



(51) International Patent Classification:

C12N 15/10 (2006.01) C12Q 1/68 (2006.01)

C12P 19/34 (2006.01)

(21) International Application Number:

PCT/US2017/031346

(22) International Filing Date:

05 May 2017 (05.05.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/332,778 06 May 2016 (06.05.2016) US

(71) Applicant: WILLIAM MARSH RICE UNIVERSITY
[US/US]; 6100 Main St., Houston, TX 77005 (US).(72) Inventors: ZHANG, David; C/o William Marsh Rice University, 6100 Main St., MS-142, Houston, TX 77005 (US).
PINTO, Alessandro; C/o William Marsh Rice University, 6100 Main St., MS-142, Houston, TX 77005 (US).

(74) Agent: HIGHLANDER, Steven, L.; Parker Highlander PLLC, 1120 S. Capital of Texas Highway, Building One, Suite 200, Austin, TX 78746 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,

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

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

Published:

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

(54) Title: STOICHIOMETRIC NUCLEIC ACID PURIFICATION USING RANDOMER CAPTURE PROBE LIBRARIES

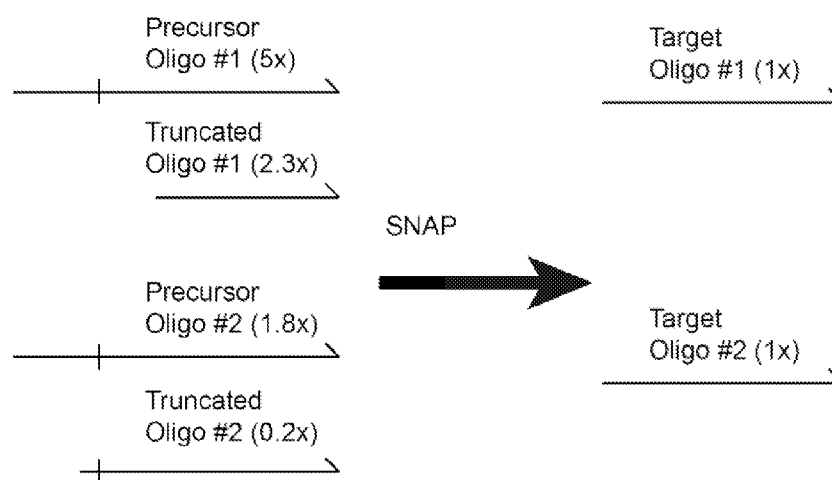


FIG. 1

(57) Abstract: This disclosure describes a method of purifying several full-length oligonucleotide targets from corresponding synthesis truncation products, in a way that ensures roughly stoichiometric equality among the targets.

DESCRIPTION**STOICHIOMETRIC NUCLEIC ACID PURIFICATION USING RANDOMER
CAPTURE PROBE LIBRARIES**

This application claims the benefit of United States Provisional Patent Application No
5 62/332,778, filed May 6, 2016, the entirety of which is incorporated herein by reference.

BACKGROUND

Technologies for writing (Gene Synthesis), editing (CRISPR/CAS) and reading (Next
generation Sequencing-NGS) large collections of nucleic acids requires an enormous (>1000)
number of oligonucleotides to be used as building blocks (writing), guides (editing) or
10 hybridization probes and primers for doing highly multiplexed enrichment and sequencing
(reading). It is not economical to synthesize, purify, and quantitate each oligonucleotide
individually. Companies such as Agilent, NimbleGen, and Twist Biosciences have developed
array-based synthesis platforms to allow highly multiplex DNA oligonucleotide synthesis,
but the oligonucleotides synthesized by these platforms include truncation products. Because
15 modern oligonucleotide synthesis occurs from 3' to 5', most impurity species are truncation
oligonucleotide products lacking a number of nucleotides at the 5' end, followed by species
with one or more internal deletions. In single-plex synthesis, these fraction of these impurity
products can be reduced through post-synthesis high pressure liquid chromatography (HPLC)
or polyacrylamide gel electrophoresis (PAGE) purification, but HPLC and PAGE cannot be
20 used to purify a pool of many different oligonucleotides. Furthermore, HPLC and PAGE are
time- and labor-intensive and cannot be easily automated to high throughput. Additionally,
even single-plex HPLC and PAGE purification of oligonucleotides result only in below 90%
purity of full-length oligonucleotide products.

The concentrations of different oligonucleotides in an array- synthesized pool will
25 vary significantly based on oligonucleotide length, oligonucleotide sequence, and synthesis
reagent age and purity. Consequently, oligonucleotides synthesis yields can vary by more
than 16-fold from the same quantity of initial synthesis reagents. The variation in
oligonucleotide concentrations can adversely affect downstream applications, *e.g.*, in the
production of long synthetic genes. In NGS, concentration variations in oligonucleotide
30 pools used for hybrid-capture enrichment result in sequencing biases that cause significant
wasted NGS reads.

