert in a vector illustratively including a bacterial artificial chromosome, yeast artificial chromosome, human artificial chromosome, cosmid, plasmid, phagemid, phage DNA or fosmid. The nucleic acid probe can be in the form of microdissected chromosomal DNA. Thus, while specific examples described herein refer to BACs as sources of nucleic acid probe DNA, other types of clones such as PACs, YACs, cosmids, fosmids, cDNAs and the like may be used.

[0065] In particular applications, nucleic acids are used as template material for amplification and the resulting amplicons, or portions thereof, are used as probes.

[0066] Nucleic acids for use in generating a nucleic acid probe, sample, or reference are obtained by methods known in the art, for instance, as described in J. Sambrook and D.W. Russell, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press; 3rd Ed., 2001 or F.M. Ausubel, Ed., Short Protocols in Molecular Biology, Current Protocols; 5th Ed., 2002. Nucleic acids may also be obtained commercially and/or using commercial kits.

[0067] Composite Probes

[0068] A composite nucleic acid probe is provided according to particular embodiments which includes nucleic acid sequences that specifically hybridize to two or more genomic loci in a genomic region of interest in a reference genome. The genomic region is characterized by a first terminus and a second terminus with an intermediate region of at least 400 kilobases disposed between the termini. The composite nucleic acid probe includes nucleic acid sequences which specifically hybridize to substantially an entire first genomic locus which includes the first terminus of the genomic region and nucleic acid sequences which specifically hybridize to substantially an entire second genomic locus which includes the second terminus. The first

genomic locus and second genomic locus each typically include at least about 100 kilobases, and may be larger.

[0069] The nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome can be derived from two or more large-insert DNA vectors. For instance, the nucleic acid sequences are derived from two or more large-insert DNA vectors such as bacterial artificial chromosomes, yeast artificial chromosomes, human artificial chromosomes, P1 derived artificial chromosomes (PAC), cosmids, plasmids, phagemids, phage DNA and fosmids. The nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome can also be derived from isolated chromosomes and/or isolated chromosome fragments.

[0070] In a particular option, the nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome are amplicons amplified from templates derived from two or more large-insert DNA vectors, isolated chromosomes, isolated chromosome fragments or a combination of these or other sources of nucleic acids.

[0071] In particular embodiments, the nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome are provided as oligonucleotides and/or polynucleotides which individually have a length in the range of about 20 – 250,000 nucleotides, inclusive and which, together, specifically hybridize to substantially an entire first genomic locus which includes the first terminus of the genomic region and to substantially an entire second genomic locus which includes the second terminus of the genomic region.

[0072] Thus, a composite probe can include a pool of 2 or more BACs, or the insert DNA derived from 2 or more BACs. Optionally, the pool can include the insert DNA derived from between 4 and about 100 BACs, inclusive. The probe DNA may be extracted from the cultured BACs or in a particular embodiment may be amplicons derived from the BAC DNA, for example by degenerate oligonucleotide primer (DOP) PCR or ligation-mediated PCR. Other large-insert clones, such as PACs, YACs, cosmids, fosmids, etc. can be used instead of BACs. Pools of large numbers of oligonucleotides can also be used.

[0073] In a particular aspect, the composite immobilized probes may be DNA derived from sorted chromosomes. Cell-sorting flow cytometers (also known as fluorescence activated cell sorters; FACS) can be used to create pools of sorted chromosomes where the sorting is based on the ratio of signals between stains specific for AT- and GC-rich regions of the genome. DNA extracted from such sorted chromosomes can be amplified by WGA and labeled with a fluorescent dye to be used as chromosome-painting probes in metaphase FISH analysis.

Similarly, chromosome arm painting probes can be prepared by amplifying the DNA from sorted chromosome arms separated from metaphase spreads by laser-capture microdissection. Immobilized whole-chromosome or chromosome-arm probes immobilized as microarray spots or onto encoded particles produce hybridization ratio signals representing the average gain or loss across the whole chromosome or arm, respectively, in a manner similar to a pool of a very large number of BACs that span an entire chromosome or arm.

[0074] Composite probes have utility in a variety of roles in a multiplex genomic gain-loss assay, as exemplified in a multiplex genomic gain-loss assay described herein. These include chromosome enumeration (detection of aneuploidy), detection of microdeletion or other syndromes, or as controls. Using composite probes as controls to determine the normal autosomal response (the nominal sample/reference ratio = 1.0 level) allows averaging of response over a larger span of the genome to reduce ratio noise caused by normal copy number variations between individuals.

[0075] Process for Making Composite Probes

[0076] A method of preparing a substrate-attached composite nucleic acid probe reagent for assay of DNA is provided herein. A first nucleic acid sequence is isolated which specifically hybridizes to substantially an entire first genomic locus which includes a first terminus of a genomic region of a reference genome. Further, at least a second nucleic acid sequence is isolated which specifically hybridizes to substantially an entire second genomic locus comprising a second terminus of the genomic region of the reference genome. The term "isolated" when used in reference to nucleic acids refers to nucleic acids substantially separated from other substances with which they are naturally found, such as cells, proteins and other nucleic acids. The first and the second nucleic acid sequences are mixed to produce a composite probe and the composite probe is then bound to a solid substrate, producing a substrate-attached composite nucleic acid probe reagent for assay of nucleic acids, such as genomic DNA.

[0077] In a particular option, the first nucleic acid sequence is isolated from a first large-insert vector and the second nucleic acid sequence is isolated from a second large-insert vector. For example, the first nucleic acid sequence is a DNA insert isolated from a first BAC and the second nucleic acid sequence is a DNA insert isolated from a second BAC.

[0078] In a further option, the first nucleic acid sequence is a human genomic DNA insert isolated from a first BAC and the second nucleic acid sequence is a human genomic DNA insert isolated from a second BAC.

[0079] Optionally, the nucleic acid sequences are amplified prior to or after mixing. For example, the first and the second, or more, nucleic acid sequences are amplified prior to or after mixing.

[0080] The number of genomic loci targeted by nucleic acid sequences in a composite probe is not limited to two and may be three, four or more, such as about 4-100, or more. Thus, the number of particular nucleic acid sequences which specifically hybridize to genomic loci is likewise not limited to two and may be three, four or more, such as about 4-100, or more.

[0081] An example of a composite probe is a composite probe made using insert DNA from five BACs mapping to cytoband 22p11.2 corresponding to the DiGeorge microdeletion syndrome. The center loci of the five BACs used span about 0.45 megabases (445 kilobases), the total span accounting for the175 kb typical length of the BACs is a little over 600 kilobases. The isolated insert DNA from each of the five BACs is amplified and the resulting amplicons are pooled to produce a composite probe. The composite probe is then bound to a set of Luminex encoded multiplex microspheres, all having the same bead encoding identification, for use as a 22p11.2 cytoband probe in a multiplex genomic gain-loss assay.

[0082] Thus, a particular example, 2, 3, 4, 5 or more, such as 6-100, or more, inserts from BAC, or other large insert vectors, can be used as nucleic acid sequences which specifically hybridize to particular genomic loci in a defined genomic region.

[0083] Substrates

[0084] A solid substrate, which includes semi-solid substrate, for attachment of a probe, including a composite probe, can be any of various materials such as glass; plastic, such as polypropylene, polystyrene, nylon; paper; silicon; nitrocellulose; or any other material to which a nucleic acid can be attached for use in an assay. The substrate can be in any of various forms or shapes, including planar, such as silicon chips and glass plates; and three-dimensional, such as particles, microtiter plates, microtiter wells, pins, fibers and the like.

[0085] In particular aspects, a solid substrate to which a probe is attached is a particle.

[0086] Particles to which a probe is bound can be any solid or semi-solid particles to which a probe can be attached, which are suitable for a nucleic acid hybridization assay and which are stable and insoluble under hybridization and detection conditions. The particles can be of any shape, such as cylindrical, spherical, and so forth, size, composition, or physiochemical characteristics. The particle size or composition can be chosen so that the particle can be separated from fluid, e.g., on a filter with a particular pore size or by some other physical property, e.g., a magnetic property.

[0087] Microparticles, such as microbeads, used can have a diameter of less than one millimeter, for example, a size ranging from about 0.1 to about 1,000 micrometers in diameter, inclusive, such as about 3-25 microns in diameter, inclusive, or about 5-10 microns in diameter, inclusive. Nanoparticles, such as nanobeads used can have a diameter from about 1 nanometer (nm) to about 100,000 nm in diameter, inclusive, for example, a size ranging from about 10-1,000 nm, inclusive, or for example, a size ranging from 200-500 nm, inclusive. In certain embodiments, particles used are beads, particularly microbeads and nanobeads.

[0088] Particles are illustratively organic or inorganic particles, such as glass or metal and can be particles of a synthetic or naturally occurring polymer, such as polystyrene, polycarbonate, silicon, nylon, cellulose, agarose, dextran, and polyacrylamide. Particles are latex beads in particular embodiments.

[0089] Particles used include functional groups for binding to nucleic acids in particular embodiments. For example, particles can include carboxyl, amine, amino, carboxylate, halide, ester, alcohol, carbamide, aldehyde, chloromethyl, sulfur oxide, nitrogen oxide, epoxy and/or tosyl functional groups. Functional groups of particles, modification thereof and binding of a chemical moiety, such as a nucleic acid, thereto are known in the art, for example as described in Fitch, R. M., Polymer Colloids: A Comprehensive Introduction, Academic Press, 1997. U.S. Pat. No. 6,048,695 describes an exemplary method for attaching nucleic acid probes, such as amplicons, to a substrate, such as particles. In a further particular example, 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride, EDC or EDAC chemistry, can be used to attach nucleic acids to encoded particles.

[0090] Encoded particles are particles which are distinguishable from other particles based on a characteristic illustratively including an optical property such as color, reflective index and/or an imprinted or otherwise optically detectable pattern. For example, the particles may be encoded using optical, chemical, physical, or electronic tags. Encoded particles can contain or be attached to, one or more fluorophores which are distinguishable, for instance, by excitation and/or emission wavelength, emission intensity, excited state lifetime or a combination of these or other optical characteristics. Optical bar codes can be used to encode particles.

[0091] In particular embodiments, each particle of a particle set is encoded with the same code such that each particle of a particle set is distinguishable from each particle of another particle set. In further embodiments, two or more codes can be used for a single particle set. Each particle can include a unique code, for example. In certain embodiments, particle encoding includes a code other than or in addition to, association of a particle and a nucleic acid probe specific for genomic DNA.

[0092] In particular embodiments, the code is embedded, for example, within the interior of the particle, or otherwise attached to the particle in a manner that is stable through hybridization and analysis. The code can be provided by any detectable means, such as by holographic encoding, by a fluorescence property, color, shape, size, light emission, quantum dot emission and the like to identify particle and thus the capture probes immobilized thereto. In some embodiments, the code is other than one provided by a nucleic acid.

[0093] One exemplary platform utilizes mixtures of fluorescent dyes impregnated into polymer particles as the means to identify each member of a particle set to which a specific capture probe has been immobilized. Another exemplary platform uses holographic barcodes to identify cylindrical glass particles. For example, Chandler et al. (U.S. Pat. No. 5,981,180) describes a particle-based system in which different particle types are encoded by mixtures of various proportions of two or more fluorescent dyes impregnated into polymer particles. Soini (U.S. Pat. No. 5,028,545) describes a particle-based multiplexed assay system that employs time-resolved fluorescence for particle identification. Fulwyler (U.S. Pat. No. 4,499,052) describes an exemplary method for using particle distinguished by color and/or size. U.S. Patent Application Publications 20040179267, 20040132205, 20040130786, 20040130761, 20040126875, 20040125424, and 20040075907 describe exemplary particles encoded by holographic barcodes. U.S. Pat. No. 6,916,661 describes polymeric microparticles that are associated with nanoparticles that have dyes that provide a code for the particles

[0094] While an embodiment described in detail herein utilizes the Luminex encoded bead platform, other types of encoded particle assay platforms may be used, such as the VeraCode beads and BeadXpress system (Illumina Inc., San Diego CA), xMAP 3D (Luminex) and the like. Magnetic Luminex beads can be used which allow wash steps to be performed with plate magnets and pipetting rather than with filter plates and a vacuum manifold. Each of these platforms are typically provided as carboxyl beads but may also be configured to include a different coupling chemistry, such as amino-silane.

[0095] Binding to Substrate

[0096] Binding of the nucleic acid probes to a substrate is achieved by any of various methods effective to bond a nucleic acid to a solid or semi-solid substrate, illustratively including adsorption and chemical bonding. The nucleic acids can be bonded directly to the material of the encoded particles or indirectly bonded to the encoded particles, for example, via bonding to a coating or linker disposed on the particles. Nucleic acids can be synthesized, and/or modified once synthesized, to include a functional group for use in bonding the nucleic acids to particles. For example, the nucleic acids sequences used as probes can include carboxyl, amine, amino,

carboxylate, halide, ester, alcohol, carbamide, aldehyde, chloromethyl, sulfur oxide, nitrogen oxide, epoxy and/or tosyl functional groups.

[0097] Probes, including composite probes, attached to a substrate can be single-stranded and/or double-stranded nucleic acids. In particular embodiments, where double-stranded nucleic acids are bound, they are denatured and rendered single stranded after immobilization to the substrate for preparation for use in certain embodiments of assay methods. Optionally, double stranded nucleic acid probes are denatured prior to immobilization and the single stranded nucleic acids are then bound to the substrate.

[0098] Amplicon Reagent Compositions

[0099] In particular embodiments, a reagent for assaying nucleic acids is provided which includes a plurality of encoded particles having attached amplicons as probes.

[00100] In further particular embodiments, the attached amplicons are amplified from more than one template, making up a composite probe.

[00101] In certain embodiments, the amplicons attached to the plurality of encoded particles each include a nucleic acid sequence identical or completely complementary to a portion of a template genomic nucleic acid and together the amplicons represent substantially the entire template genomic nucleic acid.

[00102] Figure 1 illustrates an embodiment of a process for making a reagent for assaying genomic DNA. As indicated in Figure 1, a template nucleic acid is provided, 1. The template is amplified, 2, in a first amplification reaction using degenerate oligonucleotide primers (DOP) to produce a first amplification product.

[00103] The template nucleic acid can be any nucleic acid capable of being copied using a nucleic acid amplification method.

[00104] The template DNA for this first amplification reaction is optionally genomic DNA, typically having a size in the range of about 20 – 300 kb, although the template can be smaller or larger. The term “genomic” refers to DNA of the genome of a cell or organism and includes DNA isolated directly from a cell or organism, such as microdissected chromosomal DNA, as well as DNA copied from DNA of the genome of a cell or organism, such as cloned DNA. The template DNA can encompass all or part of a genome of a cell or organism. The template DNA can encompass DNA representing one or more chromosomes, a portion of a chromosome, a genetic locus, a gene or a portion of a gene. The template DNA can be in any form, such as an insert in a vector illustratively including a bacterial artificial chromosome, yeast artificial chromosome, human artificial chromosome, cosmid, plasmid, phagemid, phage DNA or fosmid. Template DNA can be in the form of microdissected chromosomal DNA. Thus, while specific

examples described herein refer to BACs as sources of template DNA, other types of clones such as PACs, YACs, cosmids, fosmids, cDNAs and the like may be used.

[00105] Multiple templates from any of these or other sources can be amplified together or separately and the combined amplicons from the multiple templates are a composite probe according to particular embodiments.

[00106] Template genomic DNA is obtained by methods known in the art, for instance, as described in J. Sambrook and D.W. Russell, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press; 3rd Ed., 2001 or F.M. Ausubel, Ed., Short Protocols in Molecular Biology, Current Protocols; 5th Ed., 2002. Template DNA may also be obtained commercially and/or using commercial kits for isolation of genomic DNA.

[00107] Amplification of template DNA is achieved using an in vitro amplification method. The term "amplification method" refers to a method or technique for copying a template nucleic acid, thereby producing nucleic acids including copies of all or a portion of the template nucleic acid, the produced nucleic acids also termed amplicons.

[00108] Amplicons optionally contain nucleic acid sequences present in the primers and not present in the original DNA template. Such primer-derived nucleic acids add functionality such as primer binding sites for additional amplification reactions and/or a functional group for chemical bonding to a substrate.

[00109] Amplification methods illustratively including PCR, ligation-mediated PCR (LM-PCR), phi-29 PCR, and other nucleic acid amplification methods, for instance, as described in C.W. Dieffenbach et al., PCR Primer: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 2003; and V. Demidov et al., DNA Amplification: Current Technologies and Applications, Taylor & Francis, 2004.

[00110] Many combinations of particular DNA template sources and nucleic acid amplification methods may be used.

[00111] The term "oligonucleotide primer" refers to a nucleic acid that is capable of acting as a site of initiation of synthesis of a primer extension product under appropriate reaction conditions. An oligonucleotide primer is typically about 10 – 30 contiguous nucleotides in length. An oligonucleotide primer is completely or substantially complementary to a region of a template nucleic acid such that, under hybridization conditions, the oligonucleotide primer anneals to the complementary region of the template nucleic acid. Appropriate reactions conditions for synthesis of a primer extension product include presence of suitable reaction components including, but not limited to, a polymerase and nucleotide triphosphates. Design of

oligonucleotide primers suitable for use in amplification reactions is well known in the art, for instance as described in A. Yuryev et al., PCR Primer Design, Humana Press, 2007.

[00112] The term "degenerate oligonucleotide primer" refers to a primer which includes a nucleic acid having a random or semi-random nucleotide sequence. Design of degenerate
5 oligonucleotide primers suitable for particular nucleic acid amplification reactions is well known in the art for instance as described in A. Yuryev et al., PCR Primer Design, Humana Press, 2007. Random or semi-random nucleotide sequences having about 5-8 nucleotides can be used. In further embodiments, random or semi-random nucleotide hexamers are included in degenerate oligonucleotide primers used in the first amplification.

10 **[00113]** The degenerate oligonucleotide primers used in particular embodiments each include a 5' constant DNA segment, an intermediate random DNA segment and a 3' anchor segment, for example as described in Fiegler et al., Genes Chromosomes Cancer, 36(4):361-74, 2003; and Telenius, et al., Genomics 13:718-25, 1992. The 5' constant DNA segment optionally has the same nucleotide sequence in all of the DOPs. The 3' anchor segment optionally has a nucleotide
15 sequence determined to have a desired frequency of occurrence in the template nucleic acid. Analysis of frequency of occurrence of a particular nucleic acid sequence is well known in the art, for example, as described in Milosavljevic, A. and Jurka, J., 1993, Comput. Applic. Biosci., 9:407-411; Pesole, G. et al., 1992, Nucleic Acids, Res., 20 :2871-2875 ; and Hutchinson, G. B., 1996, Comput. Appl. Biosci., 12 :391-398.

20 **[00114]** In particular embodiments the DOPs include about 17-25 contiguous nucleotides, of which about 7-12 contiguous nucleotides are included in the 5' constant DNA segment, about 5-8 contiguous nucleotides are included in the random DNA segment and about 5-8 contiguous nucleotides are included in the 3' anchor segment.

[00115] The first amplification reaction yields a first reaction product containing a plurality
25 of amplicons. Each individual amplicon in the first reaction product includes a DNA sequence identical or completely complementary to a random portion of the DNA template and a DNA sequence identical to the 5' constant DNA sequence of the first reaction primers.

