

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
22 June 2006 (22.06.2006)

PCT

(10) International Publication Number
WO 2006/065732 A3

(51) International Patent Classification:
C07H 21/04 (2006.01)

(21) International Application Number:
PCT/US2005/044899

(22) International Filing Date:
12 December 2005 (12.12.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/635,904 13 December 2004 (13.12.2004) US

(71) Applicant (for all designated States except US): **AUTO-GENOMICS, INC.** [US/US]; 2251 Rutherford Road, Carlsbad, CA 92008 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **KIM, Phillip** [US/US]; 28 Exeter, Irvine, CA 92612 (US). **MAHANT, Vijay** [US/US]; 42299 Wild Mustang Road, Murrieta, CA 92562 (US).

(74) Agents: **FESSENMAIER, Martin** et al.; Rutan & Tucker, LLP, 611 Anton Blvd., Suite 1400, Costa Mesa, CA 92626 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, 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, HU, IE, IS, IT, LT, LU, LV, MC, NL, 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:

- with international search report
- with amended claims

(88) Date of publication of the international search report:
5 July 2007

Date of publication of the amended claims: 30 August 2007

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

(54) Title: COMPOSITIONS AND METHODS FOR NUCLEIC ACID ANALYSIS OF SEQUENCES WITH INSERTIONS OR DELETIONS

(57) Abstract: Contemplated methods and kits include a plurality of single stranded oligonucleotides that comprise a discriminating sequence that is encompassed on one end by a label and on the other end by a unique tag sequence. In one preferred aspect, the discriminating sequence is an RNA repeat unit, forms a heteroduplex with a complementary DNA repeat unit of the target nucleic acid, and the discriminating agent is RNaseH. Using such systems, the number of repeat units can be simply identified by hydrolysis of the discriminating sequence where such sequence is adjacent to a predetermined number of DNA repeat units in the single stranded oligonucleotide.



WO 2006/065732 A3

AMENDED CLAIMS
received by the International Bureau on 07 June 2007 (07.06.2007)

What is claimed is:

1. A test kit comprising:

a plurality of first single stranded nucleic acids, each having an unique tag sequence coupled to a targeting sequence, each having a unique number of repeat elements in the targeting sequence, and each further having a discriminating sequence coupled to the targeting sequence, wherein the discriminating sequence is further coupled to a label; and

wherein each of the discriminating sequences has a sequence suitable for at least partial hydrolysis of the single stranded nucleic acid by a discriminating agent under a discriminating condition to thereby separate the label from the unique tag sequence.
2. The kit of claim 1 wherein at least one of the single stranded nucleic acids is a chimeric molecule in which the targeting sequence comprises DNA and in which the discriminating sequence comprises RNA.
3. The kit of claim 1 wherein the discriminating agent is RNaseH.
4. The kit of claim 1 wherein each of the discriminating sequences comprises a repeat sequence, and wherein each of the discriminating sequences has a distinct number of repeats.
5. The kit of claim 1 wherein the label is selected from the group consisting of a fluorophor, a luminophor, a dye, a Raman active moiety, and an enzyme.
6. The kit of claim 1 further comprising an instruction to (a) incubate a sample comprising a nucleic acid with the plurality of single stranded nucleic acids under the discriminating condition to form a test mixture, (b) apply the test mixture to a chip having a plurality of second single stranded nucleic acids, each of the second single stranded nucleic acids having a sequence complementary to the unique tag sequence and being located in a predetermined position, (c) acquire from the chip a plurality of signals from the labels, and (d) determine from the signals a genotype.