SUMMARY

In accordance with the present disclosure, there is provided a method for generating a set of precursor nucleotide sequences comprising a target oligonucleotide molecule, wherein the precursor nucleotide sequence comprises a fifth region that comprises the nucleotide sequence of the target oligonucleotide molecule and a fourth region and a third region, wherein at least one of the fourth and third regions differs from any subsequence within the target oligonucleotide, the method comprising:

- (i) calculating for the precursor nucleotide sequence the standard free energy of hybridization between the precursor nucleotide sequence and (a) a first oligonucleotide comprising a second region that is complementary to the third region of the precursor oligonucleotide sequence, and (b) a first region that is complementary to the fourth region of the precursor oligonucleotide sequence;
- (ii) calculating the standard free energy of a capture reaction as the standard free energy of hybridization between the precursor nucleotide sequence and the first oligonucleotide and the standard free energy of hybridization between the first oligonucleotide and target oligonucleotide;
- (iii) rejecting the precursor nucleotide sequence if the standard free energy of the capture reaction does not meet a certain criterion; and
- (iv) repeating steps (i) to (iii) until a set of precursor nucleotide sequences meets the criteria; and
- (v) producing the set of precursor nucleotide sequences.

The criteria may be a negative free energy;

In another embodiment, there is provided a method for producing a set of precursor nucleotide sequences comprising a plurality of barcode sequences, comprising:

- (i) generating a set of precursor nucleotide sequences comprising each target oligonucleotide molecule, wherein each precursor nucleotide sequence comprises a third region that is conserved across all precursor nucleotide sequences, a fourth region that is unique for each target oligonucleotide molecule, and a fifth region that comprises the nucleotide sequence of a target oligonucleotide molecule;
- (ii) calculating for each precursor nucleotide sequence the standard free energy of hybridization between the precursor nucleotide sequence and (a) a first oligonucleotide, comprising a second region that is complementary to the third region

of the precursor oligonucleotide sequence, and (b) a first region that is complementary to the fourth region of the precursor oligonucleotide sequence;

(iii) calculating for each precursor nucleotide sequence the standard free energy of hybridization of folding;

5 (iv) calculating the standard free energy of a capture reaction as the standard free energy of hybridization between the precursor nucleotide sequence and the first oligonucleotide and the standard free energy of folding of the first oligonucleotide and the standard free energy of folding of the precursor oligonucleotide;

10 (v) rejecting the set of precursor nucleotide sequences if the standard free energy of the capture reaction for any precursor nucleotide sequence exceeds a certain criteria;

(vi) repeating steps (i) to (v) until a set of precursor nucleotide sequences meets the criteria; and

(vii) producing said set of precursor nucleotide sequences.

15 The criteria may be a selected from the group consisting of a maximum range of standard free energies of capture, a standard deviation of standard free energies of capture, and a difference between two ranks in a sorted list. The criteria may be a maximum range of no more than 5 kcal/mol between a lowest standard free energy of capture and a highest standard free energy of capture for the set of precursor nucleotide sequences. The maximum range may be no more than 2 kcal/mol.

20 In yet another embodiment, there is provided a method for purifying one or multiple target nucleic acid molecules from a sample comprising one or a plurality of species of precursor molecules, wherein each species of precursor molecule comprises an fifth region comprising a target nucleic acid molecule sequence, a fourth region comprising a sequence unique to the species of precursor molecule in the plurality of species of precursor molecules,
25 defined as a barcode sequence of length n , wherein 2^n is greater than or equal to the number of unique target nucleic acid molecule sequences, a third region that is conserved across all precursor molecules, the method comprising:

contacting the sample with a capture probe library at temperature and buffer conditions conducive to hybridization;

30 wherein the capture probe library comprises a plurality of capture probe species, wherein each capture probe species comprises a first oligonucleotide comprising a first region comprising a nucleotide sequence of n nucleotides in length and a second region that is conserved across all capture probe species, wherein each nucleotide in

the nucleotide sequence of n nucleotides in length is selected from two or more nucleotides and the first region is unique to each capture probe, and wherein the second region is complementary to the third region, and wherein the fourth region of each species of precursor molecule is complementary to the first region of a species of precursor molecule;

separating the plurality of species of precursor molecules hybridized to the plurality of capture probe species from the species of precursor molecules not hybridized to the plurality of capture probe species;

treating the plurality of species of precursor molecules hybridized to the plurality of capture probe species with a cleavage agent sufficient to site-specifically cleave the plurality;

of species of precursor molecules at a site to separate the fifth regions from at least a portion of the third and fourth regions;

recovering the fifth regions from the plurality of capture probe species and the at least a portion of the third and fourth regions, and

thereby producing a purified target nucleic acid molecule or molecules.

Each capture probe species may further comprise a second oligonucleotide comprising a ninth region, wherein the ninth region is complementary to the second region. Each first oligonucleotide may further comprise a seventh region, each second oligonucleotide further comprises an eighth region, and wherein the seventh region is complementary to the eighth region. Each first oligonucleotide further may comprise a chemical moiety, and wherein the separating the plurality of species of precursor molecules hybridized to the plurality of capture probe species comprises surface capture of the chemical moiety. The chemical moiety may be selected from the group consisting of biotin, a thiol, an azide, an alkyne, a primary amine and a lipid. The first nucleotide hybridized with the precursor oligonucleotide may be the preferred ligand of an antibody or other receptor that mediate the surface capture of the complexes.