[00116] The term "complementary" as used herein refers to Watson-Crick base pairing between nucleotides and specifically refers to nucleotides hydrogen bonded to one another with
30 thymine or uracil residues linked to adenine residues by two hydrogen bonds and cytosine and guanine residues linked by three hydrogen bonds. In general, a nucleic acid includes a nucleotide sequence described as having a "percent complementarity" to a specified second nucleotide sequence. For example, a nucleotide sequence may have 80%, 90%, or 100% complementarity to a specified second nucleotide sequence, indicating that 8 of 10, 9 of 10 or 10

of 10 nucleotides of a sequence are complementary to the specified second nucleotide sequence. For instance, the nucleotide sequence 3'-TCGA-5' is 100% complementary to the nucleotide sequence 5'-AGCT-3'. Further, the nucleotide sequence 3'-TCGA- is 100%, or completely, complementary to a region of the nucleotide sequence 5'-TTAGCTGG-3'.

5 **[00117]** Referring to Figure 1, a second amplification reaction, 3, is performed using the first reaction product amplicons as template DNA. The second amplification reaction, 3, includes a "universal" oligonucleotide primer, so-called since the universal primer is identical or completely complementary to the 5' constant DNA segment of the DOP used in the first amplification reaction. A universal oligonucleotide primer includes the 5' constant DNA
10 segment of the DOP used in the first amplification reaction positioned at the 3' end of the universal primer. A universal oligonucleotide primer optionally includes additional contiguous nucleotides at the 5' end of the primer.

[00118] In a particular option, a universal oligonucleotide primer includes a functional group at the 5' terminus of the primer for attachment of the amplicons resulting from the second
15 amplification reaction to an encoded solid or semi-solid substrate such as encoded particles. For example, the universal oligonucleotide primers include an amine group at the 5' terminus of the primer. In a further option, amplicons resulting from the second amplification reaction can be modified to include a functional group for bonding to a solid or semi-solid substrate Modification of a nucleic acid to include a functional group capable of bonding to a solid or
20 semi-solid substrate is well known in the art.

[00119] In particular embodiments, each individual amplicon attached to a particle includes a DNA segment identical to a random portion of the template DNA sequence. Each individual amplicon also contains a constant DNA segment contiguous with the DNA segment identical to a random portion of the template DNA sequence. The constant DNA segment of the amplicon
25 optionally includes a terminal functional group for attachment of the amplicon to an encoded particle. In a particular embodiment, the constant DNA segment of the amplicon includes a 5' terminal amine group for attachment of the amplicon to an encoded particle.

[00120] As shown in Figure 1, the amplicons of the second reaction product are immobilized, 4, on a first plurality of encoded particles. Binding of the amplicons of the second amplification
30 reaction to the encoded particles is achieved by any of various methods effective to bond a nucleic acid to a solid or semi-solid substrate, illustratively including adsorption and chemical bonding. The amplicons can be bonded directly to the material of the encoded particles or indirectly bonded to the encoded particles, for example, via bonding to a coating or linker disposed on the particles. Amplicons can be synthesized, and/or modified once synthesized, to

include a functional group for use in bonding the amplicons to particles. For example, amplicons can include carboxyl, amine, amino, carboxylate, halide, ester, alcohol, carbamide, aldehyde, chloromethyl, sulfur oxide, nitrogen oxide, epoxy and/or tosyl functional groups.

[00121] In general, the amplicons which are the product of the second amplification reaction are double stranded and the double stranded amplicons are attached to the particles. Thus, both strands of the double stranded amplicons are represented on each particle. The amplicons are denatured and rendered single stranded after immobilization to the particles for preparation for use in particular embodiments of assay methods. Optionally, double stranded amplicons are denatured prior to immobilization and the single stranded amplicons are then bound to particles.

[00122] As described, each individual amplicon of both the first and second amplification reactions contains a nucleic acid sequence identical to a random portion of the template DNA sequence such that the amplicons produced by the first amplification reaction together represent substantially the entire template DNA sequence and the amplicons produced by the second amplification reaction together represent substantially the entire template DNA sequence.

[00123] Encoded particles having bound amplicons which are the product of a second amplification reaction and which together represent substantially the entire genomic DNA sequence used as a template in the first amplification reaction are a first particle set and a first reagent for assaying genomic DNA.

[00124] In particular embodiments, each individual amplicon attached to a particle has a length in the range of about 500 – 1200 nucleotides, inclusive. Thus, a relatively large template nucleic acid is represented substantially entirely on a set of encoded particles by the attached relatively smaller amplicons amplified from the template.

[00125] As noted above, each particle set includes encoded particles having bound amplicons which are the product of a second amplification reaction and which together represent substantially the entire genomic DNA sequence used as a template in a first amplification reaction. The number of particles including amplicons which is sufficient to together represent substantially the entire genomic DNA sequence used as a template in the first amplification reaction depends on a number of factors such as the size of the template, the size of the amplicons and the number of binding sites available for binding an amplicon on a particle. In general, the number of particles sufficient to together represent substantially the entire genomic DNA sequence used as a template in the first amplification reaction is in the range of about 1-10,000, inclusive.

[00126] Additional particle sets are generated by amplification using a second genomic DNA template and binding the amplicons which are the reaction product of a second amplification

reaction as described above to a second plurality of encoded particles. The second plurality of encoded particles is detectably different than the first plurality of encoded particles, thereby generating a second encoded particle set and a second reagent for assaying genomic DNA.

[00127] Similarly, a third or subsequent genomic DNA template is used to generate the reaction product of an amplification reaction and the reaction product is bound to a third or subsequent plurality of encoded particles. Each of the third or subsequent plurality of encoded particles is detectably different than each other plurality of encoded particles, yielding a third or subsequent encoded particle set and a third or subsequent reagent for assaying genomic DNA.

[00128] Multiplex reagent

[00129] A multiplex reagent for assaying genomic DNA is provided according to certain embodiments which includes a mixture of two or more particle sets. The individual encoded particles of each encoded particle set are detectably distinguishable from individual encoded particles of each other encoded particle set in particular embodiments.

[00130] In particular embodiments, at least one particle set includes a composite probe. Optionally, more than one particle set include a composite probe.

[00131] In particular embodiments, each encoded particle set has attached amplicons which are the product of a second amplification reaction as described herein and which together represent substantially the entire genomic DNA sequence used as a template in a first amplification reaction, wherein a different genomic template is represented by amplicons attached to each other encoded particle set.

[00132] A multiplex reagent according to a specific embodiment includes a first encoded particle set having attached amplicons which together represent substantially an entire template DNA sequence inserted in the first bacterial artificial chromosome and a second encoded particle set having attached amplicons which together represent substantially an entire template DNA sequence inserted in the second bacterial artificial chromosome.

[00133] For example, a first encoded particle set has attached amplicons including nucleic acid sequences identical to a portion of human chromosome 13 DNA and a second encoded particle set has attached amplicons including nucleic acid sequences identical to a portion of chromosome 18 human DNA. Third or subsequent encoded particle set have attached amplicons including nucleic acid sequences identical to human DNA from another chromosome or another non-overlapping region of a chromosome.

[00134] A multiplex reagent described herein allows for simultaneous assay of multiple targets, such as multiple genomic loci, in a single assay.

[00135] A multiplex reagent for assaying genomic DNA is generated by mixing at least a first encoded particle set and a second encoded particle set.

[00136] Figure 2 illustrates an embodiment of a method of generating a multiplex reagent. As indicated in the figure, any number, "m" of encoded particle sets can be included in the multiplex reagent. Thus, for example, "m" can be at least 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or 200 different encoded particle sets. A set of encoded particles having bound amplicons is combined with one or more additional sets of encoded particles having bound amplicons to generate a multiplex reagent for assay of genomic gain and loss in a sample.

[00137] In a particular embodiment, a set of encoded particles including a bound composite probe is combined with one or more additional sets of encoded particles having bound composite or non-composite probes to generate a multiplex reagent for assay of genomic gain and loss in a sample.

[00138] Assay Methods

[00139] A method of assaying a DNA sample is provided according to particular embodiments which includes providing a substrate-attached composite nucleic acid probe. The composite nucleic acid probe includes nucleic acid sequences that specifically hybridize to two or more genomic loci in a genomic region of a reference genome. The genomic region is characterized by a first terminus and a second terminus with an intermediate region of at least 400 kilobases disposed between the termini. The composite nucleic acid probe includes nucleic acid sequences which specifically hybridize to substantially an entire first genomic locus which includes the first terminus of the genomic region and nucleic acid sequences which specifically hybridize to substantially an entire second genomic locus which includes the second terminus. The first genomic locus and second genomic locus each typically include at least about 100 kilobases. Optionally, nucleic acid sequences of the composite probe specifically hybridize to additional loci within the genomic region. In a further option, the reference genome is one or more human genomes.

[00140] The substrate-attached composite nucleic acid probe is hybridized with sample genomic DNA at a stringency sufficient to achieve specific hybridization. The substrate-attached composite nucleic acid probe is also hybridized with reference genomic DNA under the same or similar conditions to allow for comparison.

[00141] A first signal indicating specific hybridization of the substrate-attached composite nucleic acid probe with the sample genomic DNA is detected along with a second signal indicating specific hybridization of the substrate-attached composite nucleic acid probe with the

reference genomic DNA. The first signal and the second signal are compared to detect differences between the first and second signals, indicative of differences between the sample DNA and the reference DNA.

[00142] Particular embodiments of methods of assaying sample nucleic acid include providing a multiplex reagent including a mixture of two or more encoded particle sets encoded such that each particle of each encoded particle set is detectably distinguishable from each particle of each other encoded particle set. The encoded particles include attached nucleic acid sequences which specifically hybridize to at least one genomic locus of a reference genome, and at least one encoded particle set includes an attached composite nucleic acid probe.

[00143] The multiplex reagent is hybridized with sample genomic nucleic acid and with reference nucleic acid, together or in parallel.

[00144] A first signal is detected which indicates specific hybridization of the attached nucleic acid sequences with detectably labeled sample nucleic acid and a second signal is detected indicating specific hybridization of the attached nucleic acid sequences with detectably labeled reference nucleic acid. The encoded particles are then identified so as to associate particle encoding with the first signal or with the second signal. The first signal and the second signal for each encoded particle set are then compared and differences in the first and second signals are indicative of differences between the sample and reference nucleic acids.

[00145] In particular embodiments, methods of assaying genomic DNA include providing encoded particles having attached amplicons which together represent substantially an entire template genomic nucleic acid. In particular embodiments, encoded particles having attached amplicons are provided which together represent more than one copy of substantially an entire template genomic nucleic acid.

[00146] In particular embodiments, a sample of genomic DNA to be assayed for genomic gain and/or loss is labeled with a detectable label. Reference DNA is also labeled with a detectable label for comparison to the sample DNA. The sample and reference DNA can be labeled with the same or different detectable labels depending on the assay configuration used. For example, sample and reference DNA labeled with different detectable labels can be used together in the same container for hybridization with amplicons attached to encoded particles in particular embodiments. In further embodiments, sample and reference DNA labeled with the same detectable labels can be used in separate containers for hybridization with amplicons attached to particles.

[00147] The term “detectable label” refers to any atom or moiety that can provide a detectable signal and which can be attached to a nucleic acid. Examples of such detectable

labels include fluorescent moieties, chemiluminescent moieties, bioluminescent moieties, ligands, magnetic particles, enzymes, enzyme substrates, radioisotopes and chromophores.

[00148] Any of various methods of labeling sample and reference nucleic acids, such as DNA, may be used in the assay, such as nick translation or chemical labeling of the nucleic acids. For example, a detectable label can be introduced by polymerization using nucleotides that include at least some modified nucleotides, such as nucleotides modified to include biotin, digoxygenin, fluorescein, or cyanine. In some embodiments, the detectable label is introduced by random-priming and polymerization. Other examples include nick translation (Roche Applied Science, Indianapolis Ind.; Invitrogen, Carlsbad Calif.) and chemical labeling (Kreatech ULS, Amsterdam NL). Detectable labeling of nucleic acids is well known in the art and any labeling method appropriate for labeling nucleic acids, such as genomic DNA, can be used.

[00149] In yet another embodiment, covalent labeling of sample and reference nucleic acids, such as DNA, individually with a detectable label is avoided. For example, unlabeled genomic DNA samples are hybridized to the amplicons immobilized to the encoded particles. Pre-labeled reporter sequences are also hybridized to the amplicon-sample DNA complexes and amplicon-reference DNA complexes at sequences adjacent to but not overlapping the sequences of the capture probes of the amplicons. These labeled reporter sequences can be hybridized in the same or in a different hybridization reaction. In this manner the labeled reporter sequences can be manufactured in bulk in a larger-scale environment, lowering the cost per assay compared to individually labeling each sample at the time of the assay.

[00150] The “sample” and “reference” nucleic acids, such as genomic DNA, can be obtained from any suitable source. Particular methods described herein involve using sample genomic DNA from an individual subject. Genomic sample and/or reference DNA can be extracted from almost any tissue including, but not limited to, blood, amniotic fluid, solid tumors, organ biopsies, cheek swabs, chorionic villae, blastocysts and blastomeres, products of conception, saliva, urine and the like. Archived samples extracted from formalin-fixed, paraffin-embedded (FFPE) pathology samples are also sources of sample genomic DNA assayed by this method. Sample and/or reference genomic DNA can also be obtained from in vitro sources such as cell lines. Methods of obtaining genomic DNA from these or other sources are well known in the art.

[00151] In particular embodiments, reference DNA is characterized with respect to a particular characteristic of the sample DNA to be assayed. For example, where sample DNA is to be assayed to detect duplication of a particular gene or chromosomal locus, the reference DNA is characterized so that it is known how many copies of the gene or locus are contained in the reference DNA. In general, sample and reference DNA from the same species are used.

[00152] Reference DNA can be a pooled mixture of genomic DNA derived from a plurality of normal subjects, particularly human subjects, of the same gender. DNA pooled from a plurality of normal subjects can be obtained commercially.

[00153] In some embodiments, more than one reference DNA is used and additional
5 information is obtained in a process using additional reference DNAs.

[00154] Thus, for example, in a particular embodiment, two reference genomic DNA samples are compared to a test sample of genomic DNA. A first reference genomic DNA obtained from a male subject and a second reference genomic DNA obtained from a female subject are compared to a test sample of genomic DNA obtained from a subject whose gender is to be
10 determined, such as a prenatal fetus.

[00155] A particular characteristic of an assay of the present invention which includes comparision of at least two reference genomic DNA samples to a test sample of genomic DNA is decreased ambiguity and increased confidence in the assay results. For example, as shown in Figure 14, an ambiguous result which would have been achieved with only a male-specific
15 reference is clarified when the assay is performed using both a male-specific reference and a female-specific reference.

[00156] An assay described herein can be used to detect or characterize disorders associated with chromosomal gains or losses. Constitutional, or inborn, disorders include trisomies of entire chromosomes, amplifications or deletions of smaller genomic loci (approximately 200
20 kilobases to 20 megabases), and amplifications or deletions in the sub-telomeric or centromeric regions. Various cancers are also characterized by chromosomal gains and losses that may correlate with type, stage, drug resistance, or therapy response. Laboratory cell lines, including stem-cell lines, may be characterized for chromosomal stability using the present method.

[00157] Thus, two or more reference genomic DNA samples can be used which represent
25 different stages of a progressive disease, condition or disorder which affects genomic DNA and the two or more references are compared with a test sample of genomic DNA of an individual subject whose condition relative to the references is to be determined. For example, various cancers, other disorders and/or age are associated with progressive deletion or genomic DNA, such as mitochondrial DNA deletions and telomere shortening.

30 [00158] While methods and compositions are described herein primarily with reference to nucleic acids derived from humans, it is appreciated that methods and compositions described herein may be used to assay sample genomic DNA from any of various organisms including, but not limited to, non-human primates, rodents, rabbits, dogs, cats, horses, cattle, pigs, goats and sheep. Non-mammalian sources of sample DNA can also be assayed, illustratively including

fish and other aquatic organisms, birds, poultry, bacteria, viruses, plants, insects, reptiles, amphibians, fungi and mycobacteria. Similarly, reference DNA can be human DNA or DNA from any of various organisms including, but not limited to, non-human primates, rodents, rabbits, dogs, cats, horses, cattle, pigs, goats, sheep and non-mammalian sources illustratively
5 including fish and other aquatic organisms, birds, poultry, bacteria, viruses, plants, insects, reptiles, amphibians, fungi and mycobacteria.

[00159] The substrate-attached nucleic acid probes are hybridized with detectably labeled sample genomic DNA of an individual subject so as to achieve specific hybridization of the substrate-attached nucleic acid probes and the sample and/or reference nucleic acids.

10 **[00160]** In particular embodiments of assays described herein, amplicons attached to the encoded particles are hybridized with detectably labeled sample genomic DNA of an individual subject so as to achieve specific hybridization of the amplicon DNA and the detectably labeled sample genomic DNA. In addition, DNA sequences attached to the encoded particles are hybridized with detectably labeled reference genomic DNA so as to achieve specific
15 hybridization of the amplicon DNA and the detectably labeled reference genomic DNA.

[00161] The terms "hybridization" and "hybridized" refer to pairing and binding of complementary nucleic acids. Hybridization occurs to varying extents between two nucleic acids depending on factors such as the degree of complementarity of the nucleic acids, the melting temperature, Tm, of the nucleic acids and the stringency of hybridization conditions, as
20 is well known in the art. The term "stringency of hybridization conditions" refers to conditions of temperature, ionic strength, and composition of a hybridization medium with respect to particular common additives such as formamide and Denhardt's solution. Determination of particular hybridization conditions relating to a specified nucleic acid is routine and is well known in the art, for instance, as described in J. Sambrook and D.W. Russell, Molecular
25 Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press; 3rd Ed., 2001; and F.M. Ausubel, Ed., Short Protocols in Molecular Biology, Current Protocols; 5th Ed., 2002. High stringency hybridization conditions are those which only allow hybridization of substantially complementary nucleic acids. Typically, nucleic acids having about 85-100% complementarity are considered highly complementary and hybridize under high stringency conditions.
30 Intermediate stringency conditions are exemplified by conditions under which nucleic acids having intermediate complementarity, about 50-84% complementarity, as well as those having a high degree of complementarity, hybridize. In contrast, low stringency hybridization conditions are those in which nucleic acids having a low degree of complementarity hybridize. The terms "specific hybridization" and "specifically hybridizes" refer to hybridization of a particular

nucleic acid to a target nucleic acid without substantial hybridization to nucleic acids other than the target nucleic acid in a sample.

[00162] Assays described can be performed in any suitable container. In particular embodiments, for example, where multiple samples are to be assayed, a multi-chamber container can be used. Multi-chamber containers illustratively include multi-depression substrates such as slides, silicon chips or trays. In some embodiments, each sample is disposed in a different well of a multi-well plate. For example, a multi-well plate can be a 96-well, 384-well, 1024-well or 1536-well assay plate.

[00163] Further included is detection of a first signal indicating specific hybridization of the attached DNA sequences with detectably labeled genomic DNA of an individual subject and detection of a second signal indicating specific hybridization of the attached DNA sequences with detectably labeled reference genomic DNA.

[00164] Any appropriate method, illustratively including spectroscopic, optical, photochemical, biochemical, enzymatic, electrical and/or immunochemical is used to detect a signal in an assay described herein.