7. The kit of claim 1 further comprising a chip having a plurality of second single stranded nucleic acids, each of the second single stranded nucleic acids having a sequence complementary to the unique tag sequence and being located in a predetermined position.
8. The kit of claim 1 further comprising at least one of a reagent that includes RNaseH, a reagent that includes a polymerase or terminal nucleotidyl transferase, and a reagent that includes a labeled nucleotide.
9. A method of determining a genotype of a nucleic acid, comprising:
incubating a sample nucleic acid with a plurality of first single stranded nucleic acids to thereby form a test mixture;
wherein each of the first single stranded nucleic acids has a unique tag sequence that is coupled to a targeting sequence and further has a discriminating sequence that is coupled to the targeting sequence, wherein the discriminating sequence is further coupled to a label;
adding under a discriminating condition a discriminating agent to the test mixture to thereby separate the label from at least one of the first single stranded nucleic acids;
determining separation of the label from the at least one of the first single stranded nucleic acids using the unique tag sequence; and
deducing the genotype from the step of determining.
10. The method of claim 9 wherein the sample nucleic acid comprises an optionally methylated nucleic acid selected from the group consisting of an amplicon, a cDNA, and a genomic DNA.
11. The method of claim 9 wherein the unique tag sequence and the targeting sequence comprise a DNA and wherein the discriminating sequence comprises a RNA.
12. The method of claim 9 wherein the label is selected from the group consisting of a fluorophor, a luminophor, a dye, a Raman active moiety, and an enzyme.

13. The method of claim 9 wherein the discriminating condition is a condition that allows hybridization of the sample nucleic acid with all of the plurality of first single stranded nucleic acids.
14. The method of claim 9 wherein the discriminating agent is RNaseH or a methylation sensitive restriction endonuclease.
15. The method of claim 9 wherein the step of determining separation comprises binding the plurality of the first single stranded nucleic acids to a chip in predetermined positions using the unique tag sequence, and querying for a signal from the label in the predetermined positions.
16. A method of determining a copy number of a repeat unit in a nucleic acid, comprising:
combining a plurality of single stranded nucleic acids of the general formula $A-T-R_n-R'-L$ with the nucleic acid under hybridization conditions to form a duplex;
wherein A is a unique tag DNA sequence, T is a targeting DNA sequence, R is a DNA repeat sequence, n is an integer between 1 and 500, inclusive, R' is a RNA repeat sequence, and L is a label;
hydrolyzing at least part of R' using RNaseH in a duplex where R' and the repeat unit in the nucleic acid form a complementary double strand to thereby separate L from A;
binding the single stranded nucleic acids onto a chip in predetermined positions using A; and
measuring a signal from the predetermined positions.
17. The method of claim 16 further comprising a step of selectively labeling the single stranded nucleic acids in which R' was at least partially hydrolyzed, wherein the step of labeling is performed using a second label that is distinguishable from L.
18. The method of claim 16 wherein T has a length of at least 12 bases and has at least 90% complementarity with a portion of the nucleic acid.
19. The method of claim 16 wherein the step of hydrolyzing is performed using RNaseH.

20. A chimeric oligonucleotide having the formula A-T-R_n-R'-L, wherein A is a unique tag DNA sequence, T is a targeting DNA sequence, R is a DNA repeat sequence, n is an integer between 1 and 500, inclusive, R' is a RNA repeat sequence, and L is a label.
21. A method of determining a copy number of a repeat unit in a nucleic acid, comprising:
- combining a plurality of single stranded nucleic acids of the general formula T-R_n-R'-L* with the nucleic acid under hybridization conditions to form a duplex;
- wherein T is a targeting DNA sequence, R is a DNA repeat sequence, n is an integer between 1 and 500, inclusive, R' is a RNA repeat sequence, and L* is an optional label, and wherein each of the single stranded nucleic acids has not a 3'-OH terminus;
- hydrolyzing at least part of R' using a discriminating agent when R' and the repeat unit in the nucleic acid form a complementary double strand to thereby generate a 3'-OH terminus;
- (a) adding at least one optionally labeled nucleotide to the 3'-OH terminus, or (b) adding a plurality of nucleotides to the 3'-OH terminus and labeling the added nucleotides; and
- detecting the label in solution.
22. The method of claim 21 wherein the step of detecting is performed in real time.