Recovering the fifth regions from the plurality of capture probe species and the at least a portion of the third and fourth regions may comprise a treatment selected from the group consisting of heating, introducing denaturants, washing with low salinity buffers, and introducing a nuclease. The site-specific cleavage may comprise a treatment selected from the group consisting of changing the temperature, changing the pH, and illuminating the plurality of species of precursor molecules hybridized to the plurality of capture probe

species at a specific wavelength. The standard free energies of binding between each first oligonucleotide and a DNA sequence complementary to the entire sequence of the first oligonucleotide may be within 5 kcal/mol of each other. The two or more nucleotides at each nucleotide in the nucleotide sequence of n nucleotides in length maybe A or T, or may be G or C. The two or more nucleotides at each nucleotide in the nucleotide sequence of n nucleotides in length may be G or C for one or more nucleotides in the nucleotide sequence and A or T for one more nucleotides in the nucleotide sequence. The first region may comprise between 3 and 25 nucleotides. n may be between 3 and 60, 3 and 18, or 3 and 10 and not greater than the number of nucleotides in the first region. The first region may further comprise at least one nucleotide in addition to the nucleotide sequence of n nucleotides in length. The second region may further comprise between 8 and 200 nucleotides.

The barcode sequence of each species of precursor molecule is assigned based on a method comprising:

generating a set of precursor nucleotide sequences comprising each target oligonucleotide molecule, wherein each precursor nucleotide sequence comprises a third region that is conserved across all precursor nucleotide sequences, a fourth region that is unique for each target oligonucleotide molecule, and an fifth region that comprises the nucleotide sequence of the target oligonucleotide molecule;

calculating for each precursor nucleotide sequence the standard free energy of hybridization between the precursor nucleotide sequence and a first oligonucleotide, comprising a second region that is complementary to the third region of the precursor oligonucleotide sequence, and a first region that is complementary to the fourth region of the precursor oligonucleotide sequence;

calculating for each precursor nucleotide sequence the standard free energy of hybridization of folding;

calculating the standard free energy of a capture reaction as the standard free energy of hybridization between the precursor nucleotide sequence and the first oligonucleotide and the standard free energy of folding of the first oligonucleotide and the standard free energy of folding of the precursor oligonucleotide;

rejecting the set of precursor nucleotide sequences if the standard free energy of the capture reaction for any precursor nucleotide sequence exceeds a certain criteria; and repeating the method until a set of precursor nucleotide sequences meets the criteria.

In still yet another embodiment, there is provided a capture probe library comprising:
a plurality of oligonucleotides comprising a first plurality of oligonucleotides wherein
each oligonucleotide of the first plurality of oligonucleotides comprises:

a first region comprising a first nucleotide sequence comprising at least 3 variable
positions, wherein each variable position comprises a nucleotide selected from at least
two possible nucleotides,

wherein the first nucleotide sequence comprising at least 3 variable positions is
unique to each oligonucleotide, and

a second region comprising a second nucleotide sequence, wherein the second
nucleotide sequence of the second region is conserved across each oligonucleotide in
the first plurality of oligonucleotides,

A capture probe library comprising:

a plurality of oligonucleotides comprising a first plurality of oligonucleotides and a
second plurality of oligonucleotides, wherein each oligonucleotide in the first
plurality of species of oligonucleotides comprises

a first region comprising a nucleotide sequence comprising at least 3 variable
positions, wherein each variable position comprises a nucleotide selected from at least
two possible nucleotides,

wherein the nucleotide sequence comprising at least 3 variable positions is unique to
each species of oligonucleotide, and

a second region comprising a nucleotide sequence, wherein the nucleotide sequence
of the second region is conserved across each oligonucleotide in the first plurality of
oligonucleotides,

wherein each oligonucleotide in the second plurality of oligonucleotides comprises a
third region, wherein the third region is complementary to the second region.

The standard free energies of binding between each oligonucleotide in the first plurality of
oligonucleotides and a DNA sequence complementary to the entire sequence of the respective
oligonucleotide in the first plurality of oligonucleotides may be within 5 kcal/mol of each
other. The at least two possible nucleotides at each nucleotide in may be A or T, or may be G
or C. The

at least two possible nucleotides at each nucleotide in may be G or C for one or more
nucleotides in the nucleotide sequence and A or T for one more nucleotides in the nucleotide
sequence. The concentration of a second oligonucleotide may be greater than a sum of the
concentrations of each oligonucleotide of the first plurality of oligonucleotides. The first

region may comprise between 3 and 25 nucleotides. The number of variable positions in the first region may be between 3 and 60, between 3 and 18, or between 3 and 10, and not greater than a total number of nucleotides in the first region. The first region further may comprise at least one nucleotide in addition to the at least 3 variable positions. The second region may
5 comprise between 8 and 200 nucleotides. The at least three variable regions may be contiguous or non-contiguous.