[00165] Signals that are indicative of the extent of hybridization can be detected, for each particle, by evaluating signal from one or more detectable labels. Particles are typically evaluated individually. For example, the particles can be passed through a flow cytometer. Exemplary flow cytometers include the Coulter Elite-ESP flow cytometer, or FACScan.TM. flow cytometer available from Beckman Coulter, Inc. (Fullerton Calif.) and the MOFLO.TM. flow cytometer available from Cytomation, Inc., Fort Collins, Colo. In addition to flow cytometry, a centrifuge may be used as the instrument to separate and classify the particles. A suitable system is that described in U.S. Pat. No. 5,926,387. In addition to flow cytometry and centrifugation, a free-flow electrophoresis apparatus may be used as the instrument to separate and classify the particles. A suitable system is that described in U.S. Pat. No. 4,310,408. The particles may also be placed on a surface and scanned or imaged.

[00166] In certain embodiments, a first signal is detected indicating specific hybridization of the substrate-attached nucleic acid sequences with detectably labeled sample nucleic acids, such as genomic DNA of an individual subject. A second signal is also detected indicating specific hybridization of the substrate-attached nucleic acid sequences with detectably labeled reference genomic DNA.

[00167] In further embodiments, a first signal is detected indicating specific hybridization of the encoded particle attached DNA sequences with detectably labeled genomic DNA of an

individual subject. A second signal is also detected indicating specific hybridization of the encoded particle attached DNA sequences with detectably labeled reference genomic DNA.

[00168] The first signal and the second signal are compared, yielding information about the sample and reference nucleic acids. In certain embodiments, first signal and the second signal are compared, yielding information about the genomic DNA of the individual subject compared to the reference genomic DNA.

[00169] In particular embodiments, a ratio of the signals from the detectable labels of the reference DNA and the sample DNA hybridized to the amplicons of one or more particle sets is used to evaluate differences between the sample and reference DNA, indicative, for instance, of genomic gain and/or loss.

[00170] In certain embodiments, the reference DNA and the sample DNA are hybridized to the probes, such as amplicons, of one or more particle sets in the same container, such as a well of a multi-well plate. After hybridization, the two labels are analysed together, i.e. both detectable labels are detected in the hybridized material or the hybridized material is divided into two (or more) portions and each portion is evaluated separately to detect the detectable labels. Results from the evaluation can be used to provide the ratio of signals from the two detectable labels. This approach allows use of competitive hybridization to normalize any variation between assays: both of the reference and experimental samples are assayed simultaneously in the same vessel mixed with the same particles.

[00171] Optionally, the detectably labeled reference DNA and the detectably labeled sample DNA are hybridized to one or more particle sets in the different containers, such as different wells of a multi-well plate. In a further example, the detectably labeled reference DNA and the detectably labeled sample DNA are hybridized to one or more probes attached at different locations on a planar array. A ratio of signals from the two detectable labels can be obtained to evaluate differences between the sample DNA and reference DNA. When this approach is utilized, a single reference sample can be shared between several or many experimental samples. For experiments involving multiple samples per day there can be a savings on reagent cost and labor by avoiding the labeling of multiple duplicate normal samples. Also it is unnecessary to manipulate the sample to obtain different portions for separate analysis. Each sample can be evaluated only once.

[00172] Encoded particles are identified by their encoded information so as to associate particle encoding with the first signal and with the second signal in particular embodiments. Thus, for example, first and second signals are associated with encoded particles of a first encoded particle set containing human DNA from chromosome 13. The first signal and the

second signal associated with the first encoded particle set are compared, yielding information about the chromosome 13 DNA of the individual subject compared to the chromosome 13 reference DNA. Similarly, first and second signals are associated with a second encoded particle set containing human DNA from chromosome 18. The first signal and the second signal associated with the second encoded particle set are compared, yielding information about the chromosome 18 DNA of the individual subject compared to the chromosome 18 reference DNA.

[00173] The figures and descriptions herein illustrate the best mode but many alternative materials and processes can be substituted. One of skill in the art will recognize appropriate alternative materials and processes and will be able to make and use the compositions and methods described without undue experimentation.

[00174] The compositions of the various buffers and other assay components may be substituted.

[00175] The conditions for culturing, purification amplification, denaturation, coupling, hybridization, reporter binding, washing, and bead handling can all be varied by the user to suit particular types of cells, template genomic DNA, samples, selected reporters and the like.

[00176] The assay in examples herein performs well with as little as 30 ng of sample DNA. In situations where the biological source yields insufficient DNA for the described assay the sample can be amplified by a variety of whole-genome amplification (WGA) methods, such as DOP PCR or phi-29 PCR. When utilizing WGA-processed samples, the reference DNA can be processed by the same method so that any sequence-specific amplification bias will be largely corrected by the sample/reference ratio of signals.

[00177] Kits for assaying DNA are provided. In particular embodiments, a kit is provided which includes an encoded particle set and/or a mixture of two or more encoded particle sets. Instructional material for use of the encoded particle set and/or multiplex reagent including two or more encoded particle sets is optionally included in a kit. An ancillary reagent such as buffers, enzymes, washing solutions, hybridization solutions, detectable labels, detection reagents and the like are also optionally included.

[00178] Embodiments of assay compositions and methods are illustrated in the following examples. These examples are provided for illustrative purposes and are not considered limitations on the scope of compositions and methods.

Examples

[00179] Example 1

[00180] Preparation of a bead set reagent for genomic DNA assay

[00181] Fig. 1A shows a flowchart illustrating preparation of BAC amplicons from a single BAC clone and immobilizing the amplicons as probes onto a set of encoded beads. In this example, the beads in the set all have the same ID code.

[00182] The starting material is living BAC clone material, 10, a long (100-200 kilobases typically) human DNA sequence inserted into the genome of an E. coli bacteria cell. A small chip of frozen BAC glycerol stock material is picked and used as the starting material for a standard bacterial cell culture process, 11. The cells are cultured in 35 ml medium in 50 ml tubes overnight at 37°C with a selective antibiotic according to a standard BAC culture protocol. The cultured cells are then centrifuged to the bottom of the tube at 4°C for 20 minutes and the supernatant withdrawn and discarded. The cell pellet is resuspended in a buffer containing RNase, and then lysed using LyseBlue (Qiagen, Valencia CA) and SDS. The lysate, 12, is centrifuged, 13, at approximately 20,000 g for 30 minutes, and the supernatant, containing the DNA in solution, is collected and the pellet discarded. The centrifugation is repeated for 15 minutes on the supernatant. The clear supernatant containing the dissolved BAC DNA is collected, while the cellular debris, proteins and other impurities are driven to the bottom of the tube and discarded. The BAC DNA is extracted and purified, 17, from the supernatant using a Qiagen Genomic-Tip 20/G column purification kit. This kit comprises purification columns, 15, and wash and elution buffers, 16. After elution, the now highly purified BAC DNA is precipitated and into pellets by isopropanol, 19, precipitation. The yield is typically 20 to 200 ng of purified BAC DNA, 18. This BAC DNA can be stored as a dried pellet or resuspended in water for use immediately in the next steps.

[00183] A quantity of PCR amplicons representing substantially the entire sequence content of each BAC DNA is then produced using two rounds of polymerase chain reaction (PCR) amplification. The first round of PCR, 20, is non-specific degenerate oligonucleotide primer (DOP) PCR using a DOP primer mix, 21, a DOP PCR polymerase, 22, and DOP PCR buffer, 23, with the above prepared BAC DNA, 18, used as template. The second round of PCR amplification, 25, utilized a single primer directed at the known sequence motifs of the DOP primers. Two rounds of PCR are used to generate yields of approximately 20 μg of final amplicon product, 29, for subsequently coupling, 32, the amplicons, 29, to encoded beads, 30.

[00184] The amplicons are prepared as follows.

[00185] A first 50 μl DOP PCR mix is made for each BAC DNA comprised of:

10X DOP PCR Buffer	5.0 μl
10mM dNTP's (each)	1.0 μl
50mM MgCl	5.0 μl
10uM DOP Primer Mix (each)	10.0 μl

20 to 50 ng BAC DNA Template	2.0 µl
Platinum Taq polymerase	0.5 µl
Water	21.5 µl
Total Volume	50.0 µl

5

[00186] The DOP PCR buffer, 23, included 20 mM Tris HCL (pH 8.4), 50 mM KCl and 5 mM MgCl. The dNTPs (Amersham Biosciences, Piscataway NJ) are at a concentration of 200 µM. The platinum TAQ polymerase (Applied BioSystems) is at a concentration of 5 units/µl. The DOP primer mix, 21, see Fiegler et al. 2003, Genes Chromosomes Cancer, 36(4):361-74, 10 included three sets of degenerate oligonucleotides of the following 22-mer sequences (Operon Biotechnologies, Huntsville AL), wherein the Ns represent randomized nucleotides:

5' CCGACTCGAGNNNNNNCTAGAA 3' SEQ ID No. 1

5' CCGACTCGAGNNNNNNTAGGAG 3' SEQ ID No. 2

5' CCGACTCGAGNNNNNNTTCTAG 3' SEQ ID No. 3

15 wherein N denotes random nucleotides.

[00187] The BAC DNA template, 18, dissolved in water, is purified by column purification, 17, using Qiagen Genomic-Tip 20/G column purification kit. The Platinum Taq polymerase, 22 (Invitrogen, Carlsbad CA) is at a concentration of 5 units/µl.

[00188] The first-round amplification, 20, is performed in a GeneAmp 9700 themocycler 20 (Applied BioSystems, Foster City CA) according to the following temperature/time profile:

3.0 min	94°C	
1.5 min	94°C	
2.5 min	30°C	9 Cycles
0.10C/sec	72°C (ramp)	
3.0 min	72°C	
1.0 min	94°C	
1.5 min	62°C	30 Cycles
2.0 min	72°C	
8.0 min	72°C	
4.0°C		(steady state)

25

30

[00189] The amplicon products, 24, from this first round of DOP PCR, 20, are then used as the templates for a second round of PCR, 25. The single primer, 26, in the second round is specific to the common sequence portions of the DOP primers, 21 used in the first round, 20. 35 This primer, 26, is amine-modified so that the resulting amplicons, 29, would also have an amine group on one end to facilitate simple coupling to the encoded beads in a subsequent step, 32.

[00190] The second round PCR is performed as follows.

[00191] A second 100 µl PCR mix is made for each BAC amplicon template including:

10X PCR Buffer	10.0 µl

10mM dNTP's (each)	2.0 μl
50mM MgCl	10.0 μl
10uM Amine Primer	15.0 μl
Template (from PCR #1)	2.0 μl
Platinum Taq	0.5 μl
Water	58.5 μl
Total Volume	100.0 μl

[00192] The PCR 2 buffer, 28, included 20 mM Tris HCL (pH 8.4), 50 mM KCl and 5 mM MgCl. The dNTPs (Amersham Biosciences, Piscataway NJ) are at a concentration of 200 μM. The platinum TAQ polymerase (Applied BioSystems) is at a concentration of 5 units/μl.

[00193] The amine-linked primer (Operon) had the following sequence.

5'-GGAAACAGCCCGACTCGAG-3' SEQ ID No. 4

[00194] The templates in reaction, 25, are the DOP amplicons, 24, from the previous DOP PCR round, 20. The second-round amplification, 25, is performed in a GeneAmp 9700 themocycler (Applied BioSystems) according to the following temperature/time profile:

10 min	95°C	
1.0 min	95°C	
1.5 min	60°C	35 Cycles
7.0 min	72°C	
10 min	72°C	
4.0°C		(steady state)

[00195] This second PCR product, 29, is then purified using a magnetic-bead based kit, 9, (PCR Clean Beads, Agencourt Bioscience Corp., Beverley MA) according to the manufacturer's protocol. The purified amplicons, 29, are then resuspended in 40 μl water and stored at -20°C until used in the bead coupling step as described below.

[00196] The encoded bead coupling process, 32, to immobilize the amplicon product, 29, as probe DNA onto the surface of encoded beads is performed on Luminex carboxy beads, 30 (Luminex, Austin TX) at a scale of 50 μl of the standard bead concentration, yielding approximately 650,000 beads. The beads are made of polystyrene, approximately 5.6 μm in diameter, and encoded with controlled amounts of two or three fluorescent dyes to facilitate their bead ID being detected in a purpose-built flow cytometer reading instrument. 50 μl of suspended beads, 30, all of one bead ID or region, are transferred from the Luminex tube in which they are delivered to a 1.5 ml Eppendorf tube for the coupling, 32, with vortexing and sonication used to ensure suspension. The beads are then spun down at 12,000 RPM for 3 minutes and the bead buffer supernatant removed without disturbing the bead pellet. 25 μl of MES buffer is added to each tube of beads, followed by vortexing and sonication. Separately, 10 μg of PCR 2 amplicons, 29, from each BAC are then added to a second set of 1.5 ml

centrifuge tubes, and the DNA in each tube is then dried down completely in a SpeedVac (ThermoFisher Scientific, Waltham MA). One bead suspension is then transferred into each DNA tube, vortexed and sonicated for 5 seconds each to mix, keeping careful track of the bead ID (region) associated with each BAC.

5 **[00197]** Next, 1.5 μl of freshly dissolved EDC, 31, (1-ethyl-3-[dimethylaminopropyl]-carbodiimide hypocloride, Pierce, Rockford IL) at 10 mg/ml is added to each tube, vortexed immediately, and incubated for 30 minutes at room temperature in the dark (to preserve the Luminex beads' fluorescent encoding). Remixing is performed at the 15-minute point. The EDC addition, incubation, and remixing is then repeated for a second time.

10 **[00198]** 500 μl of TNT buffer (0.1M Tris pH 7.5, 0.15M NaCl, 0.02% Tween 20) is then added to each tube and vortexed. The tubes are then spun on a microfuge for 4 minutes at 12,000 RPM to drive the beads to the bottom and the supernatant carefully removed. Next, 500 μl of 0.1% SDS is added, and the beads again spun down for 4 minutes at 12,000 RPM and the supernatant carefully removed. Finally, 50 μl of 1x TE buffer (10 mM Tris pH 7.5, 1 mM

15 EDTA) to each tube and vortexed.

[00199] The bead set, 33, with immobilized amplicon probes, 29, can be included as a component of a multiplex bead set for use in assays of genomic DNA.

[00200] Example 2

[00201] Preparation of a multiplex encoded bead set reagent for DNA assay

20 **[00202]** Fig. 2 is a flowchart illustrating mixing *m* different encoded bead sets, each with its respective immobilized BAC-amplicon probe DNA, together to make a multiplexed encoded bead set.

[00203] Encoded bead sets 34, 35, 36, and 37 are forced into suspension by sonication, rotation of a tube container, vortexing or a similar method. A pipette is then used to transfer

25 aliquots of each bead set into another vessel where the individual bead sets are combined and mixed, followed by denaturation, 38, to facilitate subsequent hybridization to the probe DNA immobilized on the beads in an assay.

[00204] In a detailed example, the 50 μl contents of 2 or more bead sets, each in an individual tube, each encoded bead set with immobilized probe DNA, 33, are combined in

30 batches into one 1.5 ml centrifuge tube. After combining approximately 10 bead sets, the tube is spun down and the supernatant carefully removed, in order to keep the volume down. This is repeated until all of the bead sets are combined (up to 100 encoded bead IDs or regions are supported by the Luminex 200 system, for example).

[00205] After all of the bead sets are combined into a multiplex bead set the immobilized probe DNA is denatured. After spinning down the beads and removing the supernatant, 500 µl 0.1N NaOH is added and allowed to incubate for 2 minutes at room temperature. The beads are then spun down and the supernatant carefully removed. 500 µl of 10 mM Tris, 15 mM NaCL, 5 0.2% Tween 20 is added, the tube vortexed, then the beads spun down and the supernatant removed. This wash step is then repeated. Finally, the volume is brought to 500 µl with 1X TE buffer, and the multiplex bead set, 39, stored in the dark at 4°C until used for an assay.

[00206] Example 3

[00207] Multiplexed genomic gain and loss assay

10 [00208] Figure 3 is a flowchart illustrating an embodiment including running a multiplexed genomic gain and loss assay on *n* samples using a multiplexed encoded bead set. The flowchart shows embodiments of methods including providing labeled sample and reference DNA, 5, hybridization of the sample and reference DNA with two or more encoded bead sets, 6, detection of signals from the labeled sample and reference DNA hybridized to the encoded bead sets, 7, 15 and comparison of the signals to determine differences between the sample and reference DNA, 8.

[00209] Figure 3A is a flowchart illustrating an embodiment including running a multiplexed genomic gain & loss assay on *n* samples using a multiplexed encoded bead set.

[00210] In this example, two DNA samples and two references are being assayed in parallel. 20 In practice, several dozen samples may be run simultaneously in parallel in a microplate format. More or fewer samples and references than this number can be assayed in parallel.

[00211] In this example, the four DNA samples, 40 and 41 representing two references and 42 and 43 representing two assay samples, are enzymatically labeled with biotin and purified. Reference samples are typically normal male and female pooled samples, such as Human Female 25 Genomic DNA and Human Male Genomic DNA (Promega, Madison WI). Each DNA sample and reference is combined with biotin-labeled nucleotides, 45, (PerkinElmer, Boston MA), non-labeled nucleotides 49, (PerkinElmer), random primers, 47, (Operon, Biotechnologies, Huntsville AL) and a Klenow fragment polymerase enzyme, 46 (Epicentre Biotechnologies, Madison WI). After incubation, 44, the reaction product is cleaned up, 50, using a DNA column 30 purification kit, 49, such as a Purelink DNA Mini Kit (Invitrogen). Approximately 5 µl at approximately 200 ng/µl of labeled sample is used for subsequent hybridization in the assay.

[00212] Each biotin-labeled sample or reference, 51 – 54, is then hybridized, 55, with the probes immobilized on the beads of a multiplexed encoded bead set, 56. Approximately 500

beads from each bead set (each probe type) are used; in this 55-plex example a total of about 55 x 500 = 27,500 beads per hybridization is used.

[00213] Beads of each encoded bead set are distinguishable from beads of each of the other encoded bead sets due to the encoding. Each of the 55 bead sets includes a plurality of encoded beads having attached amplicons representing substantially an entire template genomic DNA fragment. The template DNA for each bead set represents a genomic locus listed in Figure 9.

[00214] A hybridization buffer containing Cot-1 DNA, formamide, dextran sulfate and 1.9x SSC is included in the hybridization reaction. The total volume is approximately 15 μl and the reactions are carried out in the wells of a rigid PCR-type microplate, such as the Bio-Rad HSP 9631 (Bio-Rad Laboratories, Hercules CA). The plate is sealed tightly to minimize evaporation using an aluminum foil sealer (MSF 1001, Bio-Rad). The hybridization incubation, 55, is performed overnight at 50°C in a microplate shaking incubator at 1150 rpm (Wallac NCS Incubator, PerkinElmer).

[00215] After the hybridization incubation, 55, the four multiplex bead sets hybridized to the four samples, 58-61, are ready for a hybridization wash, 53, followed by incubation with a fluorescent reporter, 65, and a reporter wash, 67. First, 100 μl wash buffer a (2X SSC, 50% formamide) is added to each well, the plate resealed and incubated in the shaking incubator with 1150 rpm agitation at 50°C for 20 minutes. The contents of each well is then transferred to a Millipore 0.46 μm HT filter plate (Millipore, Billerica MA). The liquid is then removed from each well by vacuum using a Millipore MSVMHTS00 vacuum manifold. Next, 100 μl of wash buffer b (2X SSC, 0.1% Igepal detergent) is added to each well, followed by another 20 minute 50°C shaking incubation and vacuum aspiration. Then, 100 μl of wash buffer c (0.2X SSC) is added to each well and the 20 minute 50°C shaking incubation is repeated, followed by vacuum aspiration.

[00216] 100 μl of 1X PhycoLink SA solution, the streptavidin-phycoerythrin reporter, 64, is then added to each well. This reporter solution is made from 2 μl 500X PhycoLink SA PJ13S (Prozyme, San Leandro CA) mixed into 1 ml of reporter diluent, where the diluent is 1X PBS, 0.1% BSA and 0.05% Tween 20. This reporter solution is incubated with the multiplex bead sets for 30 minutes at 25°C and 1050 RPM in the shaking incubator. After incubation, the solution is aspirated from the wells of the filter plate using the vacuum manifold as in the previous wash steps.

[00217] The beads are then washed twice, 67, with wash buffer d, 66, which is 1X PBS with 0.01% Tween 20. 100 μl is added to each well of the filter plate, then the liquid is vacuum aspirated through the filters in the bottoms of the plate wells. 100 μl is added a second time and

incubated in the shaking incubator for 2 minutes at 25°C at 1050 RPM. This second wash is not aspirated but used to suspend the beads for reading.

[00218] The four bead sets in the example, 68 – 71, are then ready to read, 72, on a Luminex 200 system (Luminex Corporation, Austin TX). The signals and bead IDs from the beads in each well are read in sequence, and the median fluorescence intensity of the first 50 beads of each bead ID (bead region) is recorded for each well or sample, and output in a data file, 73. There is no evidence of bead networking; the Luminex reader is set to analyze 50 beads of each region and no failures are recorded.

[00219] Figure 4 is a schematic diagram of a 96-well SBS-standard microplate, 80, showing example locations of duplicate references and duplicate samples for running the assay on 46 samples in parallel. Duplicate hybridizations of each labeled sample can be used to assure data generation in case of a well-sealing failure that results in evaporation of the reagents from a single well. When the duplicate is not affected data is still generated from that sample. Using this microplate and encoded bead approach a single laboratory technician can assay, for example, 46 samples and 2 references at a time, all in duplicate, labeling on a first day, hybridizing overnight, and washing & reading on the second day. The assay can alternately be run without replicates or with more than two replicates. Shown are duplicates of two references, 81 and 82, and duplicates of samples, and example of which is indicated at 83.

[00220] Figure 5 is an example of data generated using a Coriell DNA sample having a trisomy on chromosome 13, sex male;

[00221] This data is calculated from the median fluorescence values for each bead region produced by the Luminex reader. The average values of the negative control beads 29, 54, and 56 are subtracted from all other signals (see Fig 9). The signals from nine autosomal clones are then ratioed with the corresponding clone signals from the male and female reference DNAs. A normalization factor is calculated such that when the factor is applied to all of the autosomal clone signals it drove the average autosomal ratio to a value of one. This normalization factor is then applied to all of the signals for the sample.

[00222] The resulting ratios are plotted and shown in Figure 5. Note that the ratios for the chromosome 13 clones are all in the range of 1.3 to 1.6, while the clones for chromosomes 18 and 21, as well as the other autosomal clones are but one all below 1.2. The trisomy in chromosome 13 is readily apparent. Also, the ratio plot of the sample compared to male reference (square data points) is effectively flat across the X and Y sex chromosome. This is the response expected from a male sample. The plot of the sample compared to female reference (diamond data points) shifts down for X and up for Y, also as expected for a male sample.

[00223] Figure 6 is an example of data generated using a Coriell DNA sample having a trisomy on chromosome 18, sex male. The data is generated and plotted as described for Figure 5.

[00224] Figure 7 is an example of data generated using a Coriell DNA sample having a trisomy on chromosome 21, sex female. The data is generated and plotted as described for Figure 5.

[00225] Figure 8 is an example of data generated using a Coriell DNA sample having a 5-copy amplification of the X chromosome. The data is generated and plotted as described for Figure 5.

[00226] Figure 9 is a table displaying the BAC clones having human genomic DNA inserts used to generate amplicons in the example assays, their chromosome and cytoband locations, the sequence of the negative control oligonucleotides, and the bead ID (Luminex bead region) for the bead set to which each amplicon probe is immobilized. Sequentially numbered plotted points on the x-axis in Figures 5-8 are associated with BACs listed top-to-bottom in Figure 9. BAC RP11-186J16 is immobilized to two different bead regions (42 and 86).

[00227] For a negative control, an oligonucleotide that has no sequence homology to the human genome is selected. Specific negative control oligonucleotides used are

5' GTCACATGCGATGGATCGAGCTC 3' SEQ ID No. 5
5' CTTTATCATCGTTCCCACCTTAAT 3' SEQ ID No. 6
5' GCACGGACGAGGCCGGTATGTT 3' SEQ ID No. 7

[00228] The signals generated by the three bead regions 29, 54, and 56 having attached negative control oligonucleotides are averaged and subtracted from all other bead signals prior to calculating ratios.

[00229] Example 4

[00230] Figure 10A is a schematic flowchart illustrating a process for making a composite probe according to one aspect described herein. Probe DNA, 92, from one source and probe DNA, 93, from a second source are optionally amplified by PCR separately, 94 and 95, to produce amplicon probes which are then mixed to form a composite probe, 96. The composite probe is attached, 97, to a substrate to form a substrate-attached composite probe, 98.

[00231] Figure 10B is a schematic flowchart illustrating a process for making a composite probe according to one aspect described herein in which the probe DNA, 92 and 93, can be pooled to form a composite mixture 99, prior to optional PCR amplification 100, to produce the composite probe material, 96, which is attached is attached, 97, to a substrate to form a substrate-attached composite probe, 98.

[00232] Figure 11 is a schematic flowchart showing a process for making composite probes according to one aspect of a process described herein. An ideogram, 101, showing the cytobands, 102, of the chromosome of interest (chromosome 22 in this case) is shown. A set of five BACs, 103, with genome loci in the region of interest for the assay is shown with each
5 BAC's genome locus approximately placed on the ideogram. In this case, DNA from five BACs mapping to cytoband 22p11.2 corresponding to the DiGeorge microdeletion syndrome were used to make the composite probe. The ideogram in this figure is schematic to show the genomic proximity of the five example BACs selected from one cytoband; the BAC DNA is not extracted from a human chromosome in this process.

10 [00233] DNA was extracted and purified from each of the five cultured BACs utilizing conventional protocols. The cultured bacterial cells were lysed, the DNA was precipitated and then purified, 104, using centrifugation and column purification (Qiagen, Valencia CA). The purified DNA from each BAC, 105, was then used as the template for degenerate oligonucleotide primer (DOP) PCR amplification, 106. This was in turn followed by specific
15 PCR amplification of the DOP product using the DOP primer sequence as the specific PCR primer. This process produced five individual amplicon probes, 107. These individual probes, 107, were next pooled together, 108, to produce a composite probe, 109. The composite probe was then immobilized, 110, to a set of Luminex encoded multiplex microspheres, all of one bead "region"; i.e. having the same bead encoding identification, for use as an 22p11.2 cytoband
20 probe in a multiplex genomic gain-loss assay. The individual probes, 107, were each also immobilized individually, each to a bead set with a unique identifier so that their responses could be compared to that of the composite probe.

[00234] Figure 12 is a data plot from a test assay demonstrating the use of a composite probe with a DiGeorge syndrome reference DNA sample (Coriell Institute for Medical Research,
25 Camden NJ). The assay was performed on the Luminex xMAP platform using immobilized PCR-product probes on the Luminex encoded microspheres. The PCR product probes were made using DOP PCR from BAC DNA. The probes were immobilized, each probe on a microsphere set separately identifiable by the Luminex system, and a test assay run as described herein. The multiplex probe panel included eight autosomal probes from genome loci not
30 expected to show a gain or loss between the DNA sample and the male and female DNA references. It also included six X chromosome probes and five Y chromosome probes as positive controls, so that when the test sample is compared to a reference of the opposite sex the ratio response of a known gain or loss in the sex chromosomes can be observed. The panel also included five 22q11.2 probes, 123, see Table I below, at the locus of the DiGeorge syndrome

deletion. The center loci of these five BACs span about 0.45 megabases (445 kilobases) and the total span accounting for their 175 kb typical length is a little over 600 kilobases.

Table I. BACs and the Chr 22 linear mapping locations of their centers (megabases)

BACS	pter
F5	17.7765040
M51	17.8000000
RP11-16C10	17.9469780
RP11-316L10	18.1740250
RP11-186O8	18.2284550
	qter

[00235] Referring to Figure 12, there are two data series in the ratio plot of the DiGeorge syndrome sample compared to both male and female normal reference DNA. Data series 120 is the ratio response of the DiGeorge sample compared to male reference DNA; it shows a relative gain at all of the X chromosome probes, 125, and a loss at all of the Y chromosome probes, 126. This response shows that the sample is from a female. Data series 121 is the same sample compared to female reference DNA, and it shows a ratio response near 1.0 across the X and Y sex chromosome probes, confirming that the sample is from a female. The 1.0 ratio line, 122, in the plot indicates the expected result of sample / reference ratio for any given probe where the sample has no genomic gains or losses compared to the normal references. The autosomal probes, 127, were put into this panel as controls expected to produce a ratio of approximately 1.0, which they did.

[00236] Five probes, 123, from the 22q.11.2 locus were included in the panel. These five probes all show a ratio < 1 compared to both male and female reference DNA, consistent with the known genomic deletion in the sample at that locus. Finally, one multiplex bead type was coupled to a composite probe comprising a mixture or pool of the five 22q11.2 probes, 123. The composite probe ratio response, 124, also indicating a deletion, is reasonably concordant with the average of the response of the five constituent probes that were pooled.

[00237] Example 5

[00238] Figure 13 is ratio data from a Luminex bead array gain – loss assay according to one aspect of the present invention. This data is from fetal DNA extracted from a prenatal amniotic fluid sample for which the fetal sex was not definitively known. The assay was run with both male and female references assayed simultaneously in different wells of the same 96-well microplate. The legend, 136, identifies the two data displays, referenced to female (diamond plot points) and referenced to male (square plot points). The horizontal ratio = 1.0 line, 132, is centered in the plot. The ratio scale, 133, is the vertical axis of the graph. A first data plot, 130, is shown with the ratio of sample/reference generated against a female reference. For this plot

the data for X chromosome probes, 134, shows a ratio < 1 and the data for the Y chromosome probes, 135, shows a gain compared to female. This is consistent with a male sample. The second data plot, 131, utilizing the male reference clusters closely to the ratio = 1.0 line, also consistent with a male sample. Both data sets are show no significant deflection for the other 5 probes in the array to the left of the X probes, indicating a normal sample.

[00239] Table II shows BAC identity associated with sequentially numbered plotted points on the x-axis in Figures 13 and 14.

Table II.

	CytoBand Location	Clone ID		CytoBand Location	Clone ID
1	13q12.3-13q14.13	RP11-117I13	29	Auto 17p11.2	RP11-416I2
2	13q12.3-13q14.3	RP-11-186J16	30	Auto 1q25.2-1q31.1	RP11-46A10
3	13q12.3-13q14.3	RP-11-186J16	31	Auto 7q11.22	RP11-35P20
4	13q13.1-13q14.3	RP11-480G1	32	Auto 8p23.1	RP11-122N11
5	13q14.11-13q14.3	RP11-189B4	33	Auto 22q11.21	RP11-319F4
6	13q14.2	RP-11-174I10	34	Xp11.1-Xp11.23	RP11-465E19
7	13q14.3-13q21.31	RP11-142D16	35	Xp11.21	RP-11-292J24
8	13q21.1-13q21.33	RP-11-138D23	36	Xp11.23	RP11-38023
9	18p11.21	RP11-411B10	37	Xp11.3-Xp11.4	RP-11-258I23
10	18p11.31	RP11-55N14	38	Xp11.4-Xp21	RP11-495K15
11	18p11.32	RP11-78H1	39	Xp22.22	RP11-185L21
12	18q12.1	RP-11-63N12	40	Xp22.31	RP11-79B3
13	18q12.1	RP-11-63N12	41	Xp22.31	RP11-483M24
14	18q21.2	RP-11-160B24	42	Xp22.31	RP11-589J20
15	18q22-18q22	RP-11-88B2	43	Xp27.3	RP-11-963J21
16	18q23	RP11-89N1	44	Xq11-Xq11	RP-11-90N17
17	21q21.3	RP11-108H5	45	Xq12-Xq12	RP3-368A4
18	21q21.3-21q21.3	RP-11-147H1	46	Yp11.2	RP-11-375P13
19	21q22.12	RP11-17020	47	Yp11.31	RP11-400O10
20	21q22.12	RP11-17020	48	Yp11.31	RP-11-112L19
21	21q22.1-21q22.1	RP-11-79A12	49	Yq11.22	RP-11-20H21
22	21q22.3	GS-63-H24	50	Yq11.221	RP-11-71M14
23	21q22.3	RP11-190A24	51	Yq11.222	RP11-392F24
24	21q22.3	RP11-88N2	52	Yq11.223	RP11-336F2
25	Auto 10q26.3	RP11-462G8	53	Yq11.23	RP11-26D12
26	Auto 11p13	RP11-698N11	54	Yq11.23	RP11-79J10
27	Auto 12p13.33	RP11-598F7	55	Yq11.23	RP-11-214M24
28	Auto 16p13.3	RP11-568F1			

10 **[00240]** Figure 14 is ratio data from the same Luminex bead array assay utilizing a sample (Coriell Institute of Medical Research, Trenton NJ) with previously characterized genomic aberrations: trisomy 18 and XXX. Again, two data plots are displayed simultaneously for the sample referenced to female, 145, and referenced to male, 146. The probes for trisomy 18, all produce ratio data showing a gain, 142, compared to both references. The data for the X probes 15 shows a gain, 144, on the female-referenced plot, 145. This gain is of the same magnitude as the

trisomy gain, 142, which is consistent with XXX (sample) ratioed to XX (female reference). The male-referenced plot, 146, shows a much larger gain, 143, on the X probes as would be expected with XXX (sample) ratioed to X (male reference). The probes for the Y chromosome are noisy, as is common in aCGH, but the remaining probes are all clustered closely around the
5 ratio = 1.0, line 141. The ratio scale, 140, is the vertical axis of the graph.

[00241] It is apparent from these examples that the sex of a sample with a normal complement of X and Y chromosomes is immediately apparent from the ratio data generated against both male and female references. It is also apparent in the case of a multi-X aberrant sample the quantitation of the multiple copies of X is more straightforward when using both
10 references.

[00242] Any patents or publications mentioned in this specification are incorporated herein by reference to the same extent as if each individual publication is specifically and individually indicated to be incorporated by reference. U.S. Patent Application Serial No. 11/615,739, filed December 22, 2006; U.S. Patent Application Serial No. 12/055,919, filed March 26, 2008; and
15 U.S. Provisional Applications Serial Nos. 60/753,584, filed December 23, 2005, 60/753,822, filed December 23, 2005, 60/765,311, filed February 3, 2006, 60/765,355, filed February 3, 2006, and 60/992,489, filed December 5, 2007, are all incorporated herein by reference in their entirety.

[00243] The compositions and methods described herein are presently representative of
20 certain embodiments, exemplary, and not intended as limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art. Such changes and other uses can be made without departing from the scope of the invention as set forth in the claims.

CLAIMS

1. A method of assaying a DNA sample, comprising:

providing a substrate-attached composite nucleic acid probe, the composite nucleic acid probe comprising nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome, the genomic region characterized by a first terminus and a second terminus and having an intermediate region disposed between the first terminus and second terminus of at least 400 kilobases, wherein the composite nucleic acid probe comprises nucleic acid sequences which specifically hybridize to substantially an entire first genomic locus comprising the first terminus and to substantially an entire second genomic locus comprising the second terminus;

hybridizing the substrate-attached composite nucleic acid probe with sample genomic DNA;

hybridizing the substrate-attached composite nucleic acid probe with reference genomic DNA;

detecting a first signal indicating specific hybridization of the substrate-attached composite nucleic acid probe with the sample genomic DNA and a second signal indicating specific hybridization of the substrate-attached composite nucleic acid probe with the reference genomic DNA; and

comparing the first signal and the second signal to detect differences between the first and second signals, the differences of the first and second signals indicative of differences between the sample DNA and the reference DNA, thereby assaying the DNA sample.

2. The method of claim 1, wherein the substrate is a plurality of particles.

3. The method of claim 1, wherein the substrate is a plurality of encoded particles.

4. The method of claim 1, wherein the substrate is a planar substrate.

5. The method of claim 1, wherein the first genomic locus and second genomic locus each comprise at least about 100 kilobases.

6. The method of claim 1, wherein the nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome are derived from two or more large-insert DNA vectors.

5 7. The method of claim 6, wherein the two or more large-insert DNA vectors are selected from the group consisting of: bacterial artificial chromosomes, yeast artificial chromosomes, human artificial chromosomes, cosmids, plasmids, phagemids, phage DNA and fosmids.

10 8. The method of claim 1, wherein the nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome are derived from a source selected from: isolated chromosomes and isolated chromosome fragments.

9. The method of claim 1, wherein the nucleic acid sequences which specifically 15 hybridize to two or more genomic loci in a genomic region of a reference genome are amplicons derived from a source selected from: two or more large-insert DNA vectors, isolated chromosomes, isolated chromosome fragments, a large-insert DNA vector and isolated chromosomes, and a large-insert DNA vector and isolated chromosome fragments.

20 10. The method of claim 1 wherein the sample genomic DNA is detectably labeled.

11. The method of claim 1 wherein the reference genomic DNA is detectably labeled.

12. The method of claim 1, wherein the nucleic acid sequences which specifically 25 hybridize to two or more genomic loci in a genomic region of a reference genome individually have a length in the range of about 20 – 250,000 nucleotides, inclusive.

13. The method of claim 1, wherein the sample DNA is DNA obtained from an individual subject.

30

14. The method of claim 13, wherein the sample DNA is genomic DNA obtained from an individual subject.

15. The method of claim 1, wherein the sample DNA is human DNA.

16. The method of claim 1, further comprising:

hybridizing the substrate-attached composite nucleic acid probe with a second reference genomic DNA;

5 detecting a third signal indicating specific hybridization of the substrate-attached composite nucleic acid probe with the second reference genomic DNA; and

comparing the first signal and the third signal to detect differences between the first and third signals, the differences of the first and third signals indicative of differences between the sample DNA and the second reference DNA, thereby assaying the DNA sample.