A further embodiment comprise an oligonucleotide library for the multiplexed capture of a set of desired precursor nucleic acid molecules comprising:

- 10 a plurality of species of precursor molecules, wherein each species of precursor molecule comprises
 - a third region comprising a nucleotide sequence that is conserved across all species of precursor molecules,
 - a fourth region comprising a barcode sequence comprising 3-60 nucleotides,
 - an fifth region comprising a target nucleic acid molecule sequence that is unique to
15 the species of precursor molecule in the plurality of species of precursor molecules, wherein the barcode sequence of each species of precursor molecule is different and wherein 2^n is greater than or equal to the number of unique target nucleic acid molecule sequences; and
- 20 a capture probe library comprising a plurality of capture probe species, wherein each capture probe species comprises an oligonucleotide comprising a first region comprising a nucleotide sequence of n nucleotides in length, a second region that is conserved across all capture probe species, wherein each nucleotide in the nucleotide sequence of n nucleotides in length is selected from two or more nucleotides and the first region is unique to each capture probe, and wherein the second region is
25 complementary to the third region, and wherein the fourth region of each species of precursor molecule is complementary to the first region of a species of precursor molecule.

The the standard free energies of binding between each species of first oligonucleotide in the plurality of capture probe species and a DNA sequence complementary to the entire sequence
30 of the respective species of first oligonucleotide may be within 5 kcal/mol of each other. The two or more nucleotides at each nucleotide in the nucleotide sequence of n nucleotides in length may be A or T, or may be G or C. The two or more nucleotides at each nucleotide in the nucleotide sequence of n nucleotides in length may be G or C for one or more nucleotides in the nucleotide sequence and A or T for one more nucleotides in the nucleotide sequence.

The a concentration of a second oligonucleotide may be greater than a sum of the concentrations of each species of first oligonucleotide. The first region may comprise between 3 and 25 nucleotides. n may be between 3 and 60, 3 and 18, or 3 and 10, and not greater than the number of nucleotides in the first region. The first region may further
5 comprise at least one nucleotide in addition to the nucleotide sequence of n nucleotides in length. The second region may comprise between 8 and 200 nucleotides. The barcode sequence of each species of precursor molecule may be selected based on the method as set forth above. The at least one species of precursor molecule may be chemically synthesized or produced enzymatically. The enzyme used to produce the at least one species of precursor
10 molecule may be a ligase or a polymerase.

In an additional embodiment, there is provided a method for purifying multiple target nucleic acid molecules from a sample comprising a plurality of precursor molecules, wherein the method comprises the steps of:

providing a plurality of nucleic acid probes, wherein each probe has a sequence
15 complementary to a region of one of the precursor molecules and a first moiety sufficient to allow isolation of the probe;

adding the plurality of nucleic acid probes to a sample comprising the plurality of precursor molecules under conditions sufficient to promote hybridization of each nucleic acid probe to the region of the precursor molecule complimentary to the
20 sequence thereby forming a plurality of probe- precursor complexes, wherein each precursor molecule comprises a different target nucleic acid molecule, and wherein the region of each precursor molecule does not comprise the target nucleic acid molecule; and

isolating the plurality of probe-precursor complexes via interaction with the first
25 separating the target nucleic acid molecules from the plurality of probe-precursor.

Another embodiment provides for a method for producing a plurality of distinct target oligonucleotides each having a specified sequence, the method comprising:

(1) synthesizing a precursor oligonucleotide for each distinct target oligonucleotide, wherein the precursor comprises a third sequence, a fourth sequence,
30 and a fifth sequence, wherein the third sequence is identical for all precursors, the fourth sequence comprises a barcode and is distinct for all precursors, and a fifth sequence corresponds to the target sequence,

(2) synthesizing a capture probe library of distinct oligonucleotides comprising a first sequence and a second sequence, wherein the first sequence comprises

degenerate randomer nucleotides and wherein at least one first sequence is complementary to each fourth sequence, and the second sequence is complementary to the third sequence,

(3) mixing the precursors and the capture probe library in an aqueous hybridization buffer,

(4) removing precursor molecules not bound to the capture probes,

(5) enzymatically or chemically cleaving the fifth sequence from the remainder of the precursor molecules, and

(6) removing the capture probe library and the remainder of the precursor molecules.

The first sequence may comprise an S degenerate nucleotide at certain positions and/or a W degenerate nucleotide at certain positions, but may not comprise an N degenerate nucleotide at any position, such that any degenerate variant of the first sequence is complementary to one or more fourth sequences, wherein S is a strong base, W is a weak base, and N is any base. The capture probes may be functionalized with a moiety permitting for rapid binding to a ligand, such as a moiety selected from a biotin, a thiol, an azide, or an alkyne. In step (4), eliminating precursor molecules not bound to capture probes may comprise adding particles that specifically bind the capture probe, followed by removal of supernatant solution. The particles may be selected from streptavidin-coated magnetic beads or streptavidin-coated agarose beads. The precursors may further comprise a deoxyuracil nucleotide or an RNA nucleotide, and cleavage of the fifth sequence comprises introduction of a uracil DNA glycosylase or an RNase enzyme. The precursors may further comprise a photolabile or heat-labile moiety, and the cleaving of the fifth sequence comprises exposure of the solution to light of the appropriate wavelength or heating to the appropriate temperature. The fourth sequence may further comprise the "S" degenerate nucleotide at certain positions and/or the "W" degenerate nucleotide at certain positions, but does not comprise the "N" degenerate nucleotide at any position. The length of the first sequence may be between 5 and 50 nucleotides, and wherein the number of degenerate nucleotides is between 1 and 30, between 1 and 20, between 1 and 10, between 2 and 8, between 2 and 6, or between 3 and 5. The length of the second sequence may be between 5 and 50 nucleotides, and/or wherein the length of each target oligonucleotide is between 5 and 500 nucleotides.

In an additional embodiment, there is provided an oligonucleotide capture probe library comprising a first sequence and a second sequence, wherein the first sequence comprises degenerate randomer nucleotides comprising an "S" degenerate nucleotide at one

or more positions and/or a "W" degenerate nucleotide at one or more positions, but does not comprise an "N" degenerate nucleotide at any position, and wherein the length of the first sequence is between 5 and 50 nucleotides, the number of degenerate nucleotides is between 1 and 30, and the length of the second sequence is between 5 and 50 nucleotides. The second
5 sequence may comprise an "S" degenerate nucleotide at certain positions and/or a "W" degenerate nucleotide at certain positions, but does not comprise a "N" degenerate nucleotide at any position. The oligonucleotide capture probe library may be functionalized with a chemical moiety for rapid binding, selected from a biotin, a thiol, an azide, or an alkyne. The one or more of the oligonucleotide capture probes may further comprise a deoxyuracil
10 nucleotide or an RNA nucleotide. The one or more of the oligonucleotide capture probes may further comprise a photolabile or heat-labile moiety. The length of the first sequence may be between 5 and 50 nucleotides, and wherein the number of degenerate nucleotides may be between 1 and 30. The length of the second sequence may be between 5 and 50 nucleotides. The library may have at least 8, at least 32, or at least 256 members. The library may have
15 between 8 and 32 members, between 8 and 256 members, between 32 and 256 members, between 8 and 1024 members, between 32 and 1024 members, or between 256 and 1024 members. The library may be found on one or more substrates.

In still an additional embodiment, there is provided an aqueous solution comprising
20 an oligonucleotide capture probe library, a plurality of precursor oligonucleotides and a set of precursor oligonucleotides, wherein:

the capture probe library comprises a first sequence and a second sequence, wherein the first sequence comprises degenerate randomer nucleotides comprising an "S" degenerate nucleotide at certain positions and/or an "W" degenerate nucleotide at
25 certain positions, but does not comprise an "N" degenerate nucleotide at any position, and wherein the length of the first sequence is between 5 and 50 nucleotides, the number of degenerate nucleotides is between 1 and 30, and the length of the second sequence is between 5 and 50 nucleotides,

each of the plurality of precursor oligonucleotides comprises a third sequence, a
30 fourth sequence, and a fifth sequence, wherein the third sequence is identical for all precursors, the fourth sequence comprises a barcode and is distinct for all precursors, the second sequence is complementary to the third sequence, and at least one instance of a first sequence is complementary to each fourth sequence.

The second sequence of the oligonucleotide capture probe library may comprise an "S" degenerate nucleotide at certain positions and/or a "W" degenerate nucleotide at certain positions, but may not comprise a "N" degenerate nucleotide at any position. The oligonucleotide capture probe library may be functionalized with a chemical moiety for rapid binding, selected from a biotin, a thiol, an azide, or an alkyne. The one or more of the oligonucleotide capture probes further may comprise a deoxyuracil nucleotide or an RNA nucleotide. The one or more of the oligonucleotide capture probes may further comprise a photolabile or heat-labile moiety. The length of the first sequence of the oligonucleotide capture probe library may be between 5 and 50 nucleotides, and the number of degenerate nucleotides may be between 1 and 30. The length of the second sequence of the oligonucleotide capture probe library may be between 5 and 50 nucleotides. The oligonucleotide capture probe library may have at least 8, at least 32, or at least 256 members. The oligonucleotide capture probe library may have between 8 and 32 members, between 8 and 256 members, between 32 and 256 members, between 8 and 1024 members, between 32 and 1024 members, or between 256 and 1024 members. The oligonucleotide capture probe library may be found on one or more substrates. The precursors may further comprise a deoxyuracil nucleotide or an RNA nucleotide, and cleavage of the fifth sequence comprises introduction of a uracil DNA glycosylase or an RNase enzyme.

The use of the word "a" or "an" when used in conjunction with the term "comprising" in the claims and/or the specification may mean "one," but it is also consistent with the meaning of "one or more," "at least one," and "one or more than one." The word "about" means plus or minus 5% of the stated number.