10

17. A method of assaying sample DNA, comprising:

providing a multiplex reagent comprising a mixture of two or more encoded particle sets encoded such that each particle of each encoded particle set is detectably distinguishable from each particle of each other encoded particle set, the encoded particles comprising attached 15 nucleic acid sequences which specifically hybridize to at least one genomic locus of a reference genome, wherein at least one encoded particle set comprises an attached composite nucleic acid probe, the composite nucleic acid probe comprising nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome, the genomic region characterized by a first terminus and a second terminus and having an intermediate region 20 disposed between the first terminus and second terminus of at least 400 kilobases, wherein the composite nucleic acid probe comprises nucleic acid sequences which specifically hybridize to substantially an entire first genomic locus comprising the first terminus and to substantially an entire second genomic locus comprising the second terminus;

hybridizing the multiplex reagent with sample genomic DNA;

25 hybridizing the multiplex reagent with reference genomic DNA;

detecting a first signal indicating specific hybridization of the attached nucleic acid sequences with detectably labeled DNA;

detecting a second signal indicating specific hybridization of the attached nucleic acid sequences with detectably labeled reference DNA;

30 identifying the encoded particles so as to associate particle encoding with the first signal;

identifying the encoded particles so as to associate particle encoding with the second signal; and

comparing the first signal and the second signal for each encoded particle set, wherein differences in the first and second signals are indicative of differences between the sample and reference DNA, thereby assaying DNA.

5 18. A reagent for assay of DNA, comprising:

a first composite nucleic acid probe attached to a solid substrate, the first composite nucleic acid probe comprising nucleic acid sequences which specifically hybridize to two or more genomic loci in a genomic region of a reference genome, the genomic region characterized by a first terminus and a second terminus and having an intermediate region disposed between 10 the first terminus and second terminus of at least 400 kilobases, wherein the first composite nucleic acid probe comprises nucleic acid sequences which specifically hybridize to substantially an entire first genomic locus comprising the first terminus and to substantially an entire second genomic locus comprising the second terminus.

15 19. The reagent for assay of DNA of claim 18, further comprising:

a second composite nucleic acid probe attached to a solid substrate, the second composite nucleic acid probe comprising nucleic acid sequences which specifically hybridize to two or more genomic loci in a second genomic region of a reference genome, the second genomic region characterized by a first terminus and a second terminus and having an intermediate region 20 disposed between the first terminus and second terminus of at least 400 kilobases, wherein the second composite nucleic acid probe comprises nucleic acid sequences which specifically hybridize to substantially an entire first genomic locus comprising the first terminus of the second genomic region and to substantially an entire second genomic locus comprising the second terminus of the second genomic region.

25

20. The reagent of claim 18 wherein the solid substrate is a planar substrate.

21. The reagent of claim 19 wherein the solid substrate is a planar substrate further comprising the first composite nucleic acid probe.

30

22. The reagent of claim 18 wherein the solid substrate is a first plurality of particles.

23. The reagent of claim 19 wherein the solid substrate is a second plurality of particles.

24. The reagent of claim 23 wherein the first plurality of particles and second plurality of particles are distinguishably encoded.

5 25. A method of preparing a substrate-attached composite nucleic acid probe reagent for assay of DNA, comprising:

isolating a first nucleic acid sequence which specifically hybridizes to substantially an entire first genomic locus comprising a first terminus of a genomic region of a reference genome;

isolating a second nucleic acid sequence which specifically hybridizes to substantially an
10 entire second genomic locus comprising a second terminus of the genomic region of the reference genome;

mixing the first and the second nucleic acid sequences to produce a composite probe;

binding the composite probe to a solid substrate, thereby producing a substrate-attached composite nucleic acid probe reagent for assay of DNA.

15

26. The method of claim 25 wherein the first and the second nucleic acid sequences comprise a functional group for reaction with the solid substrate.

27. The method of claim 25, wherein the first nucleic acid sequence is isolated from a
20 first large-insert vector and the second nucleic acid sequence is isolated from a second large-insert vector.

28. The method of claim 25, wherein the first and the second nucleic acid sequences are amplified prior to mixing.

25

29. The method of claim 25, wherein the first and the second nucleic acid sequences are amplified after mixing.

30. A method of assaying a DNA sample, comprising:

30 providing a substrate-attached nucleic acid probe;

hybridizing the substrate-attached nucleic acid probe with sample genomic DNA obtained from a subject;

hybridizing the substrate-attached nucleic acid probe with first reference genomic DNA;

hybridizing the substrate-attached nucleic acid probe with second reference genomic DNA;

detecting a first signal indicating specific hybridization of the substrate-attached nucleic acid probe with the sample genomic DNA, a second signal indicating specific hybridization of the substrate-attached nucleic acid probe with the first reference genomic DNA and a third signal indicating specific hybridization of the substrate-attached nucleic acid probe with the second reference genomic DNA;

comparing the first signal and the second signal to detect differences between the first and second signals, the differences of the first and second signals indicative of differences between the sample DNA and the first reference DNA; and

comparing the first signal and the third signal to detect differences between the first and third signals, the differences of the first and third signals indicative of differences between the sample DNA and the second reference DNA, thereby assaying the DNA sample.

31. The method of claim 30, further comprising:

comparing the differences between the sample DNA and the first reference DNA with differences between the sample DNA and the second reference DNA, thereby assaying the DNA sample.

32. The method of claim 30 wherein the first reference genomic DNA comprises male-specific genomic DNA, wherein the second reference genomic DNA comprises female-specific genomic DNA and wherein comparison of the differences of the first and second signals and comparison of the differences of the first and third signals is indicative of gender of the subject.

33. The method of claim 30 wherein the first reference genomic DNA comprises first condition specific genomic DNA, wherein the second reference genomic DNA comprises second condition specific genomic DNA and wherein comparison of the differences of the first and second signals and comparison of the differences of the first and third signals is indicative of a disease state of the subject.

34. The method of claim 30 wherein the first reference genomic DNA comprises first condition specific genomic DNA, wherein the second reference genomic DNA comprises second condition specific genomic DNA and wherein comparison of the differences of the first and

second signals and comparison of the differences of the first and third signals is indicative of metabolic age of the subject.

35. The method of claim 30 wherein the substrate-attached nucleic acid probe comprises a plurality of encoded particles.

36. The method of claim 30 wherein the substrate-attached nucleic acid probe comprises a planar substrate.

37. The method of claim 30 wherein the substrate-attached nucleic acid probe comprises a plurality of oligonucleotides.

38. The method of claim 30 wherein the substrate-attached nucleic acid probe comprises a plurality of amplicons.

39. The method of claim 30 wherein the substrate-attached nucleic acid probe comprises insert DNA isolated from a large-insert DNA vector.

40. The method of claim 30 wherein the substrate-attached nucleic acid probe comprises isolated chromosomal DNA.

41. A method of assaying a DNA sample, comprising:

providing a first encoded particle set comprising encoded particles having attached amplicons, the amplicons comprising random nucleic acid sequences together representing substantially an entire first template DNA sequence;

hybridizing the amplicons of the first encoded particle set with detectably labeled sample DNA;

hybridizing the amplicons of the first encoded particle set with detectably labeled first reference DNA;

hybridizing the amplicons of the first encoded particle set with detectably labeled second reference DNA;

detecting a first signal indicating specific hybridization of the amplicons of the first encoded particle set with detectably labeled sample DNA, a second signal indicating specific hybridization of the amplicons of the first encoded particle set with detectably labeled first

reference DNA and a third signal indicating specific hybridization of the amplicons of the first encoded particle set with detectably labeled second reference DNA;

comparing the first signal and the second signal to detect differences between the first and second signals, the differences of the first and second signals indicative of differences between the sample DNA and the reference DNA; and

comparing the first signal and the third signal to detect differences between the first and third signals, the differences of the first and third signals indicative of differences between the sample DNA and the second reference DNA, thereby assaying the DNA sample.

42. The method of claim 41, wherein the amplicons have a length in the range of about 500 – 1200 nucleotides, inclusive.

43. The method of claim 41, wherein the detectably labeled sample DNA is detectably labeled DNA obtained from an individual subject.

44. The method of claim 43, wherein the detectably labeled sample DNA is detectably labeled genomic DNA obtained from an individual subject.

45. The method of claim 41, wherein the detectably labeled sample DNA is human DNA.

46. The method of claim 41, further comprising:

providing a second encoded particle set comprising encoded particles having attached amplicons, the amplicons comprising random nucleic acid sequences together representing substantially an entire second template DNA sequence;

hybridizing the amplicons of the second encoded particle set with detectably labeled sample DNA;

hybridizing the amplicons of the second encoded particle set with detectably labeled first reference DNA;

hybridizing the amplicons of the second encoded particle set with detectably labeled second reference DNA;

detecting a first signal indicating specific hybridization of the amplicons of the second encoded particle set with detectably labeled sample DNA, a second signal indicating specific hybridization of the amplicons of the second encoded particle set with detectably labeled

reference DNA and a third signal indicating specific hybridization of the amplicons of the second encoded particle set with detectably labeled second reference DNA;

comparing the first signal indicating specific hybridization of the amplicons of the second encoded particle set and the second signal indicating specific hybridization of the amplicons of
5 the second encoded particle set to detect differences between the first and second signals, the differences of the first and second signals indicative of differences between the sample DNA and the first reference DNA; and

comparing the first signal indicating specific hybridization of the amplicons of the second encoded particle set and the third signal indicating specific hybridization of the amplicons of the
10 second encoded particle set to detect differences between the first and third signals, the differences of the first and third signals indicative of differences between the sample DNA and the second reference DNA.

47. The method of claim 46, wherein the first and second encoded particle sets are
15 provided in a mixture and further comprising:

associating encoding of the first encoded particle set with the first signal indicating specific hybridization of the amplicons of the first encoded particle set with detectably labeled sample DNA and a second signal indicating specific hybridization of the amplicons of the first encoded particle set; and

20 associating encoding of the second encoded particle set with the first signal indicating specific hybridization of the amplicons of the second encoded particle set with detectably labeled sample DNA and the second signal indicating specific hybridization of the amplicons of the second encoded particle set.

25 48. A method of DNA assay substantially as described herein.