It is contemplated that any method or composition described herein can be implemented with respect to any other method or composition described herein. Other objects, features and advantages of the present disclosure will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating specific embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the disclosure will become apparent to those skilled in the art from this detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present disclosure. The disclosure may be better understood by reference to one or more of these drawings in combination with the detailed
5 description of specific embodiments presented herein.

Fig. 1: Schematic overview of one embodiment of Stoichiometric Nucleic Acid Purification (SNAP). A sample mixture containing two precursors of oligonucleotide targets also contains a number of truncation synthesis products that are not desired. Through the
10 course of the method described in this disclosure, full length target oligonucleotides are produced in roughly equal stoichiometry.

Fig. 2: Schematic of a capture probe library and corresponding set of target sequences. Within the targets libraries, region 3 is conserved while regions 4 and 5 are different for each target sequence. Similarly, within a library of capture probes region 2 is conserved, while
15 region 1 contains variable positions with two or more nucleosides being possibly present at each variable position. (SEQ ID NOS: 404-409)

Fig. 3: Multiplexed oligonucleotide purification workflow. The process consists of 6 sequential steps. In step 3, solid phase consists of functionalized magnetic beads that enable the phase separation in steps 4 and 6. Depending on whether the user desires target
20 sequences with or without the universal region 3 and barcode region 4, in step 5 either heat/NaOH denaturation or restriction enzyme cleavage can be used after the purification to separate the final products from the library and magnetic beads.

Fig. 4: Removal of regions 3 and 4 from target sequences using an enzyme. The desired region 5 can be enzymatically cleaved from the 5' regions 3 and 4 used for
25 purification, for applications where such sequences are undesirable. Site-specific cleavage can be implemented through use of (a) uracil DNA glycosylase alone or in a formulation containing DNA glycosylase-lyase Endonuclease VIII - commercially known as USER, (b) RNase, (c) FokI endonuclease, or other methods known to one of ordinary skill in the art. The sequence or chemical composition of regions 3 and 4 may be adapted to accommodate
30 the anticipated enzymatic cleavage process. (SEQ ID NOS: 410-417 and 251)

Fig. 5: Use of multiple barcodes for the same target sequence increases its concentration in the final purified mixture. Degenerate nucleotides (*i.e.*, S = G or C and W = A or T) within the barcode region 4 facilitates the achievement of arbitrary desired stoichiometric ratios of targets by avoiding individual synthesis of a number numbers of precursor oligonucleotides. (SEQ ID NOS: 418-421)

Fig. 6: Proof of concept results using fluorescent labelled oligonucleotides, assayed through fluorescent polyacrylamide gel electrophoresis. 160 pmol of a capture probe library biotinylated in 3' end, comprising 64 different region 1 sequences, is incubated with 64 synthetic unpurified precursor oligonucleotides each at 10 pmol. Each precursor oligonucleotide has a uniquely assigned region 4, complementary to one and only one instance of region 1 on the capture probe species. Lane 1 shows the length distribution of Precursor 1, and Lane 2 shows the length distribution of Precursor 1 after undergoing SNAP, but without cleavage of regions 3 and 4. Lane 3 shows the SNAP product after cleavage of regions 3 and 4. Lane 4 shows the length distribution of the corresponding commercially provided PAGE-purified oligo as comparison. Lanes 5-10 show that, regardless of initial stoichiometric ratio of the 3 Precursors, the final stoichiometry of the SNAP products are close to 1:1:1. The SNAP protocol used for this series of experiments is as follows: The capture probe library and Precursors are allowed to hybridize for 2 hr at 60 °C in 0.5M NaCl pH 7.5. Subsequently, 3 mg of streptavidin coated magnetic beads (pre-washed with the same incubation buffer) are added to the oligonucleotide mixture to a final volume of 100 µL, and incubated for 30 min at 60 °C, with the sample being rocked by a shaker. The supernatant is then discarded and the magnetic beads are washed twice in the incubation buffer and one time in the buffer to be used for the subsequent enzymatic cleavage. For lane 2, the full-length precursors are eluted using 50% v/v formamide at 95 °C for 5 to 15 min. For lane 3, the beads were re-suspended in USER buffer, and 2 enzymatic units of USER enzyme mix were added to achieve a final volume of 25 µL; this mix was then incubated for 1 hour at 37 °C. Finally supernatant containing the desired purified target is extracted.

Figs. 7A-B. Characterization of purity and stoichiometry for multiplexed SNAP. 64 separate Precursors (10 pmol each) were simultaneously purified via SNAP. (Fig. 7A) Next Generation Sequencing was used to characterize the purity of the 64 oligos, where the purity is operationally defined as the number of perfectly aligned reads divided by the number of reads aligned by Bowtie 2. The purity of the SNAP products are significantly higher than

that of individually PAGE purified oligonucleotides (median 79% vs. 61%). (FIG. 7B) Digital droplet PCR was used to evaluate the stoichiometries of the 64 SNAP product oligonucleotides.

Fig. 8: Characterization of purity for 256-plex SNAP purified (2.5 pmol of each Precursor).