49. A reagent for DNA assay substantially as described herein.

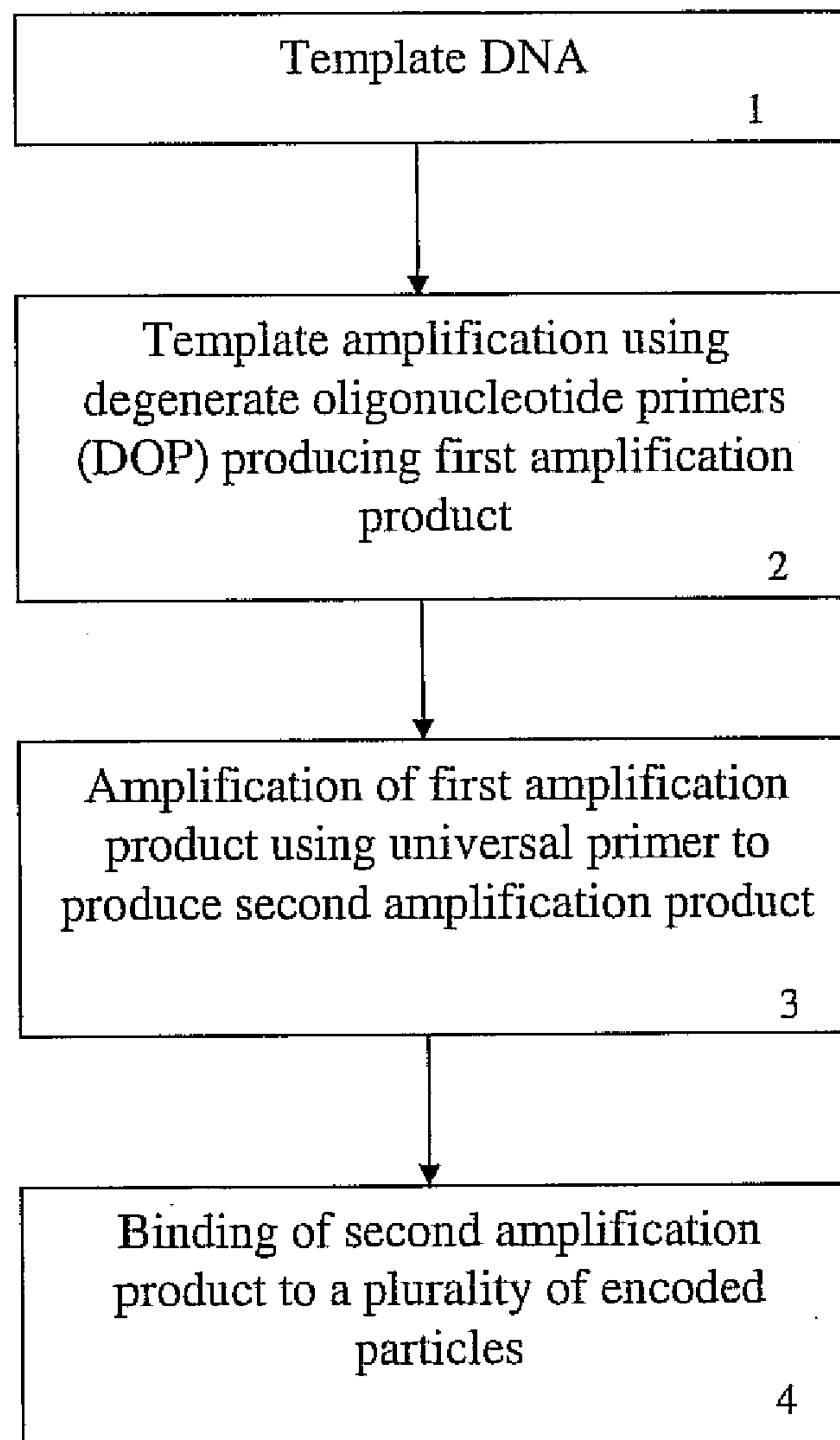

Figure 1

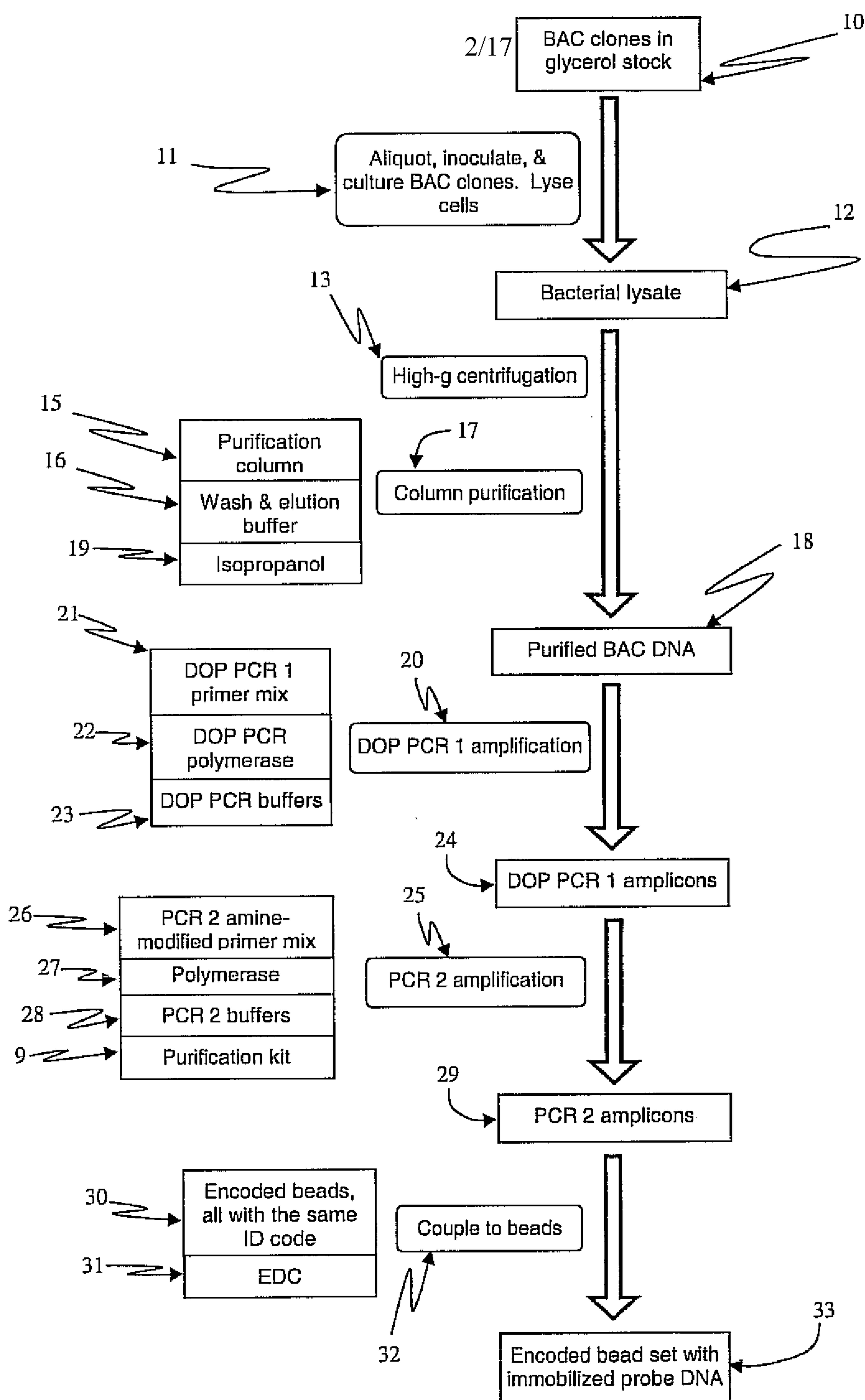

Figure 1A

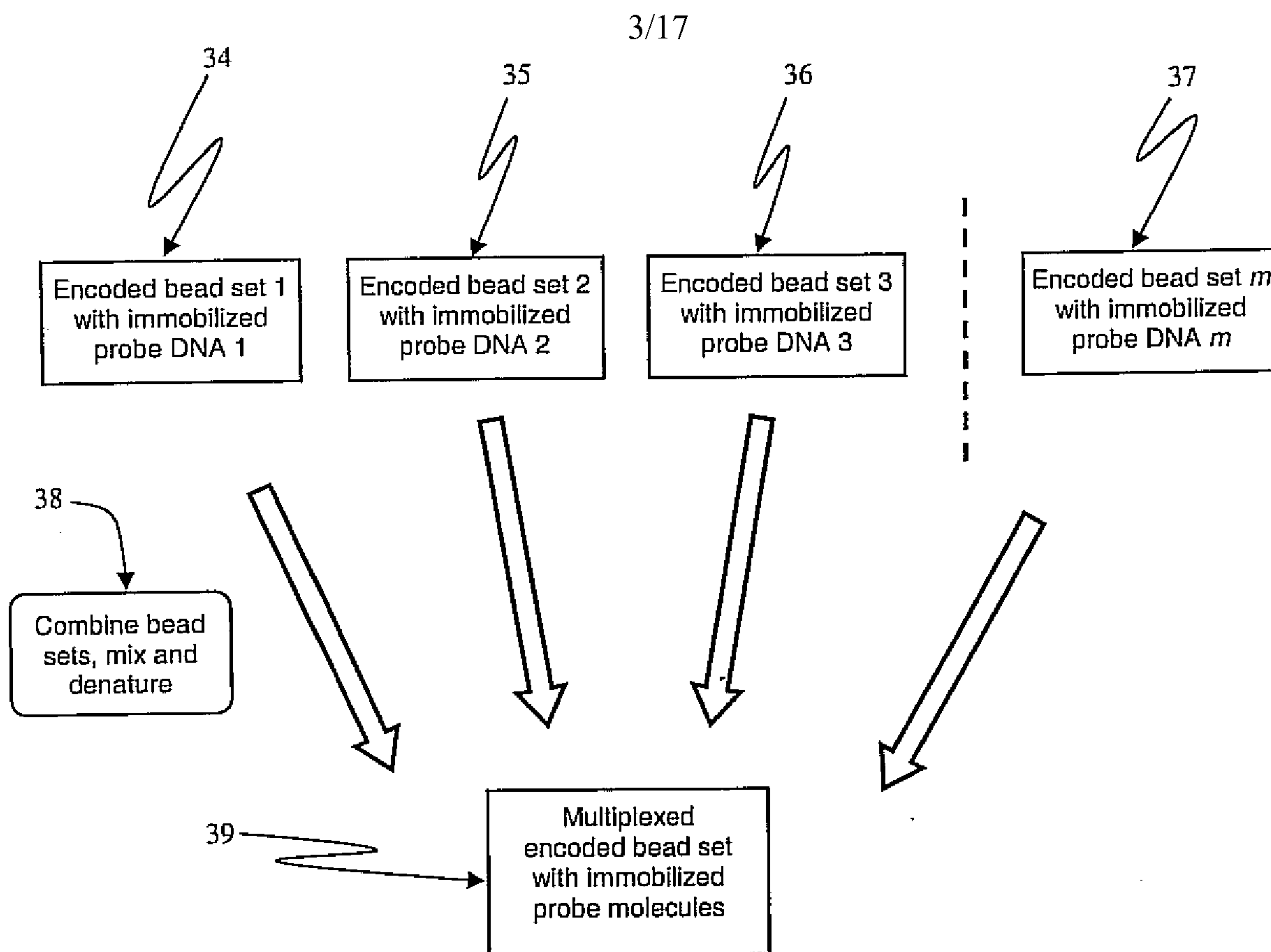

Figure 2

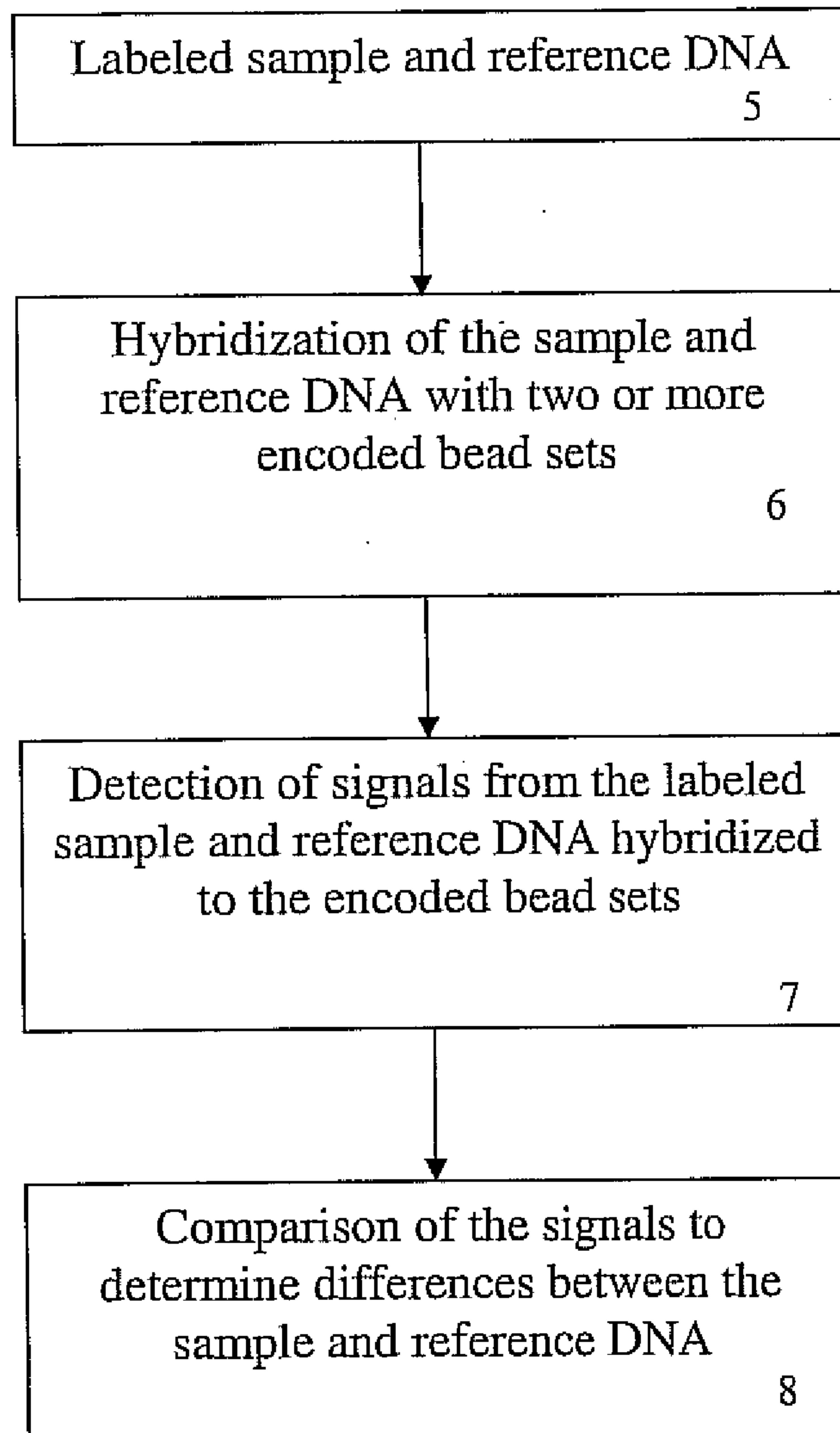

Figure 3

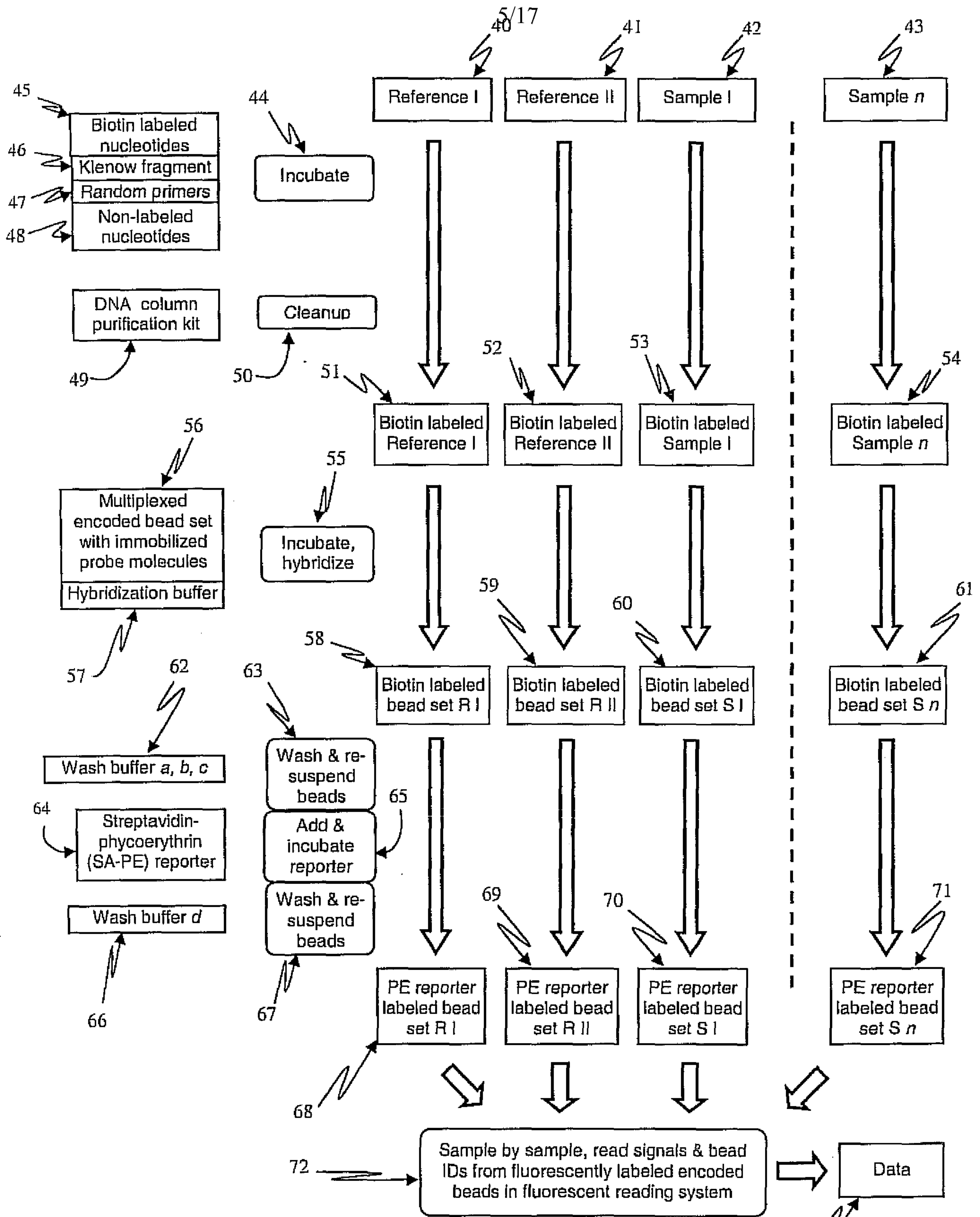

Figure 3A

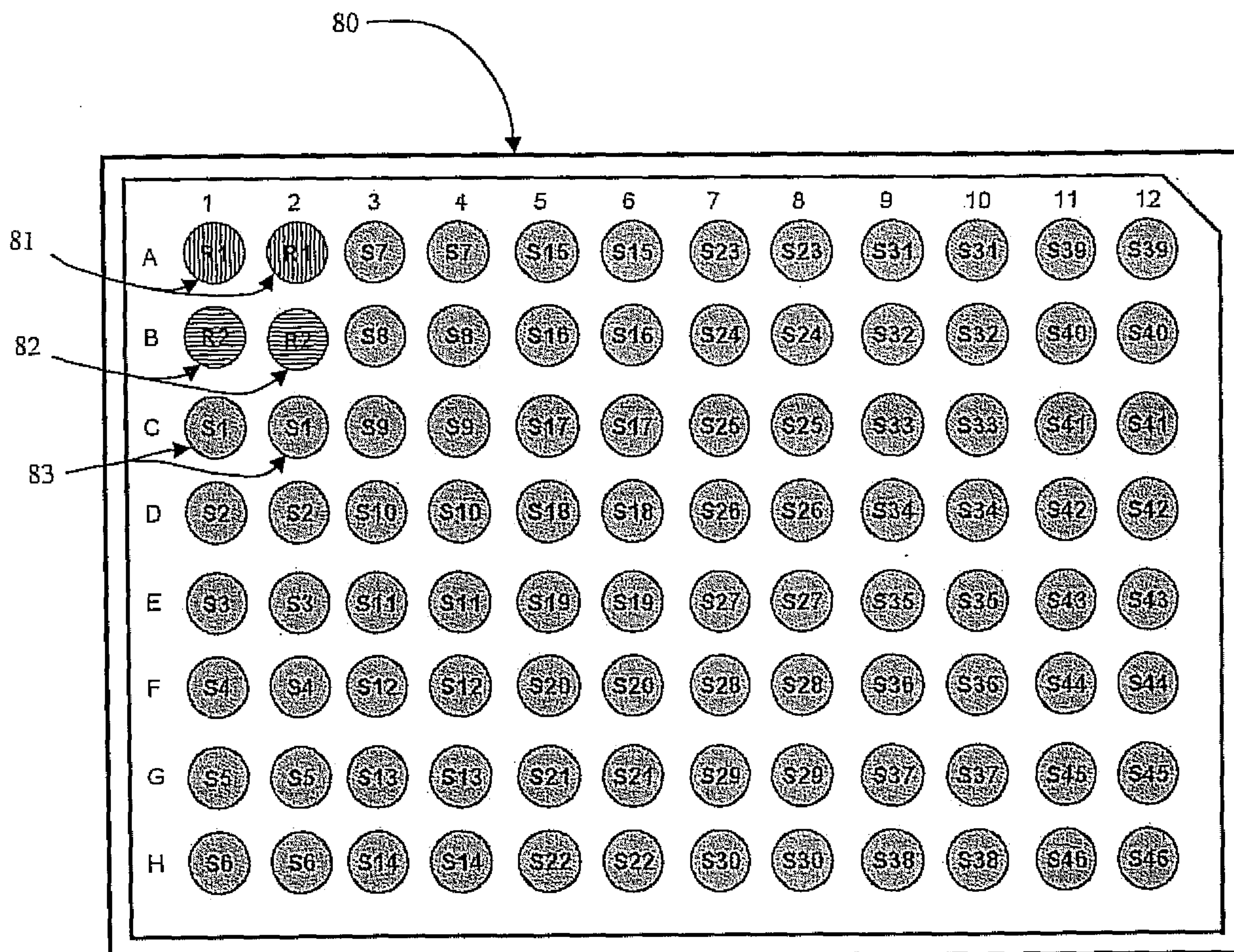

Figure 4

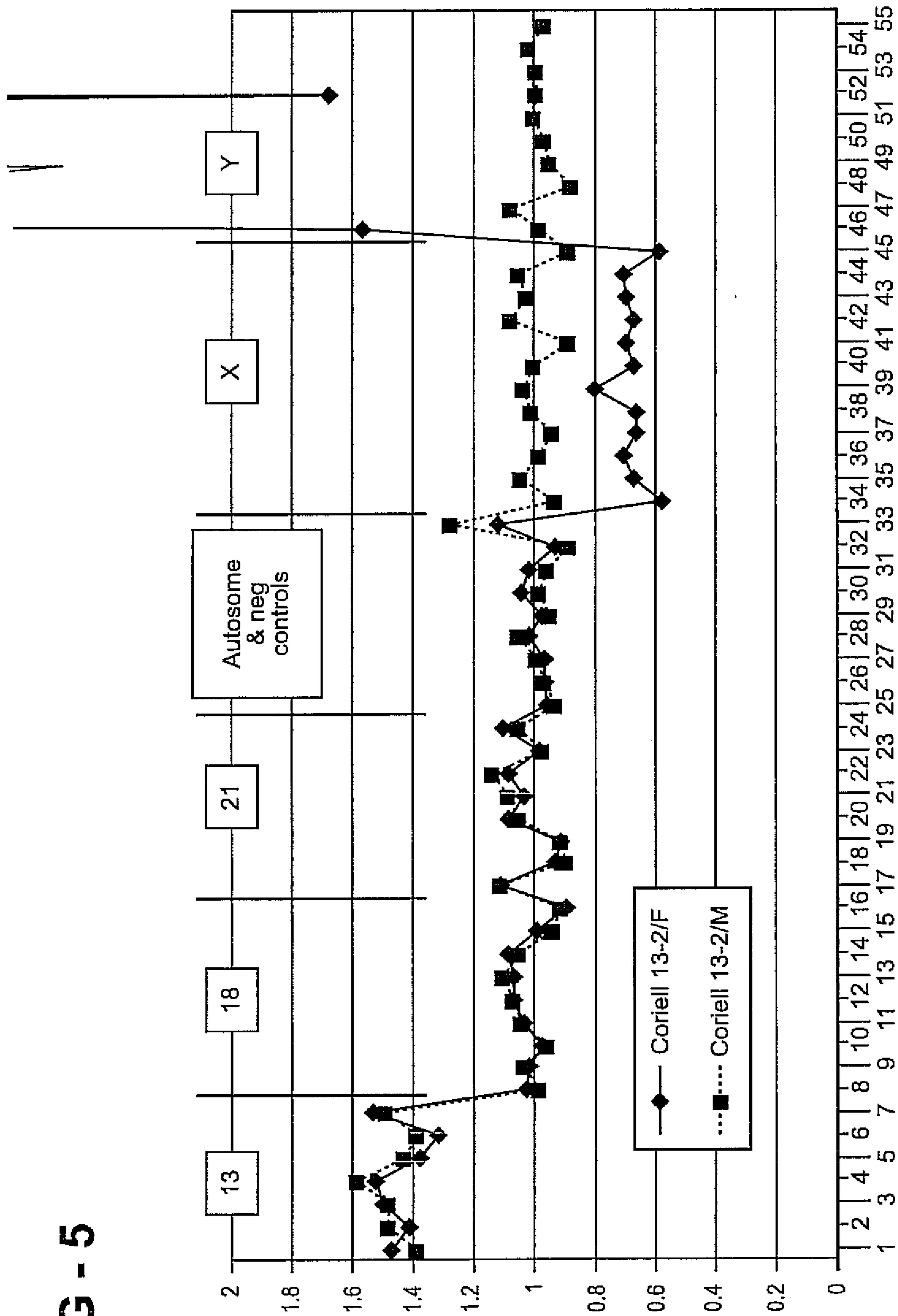

FIG - 5

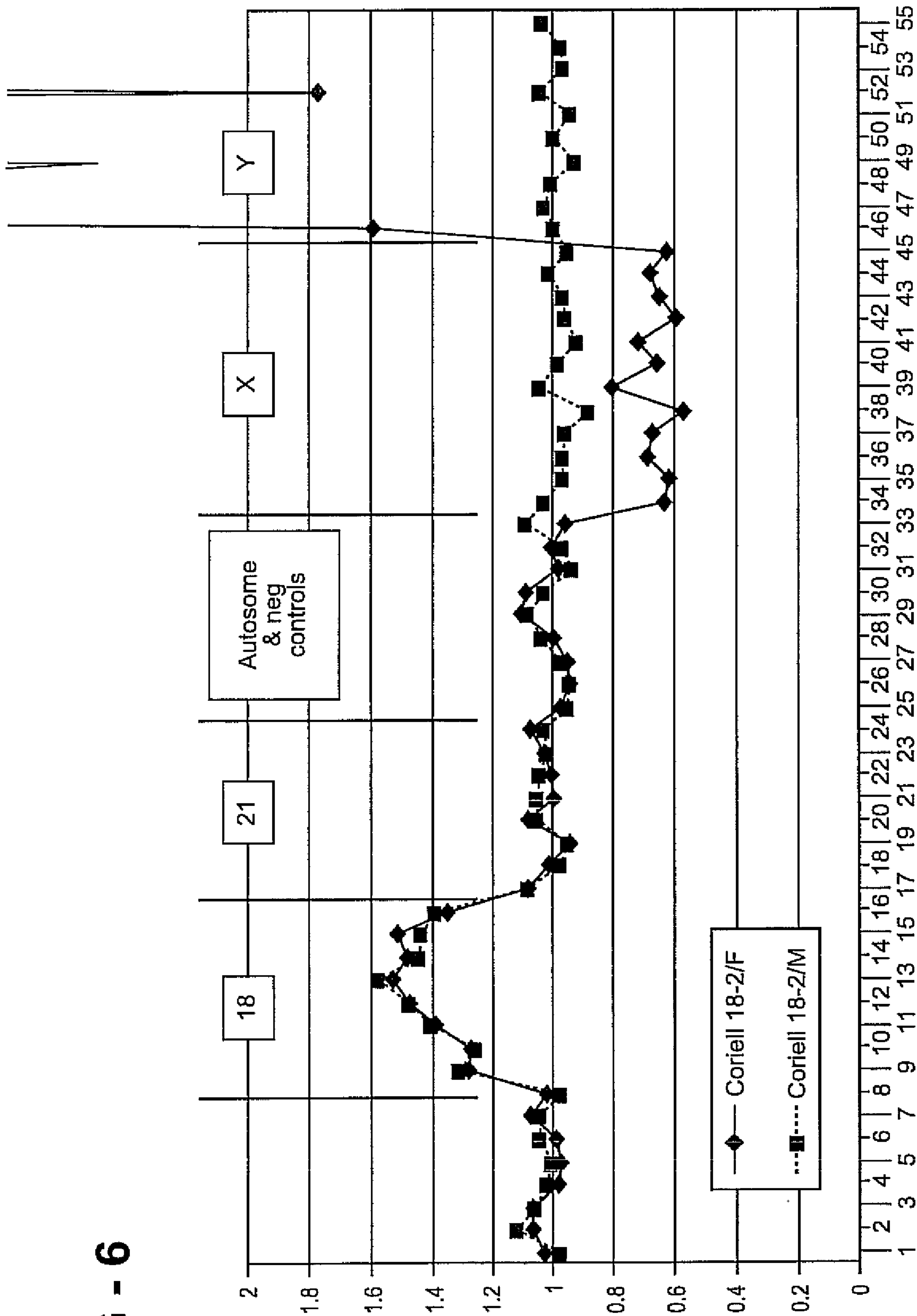
FIG - 6
18
21
Autosome & neg controls
X
Y
Coriell 18-2/F
Coriell 18-2/M
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55