Figs. 9A-D: Four possible variations in design of the capture probe libraries.

Fig.10: Experimental results on double-stranded capture probe article shown in Fig. 9B. SNAP was performed using 640 pmol of capture probe biotinylated in 5' end, 1280 pmol of protector strand, and 40 pmol of one unpurified Precursor. Note that, unlike in Fig. 6, only a single Precursor species is introduced, so that the Capture Probes are NOT saturated by Precursors. SNAP protocol is otherwise similar to Fig. 6 caption. Lane 1 corresponds to the unpurified Precursor, Lane 2 corresponds to the PAGE-purified Target oligo, Lane 3 corresponds to the SNAP product without cleavage of regions 3 and 4, and Lane 4 corresponds to the final SNAP product. Lanes 5-10 show that, regardless of initial stoichiometric ratio of the 3 Precursors, the final stoichiometry of the SNAP products are close to 1:1:1.

Fig. 11: Purification of RNA transcripts produced from a synthetic DNA template that comprises regions 3 and 4.

Fig.12 Cumulative distribution of ΔG°_{rxn} for two different randomer sequences, each producing 64 instances. Using an SWSWSW library (flanked by CA in 5' end and C in 3' end) produces a tight thermodynamics range of 0.8 kcal/mol window. In contrast, using a TGANN library results in a spread of more than 4 kcal/mol.

Figs.13A-C: Algorithm for the design of probe libraries and assignment of barcodes (region 4).

Fig. 14: Workflow for the characterization of purity of oligonucleotides libraries through NGS. The oligonucleotides share a universal region at the 3' end, which is used to align and enzymatically extend one short sequence used as primers. Upon the phosphorylation of the 5' end of the newly formed double stranded library, adaptors containing sequencing primers are enzymatically attached. Finally, the resulting library is

amplified for 3 PCR cycles, to introduce the P7 and P5 sequences used for the cluster generation and the consequent sequencing with an Illumina instrument.

Fig. 15: Workflow for the characterization of stoichiometries within oligonucleotides libraries through digital droplet PCR (ddPCR). The oligonucleotides of the library share a common region at 3' end for the priming with a reverse primer, which is added as well as the master mix contain the enzyme and the EvaGreen dye for ddPCR reaction. The reaction mixture is distributed in 64 equal aliquots, each of which receives one forward primer specific for one oligonucleotide sequence within the library. Subsequently oil is added to every sample, which undergoes to the emulsion process. Finally, upon the PCR reaction, the fluorescent reader is used to determine the ratio between the droplets-containing amplification product and those that are empty, ratio that give a statistical quantification of the template molecule that have been amplified.

DETAILED DESCRIPTION

The goal of this disclosure is outlined shown in Fig. 1: different precursor oligonucleotides are synthesized with various yields and purities, and pooled together to form an input oligonucleotide pool. Through a process of SNAP purification, an output pool of target oligonucleotides is produced with no truncation products, and exhibiting a desired stoichiometric ratio (1:1 in Fig. 1).

In certain aspects of the present disclosure, toehold probes with a randomer toehold sequence are used to capture artificially designed 5' sequence of the target oligonucleotides. Because the probes are toehold probes which are selective to single nucleotide variations, even truncated synthesis products one nucleotide shorter than the full-length product will not be efficiently captures.

I. Precursor Oligonucleotides

A full-length precursor oligonucleotide comprises three regions, labeled as 3, 4, and 5 in Fig. 2. Region 3 is also referred to as the validation region, and the sequence of region 3 is conserved across all precursor species. Region 4 is also referred to as the barcode, and the sequence of region 5 is unique to each precursor species. Region 5 is referred to as the target sequence, and is the only portion of the oligonucleotide that should remain after the purification process. Across different precursor species, some region 8's may be unique, while others may be redundant. In some embodiments, the nucleotide to the 5' of region 5 is a modified nucleotide (*e.g.*, a deoxyuracil nucleotide, or an RNA nucleotide) that can be site-specifically cleaved.

Because DNA synthesis (both chemical and enzymatic) is imperfect, there will exist truncation products in which precursors lack one or more nucleotides at either the 5' end (chemical synthesis) or the 3' end (enzymatic synthesis).

II. Capture Probe Library

Fig. 2 shows one preferred embodiment of the capture probe library used to perform SNAP purification for chemically synthesized oligonucleotides with truncations at the 5' end. In this embodiment, the library comprises two types of oligonucleotides: the probe oligonucleotides (comprising regions 1 and 2). In certain aspects, the probe oligonucleotides

are functionalized at the 3' end with a biotin, in order to allow capture by streptavidin-functionalized magnetic beads.

The sequence of region 1 is designed and synthesized as a randomer library, in which one or several positions contain a mixture of multiple nucleotides. The complement of every precursor species' barcode (region 4) should exist as an instance of the region 1 randomer library.

The sequence of region 2 is designed to be complementary to the sequence of region 3 on precursors.