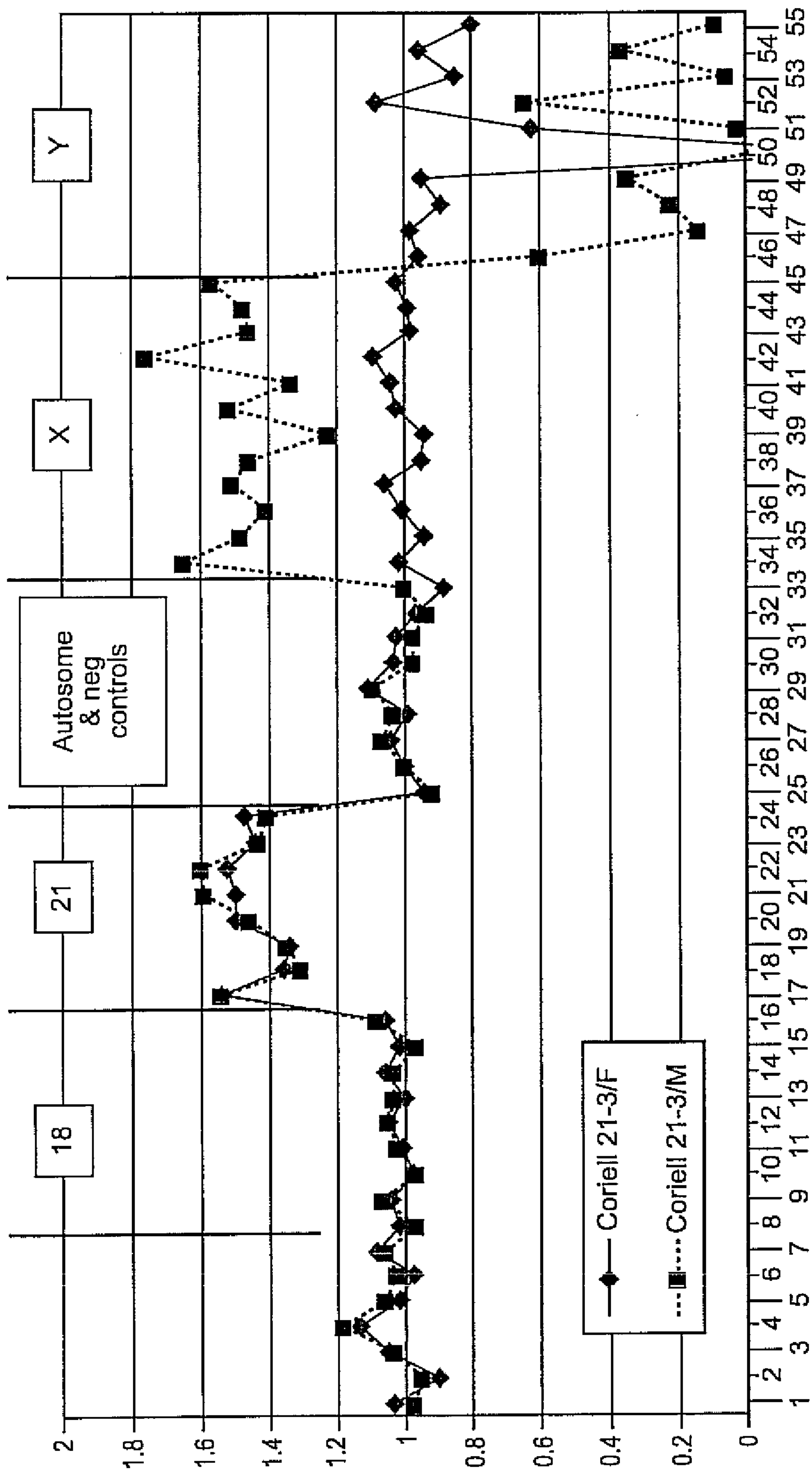

FIG - 7

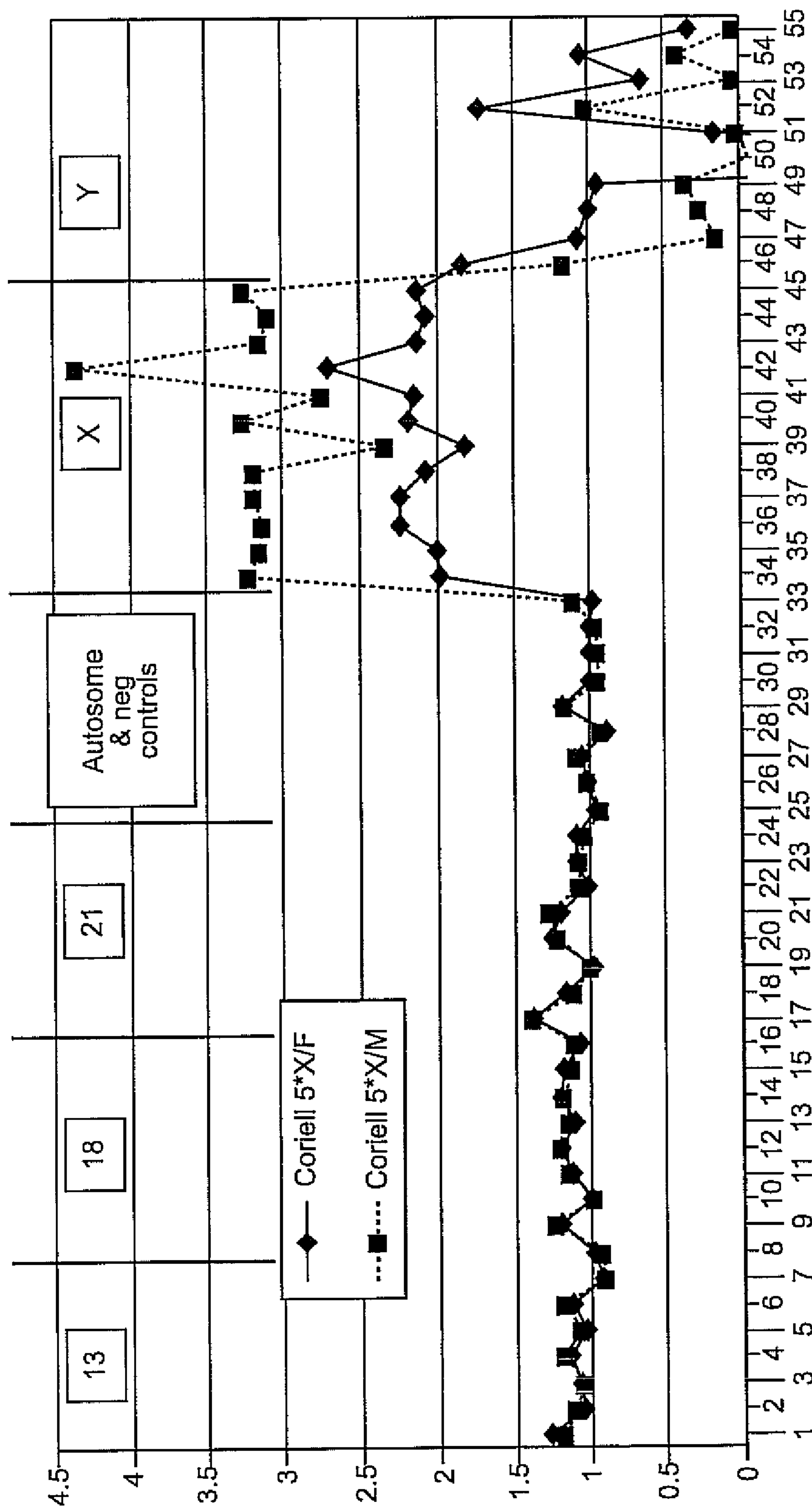

FIG - 8

BAC*	Chromosome	BAC cyto band location	Bead ID
RP11-186J16	13	13q12.3	42
RP11-186J16		13q12.3	86
RP11-117I13		13q12.3	87
RP11-480G1		13q13.1	85
RP11-189B4		13q14.11	83
RP11-174I10		13q14.2	37
RP11-142D16		13q14.3	84
RP11-138D23		13q21.1	34
RP11-411B10	18	18p11.21	88
RP11-55N14		18p11.31	57
RP11-78H1		18p11.32	90
RP11-63N12		18q12.1	45
RP11-63N12		18q12.1	89
RP11-160B24		18q21.2	36
RP11-88B2		18q22	51
RP11-89N1		18q23	48
RP11-108H5	21	21q21.3	35
RP11-147H1		21q21.3	67
RP11-17O20		21q21.3	69
RP11-17O20		21q22.1	47
RP11-79A12		21q22.12	38
RP11-190A24		21q22.12	68
RP11-88N2		21q22.3	58
GS-63-H24		21q22.3	78
RP11-46A10	Autosomal & negative controls	10q26.3)	12
RP11-35P20		11p13	13
RP11-122N11		12p13.33	14
RP11-462G8		16p13.3	6
RP11-698N11		17p11.2	7
RP11-598F7		1q25.2	32
RP11-568F1		22q11.21	8
RP11-416I2		7q11.22	41
RP11-319F4		8p23.1	11
GTCACATGCGATGGATCGAGCTC		Negative Control Oligo	29
CTTTATCATCGTTCCCACCTTAAT		Negative Control Oligo	54
GCACGGACGAGGCCGGTATGTT		Negative Control Oligo	56
RP11-38023	X	Xp11.23	72
RP11-495K15		Xp21.1-Xp21.1	73
RP11-79B3		Xp22.11	74
RP11-483M24		Xp22.31	75
RP11-589J20		Xp22.31	76
RP11-465E19		Xp11.1-Xp11.23	44
RP11-292J24		Xp11.21	43
RP11-258I23		Xp11.3-Xp11.4	63
RP1-185L21		Xp22.22	66
RP11-963J21		Xq27.3	65
RP11-90N17		Xq11-Xq11	52
RP3-368A4		Xq12-Xq12	53
RP11-400O10	Y	Yp11.31	20
RP11-336F2		Yq11.223	77
RP11-26D12		Yq11.23	17
RP11-392F24		Yq11.222	18
RP11-79J10		Yq11.23	19
RP11-375P13		Yp11.2	64
RP11-112L19		Yp11.31	33
RP11-20H21		Yq11.22	61
RP11-71M14		Yq11.221	46
RP11-214M24		Yq11.23	62

*or negative control oligo

Figure 9

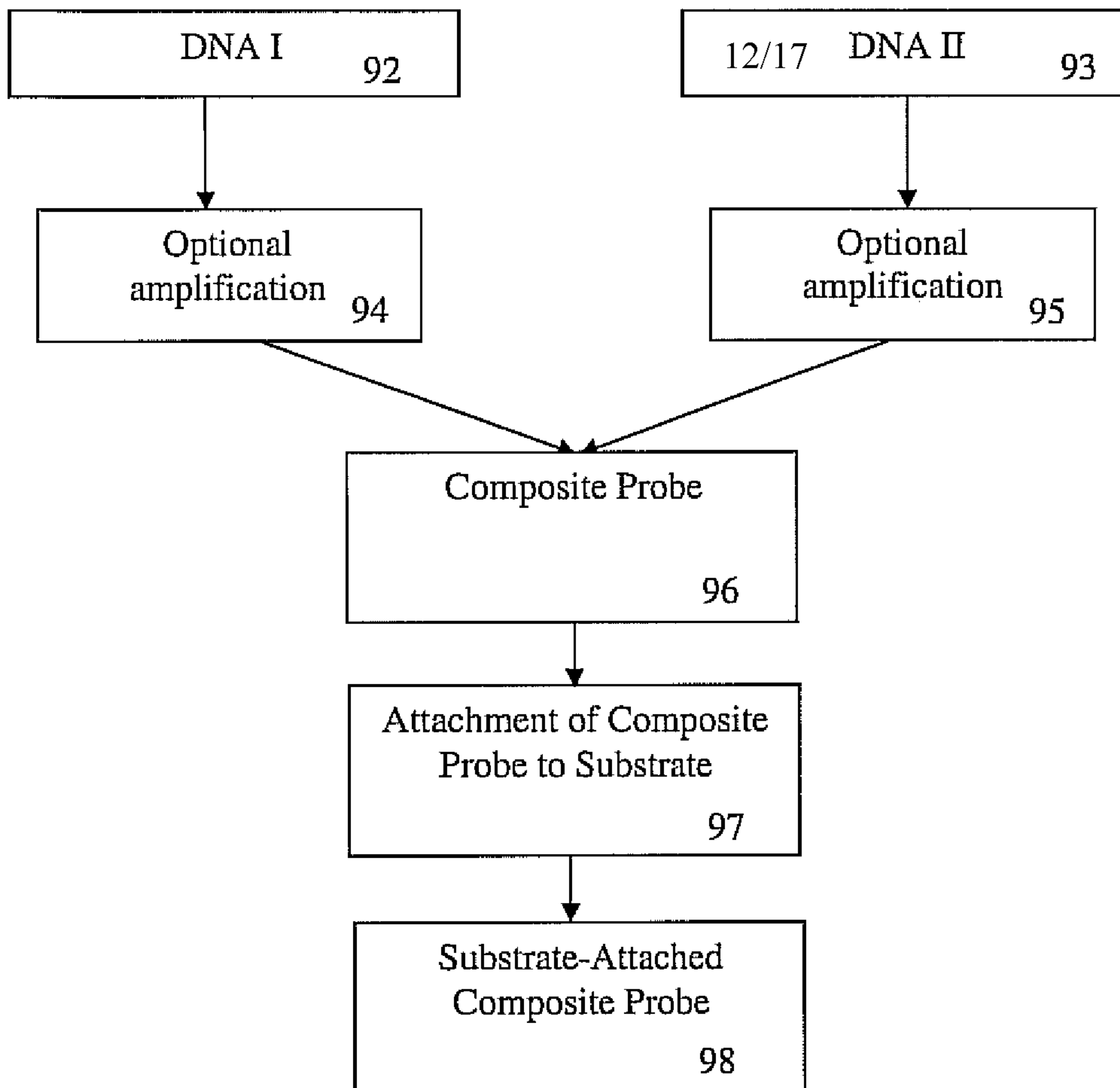

Figure 10A

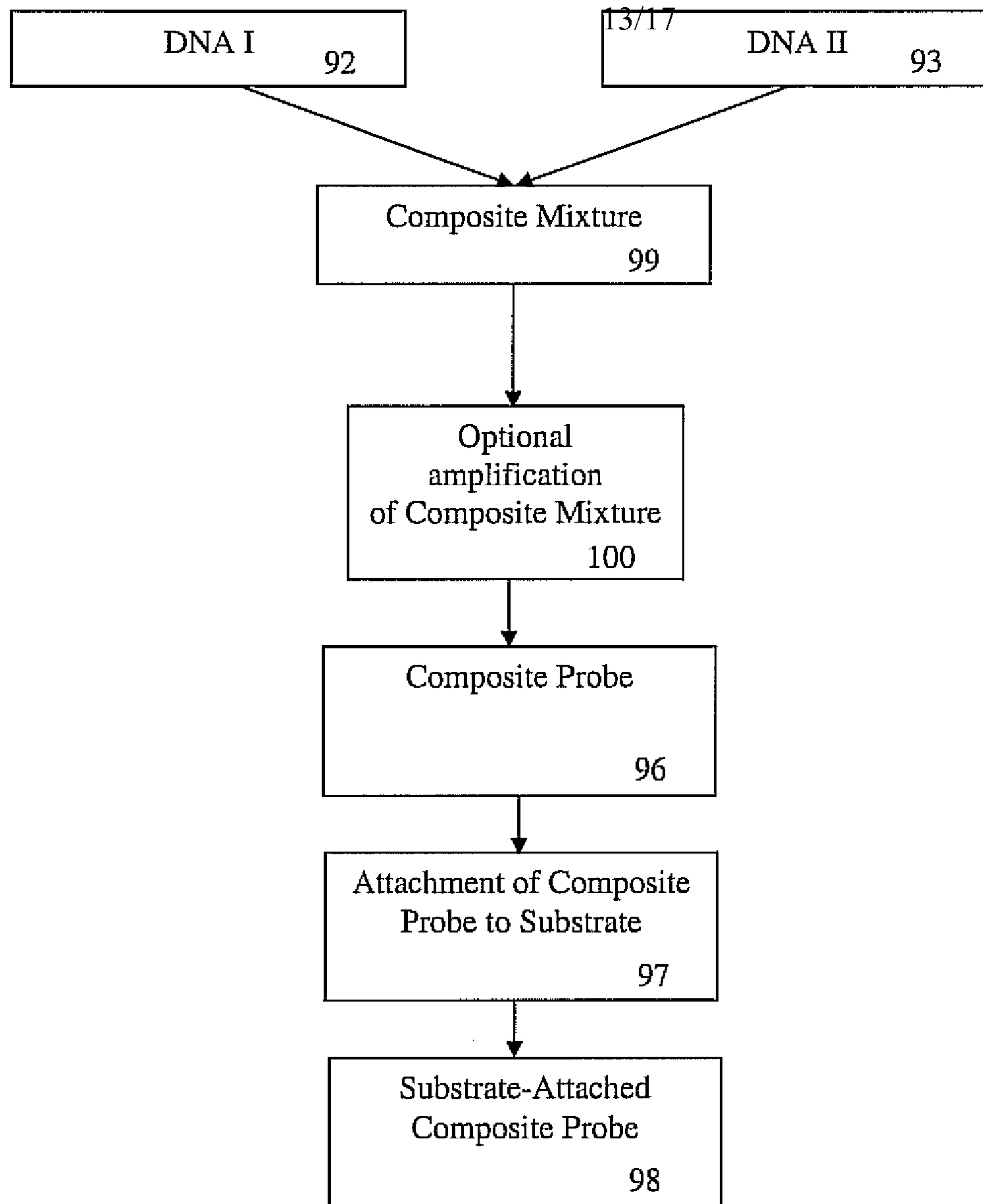

Figure 10B

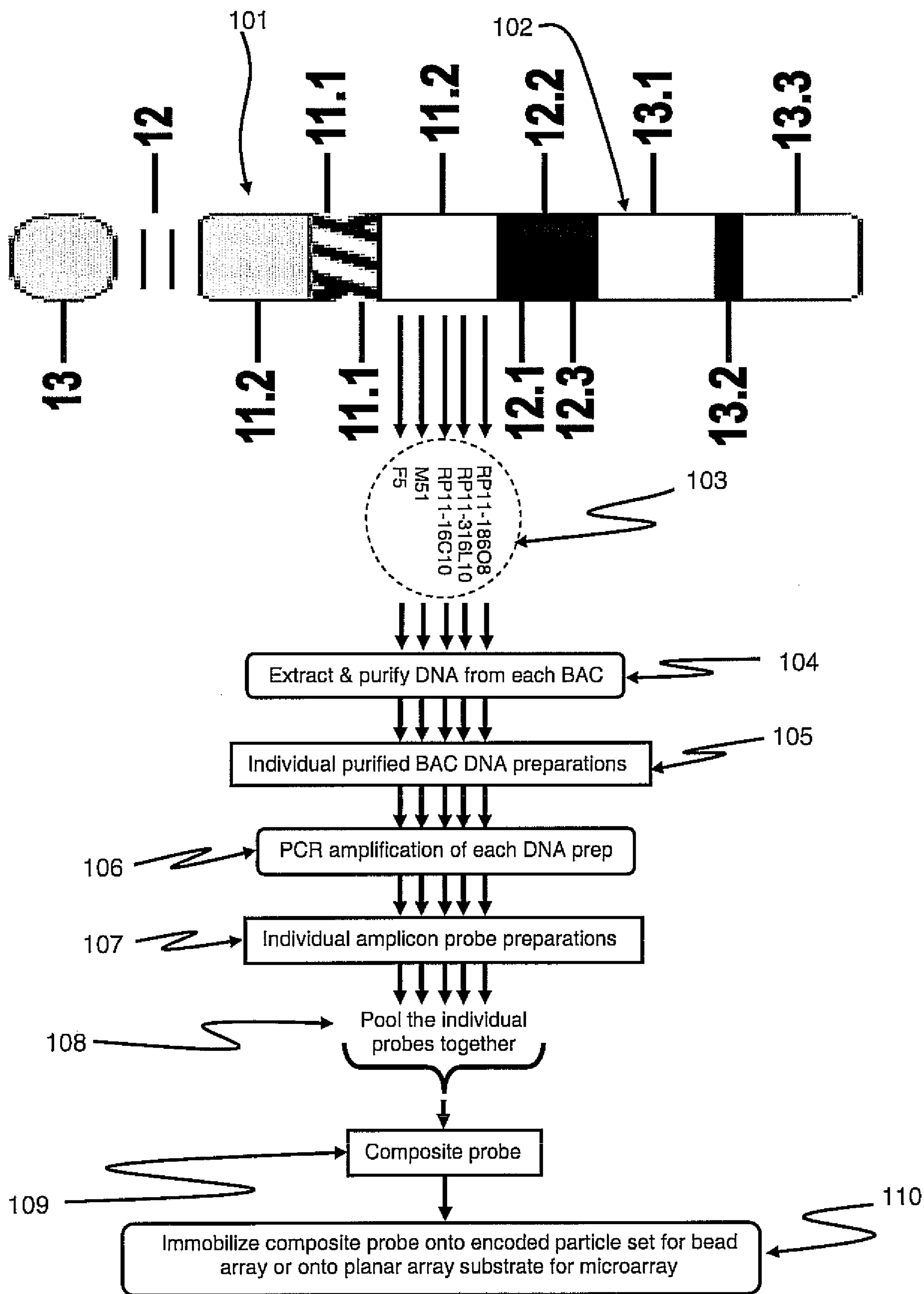

Fig. 11

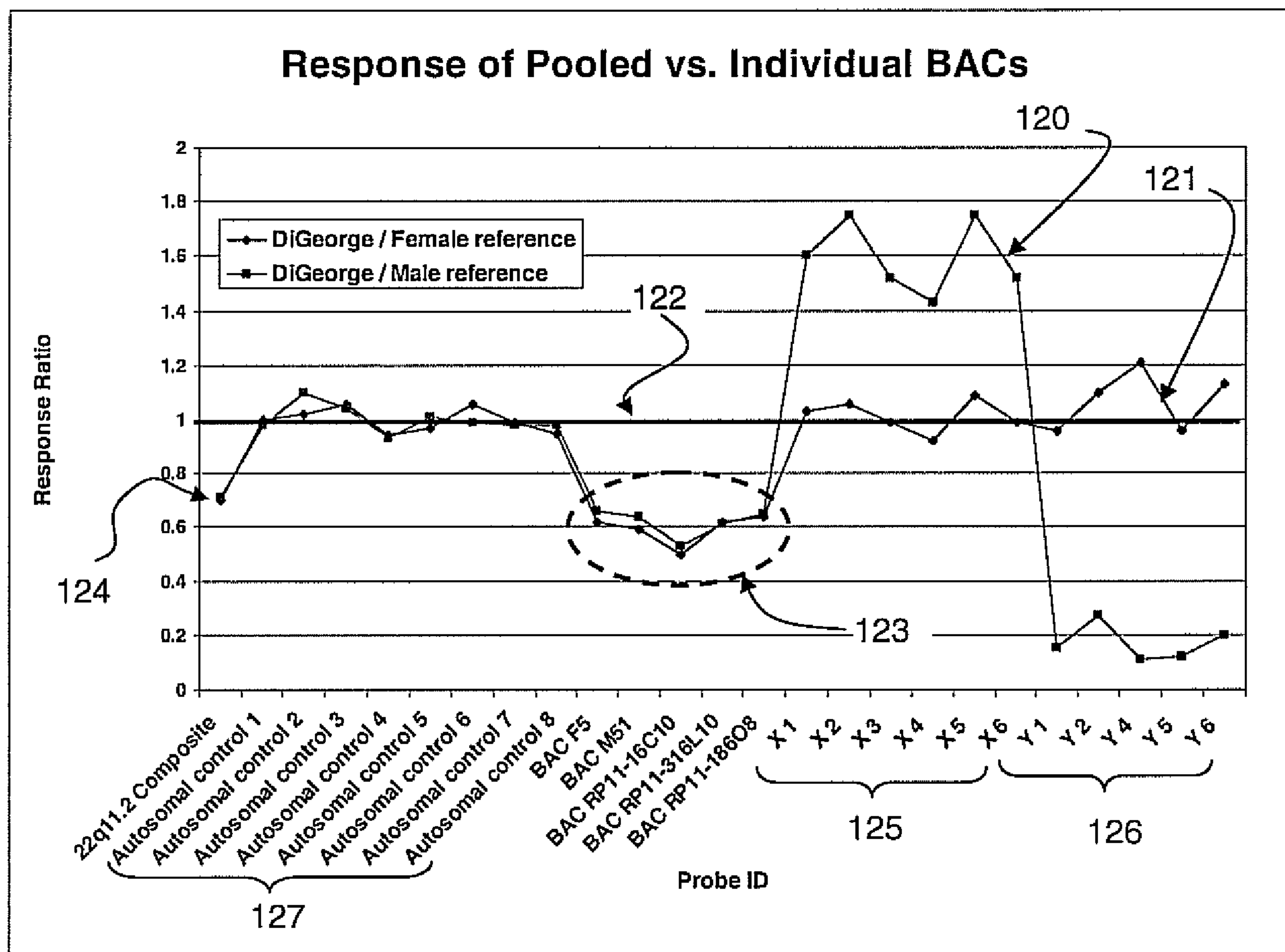

Figure 12

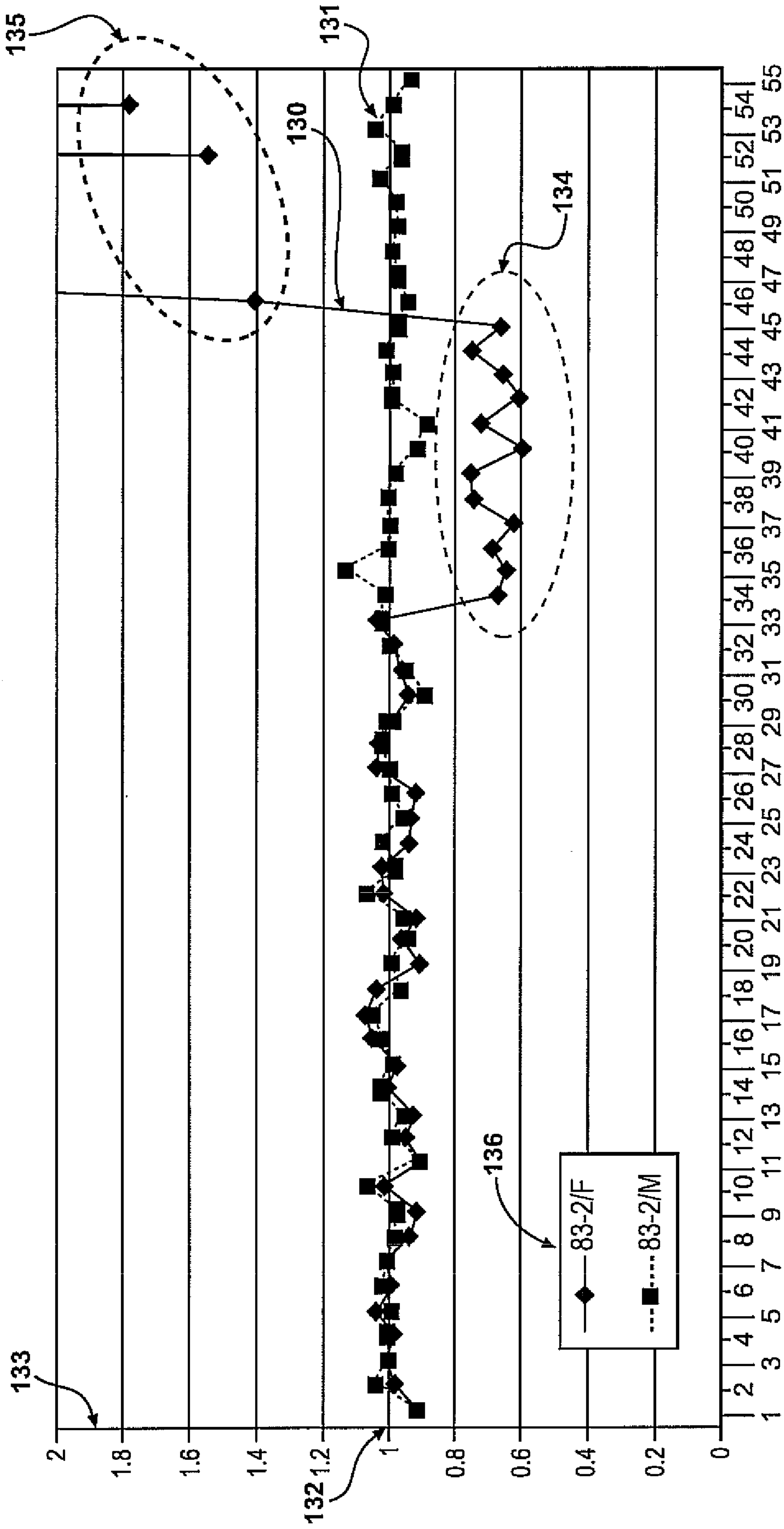

FIG - 13

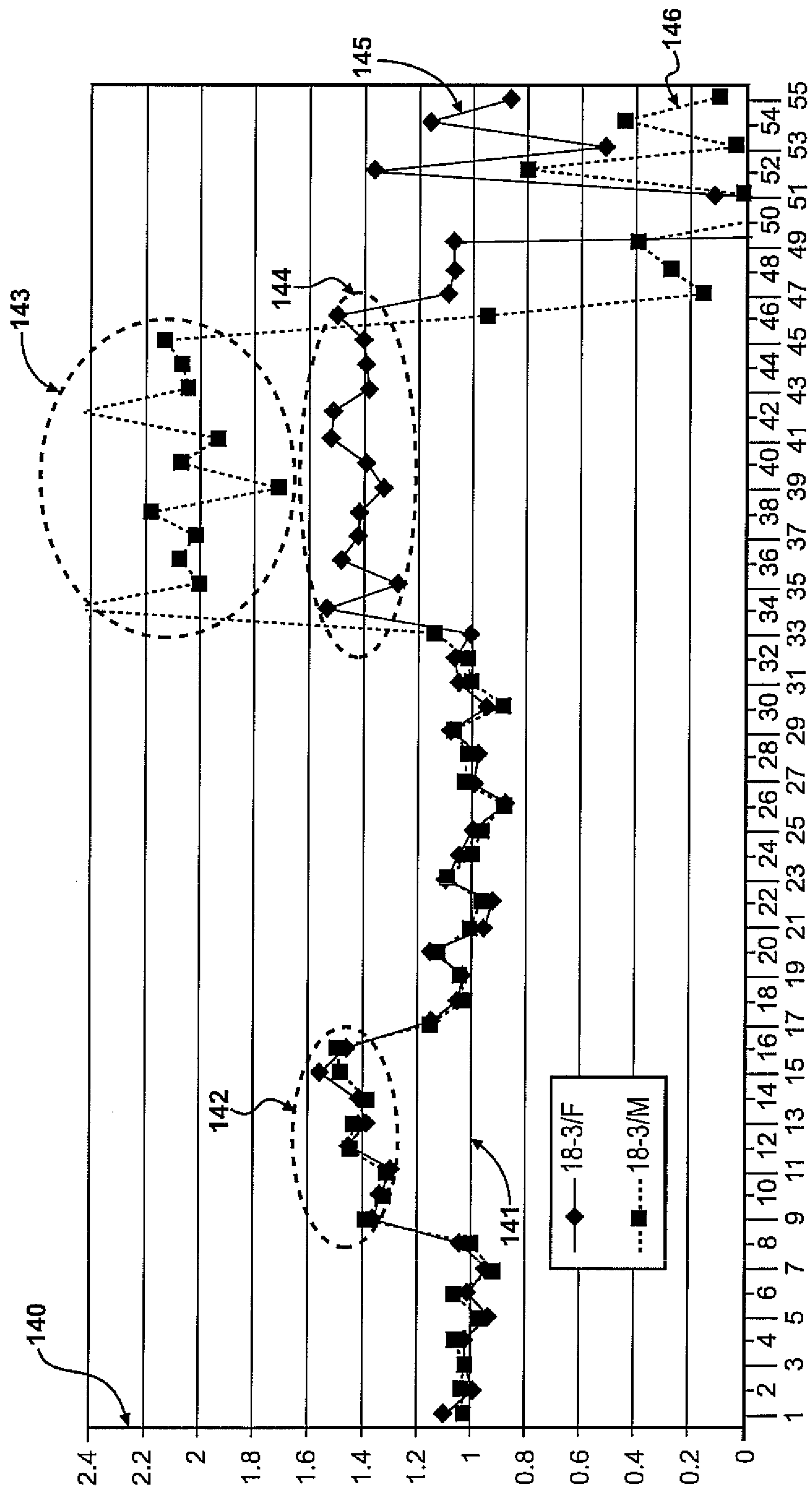

FIG - 14

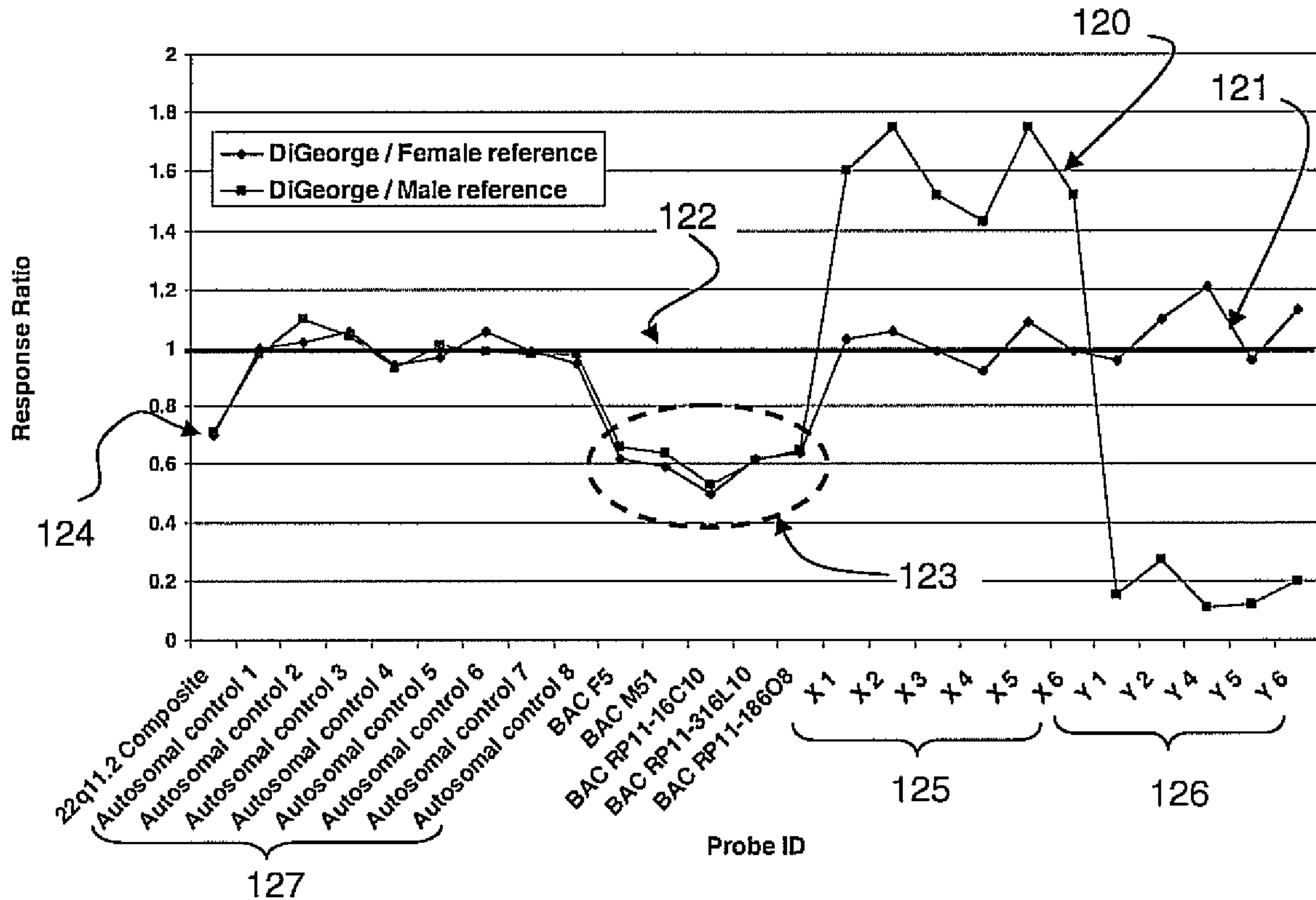

Figure 12