III. Sequence Design of Randomer Region 1

The variable positions and allowable nucleotides at each variable position should be designed such that the standard free energy of hybridization of each instance region 1 to its perfect complement are similar. In some embodiments, the sequence of region 1 comprises S (strong, mixture of G and C) and W (weak, mixture of A and T) degenerate nucleotides.

As one example of an undesirable sequence construction, if region 4 is designed as a 7nt NNNNNNN region, then both GCGCGCG and TATATAT members will be present. The ΔG_0 of these two members pairing with their complements at 37 °C in 1M Na⁺ are -13.23 kcal/mol and -4.38 kcal/mol, respectively, according to SantaLucia Jr, J., & Hicks, D. (2004). The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, 33, 415-440. This 9 kcal/mol difference can result in the GCGCGCG member capturing its target with >99.9% yield while the TATATAT member capturing its target with <0.1% yield; such a large difference in capture yield would be clearly undesirable for achieving uniform or ratiometric product quantity/concentration distributions.

For this reason, the nucleosides present at variable positions are designed to be either S or W. That is to say, some variable positions contain either an A or T nucleoside but not G or C, while other variable positions contains G or C but not A or T. Based on published literature parameters, there is only a maximum difference of 0.17 kcal/mol per base stack for SW and for WS stacks, at 37 °C in 1M Na⁺.

IV. Sequence-Specific Capture

In those instances where the number of different probes is equal the number of the target sequences, and the total concentration of probes is lower than the total concentration of target, any instance of region 1 only hybridizes to its perfectly complement in region 4, as
5 any other non-specific hybridization will be outcompeted.

Consequently, if the probe oligonucleotide library is synthesized such that all instance sequences are equally represented, and if the concentration of all precursors exceed that of their corresponding probe sequences, then the amount of precursor captured should be roughly stoichiometric, regardless of the initial stoichiometric ratio between the precursors.
10 As a numerical example, if the sequence of region 1 is "GWSWSWST", then there are $2^6 = 64$ instance sequences. Assuming a total probe concentration of $6.4 \mu\text{M}$, each sequence instance would have a concentration of approximately 100 nM . For an initial precursor pool in which the concentrations of each precursor species ranges between 200 nM and $10 \mu\text{M}$, the amount of each precursor bound to the probe will be limited by probe instance sequence
15 concentration to a maximum of 100 nM , except insofar as off-target hybridization between precursors and their non-cognate probe instance sequences hybridize.

As another mathematical example, a probe library with 12 variable position, and 2 possible nucleotides at each position is comprised of $2^{12} = 4096$ members. Assuming a total amount of 4 nanomoles (nmol) of the library, each member is expected to be present at
20 quantity of roughly 1 pmol . This library is suitable for purification of up to 4096 targets, each with quantity of $\geq 10 \text{ pmol}$. Array oligonucleotide synthesis providers often produce panels of oligonucleotides at either the 10 pmol or 100 pmol scale.

V. Separation of Precursors Bounded to Probes

The precursor oligonucleotides bound to the probe oligonucleotides are separated
25 from other precursors using the probes as marker for recovery, through the use of a solid support or enzymatic degradation of unbound molecules, for example, using an exonuclease (*e.g.*, 5'-3') for single-strand digestion. In a particular embodiment, the probe oligonucleotides are biotin-functionalized at the 3' end, and streptavidin-functionalized magnetic beads are added to solution after the hybridization reaction between the precursors,
30 protectors, and probes. Washing the magnetic bead suspension in the vicinity of a magnetic removes unbound molecules.

VI. Removal of Regions 3 and 4 from Targets

For many applications with purified pools of target oligonucleotides, the sequences of regions 3 and 4 would be an undesirable artifact. The sequence or composition of these regions may be designed to facilitate enzymatic removal of these regions from the desired target sequence after surface-based purification. Fig. 4 shows several strategies for enzymatically cleaving the captured target sequences after region 4 to remove all artifact sequences, or after region 3 to remove the purification sequence but not the barcode.

VII. SNAP Purification Workflow

Fig. 3 shows one embodiment of the overall SNAP purification workflow. Hybridization durations, buffers, and temperatures vary depending on the concentration of the probe and precursor oligonucleotides, and reasonable parameter values are known to those ordinary in the art of nucleic acid hybridization probes. The capture of biotin-functionalized probes by magnetic beads and subsequent wash protocols are typically provided by suppliers of biotin-functionalized magnetic beads (*e.g.*, Thermo Fisher Dynabeads, New England Biolabs streptavidin magnetic beads). The USER enzyme (*e.g.*, from New England Biolabs) can be used to site-specifically cleave precursor oligonucleotides at dU positions.

VIII. Ratiometric Concentrations of Purified Targets

Through the use of multiple barcodes (region 4) for the same target sequence (region 5), it is possible to adjust the stoichiometric ratios of different target sequences after SNAP purification. Fig. 5 shows an example in which 3 different target sequences are sought to be purified in a 1:2:6 ratio.

The number of available barcodes based on variable positions determines the range of available stoichiometric ratios and number of sequences possible. For example, a probe library with 12 variable positions and 2 possible nucleotides at each position contains $2^{12} = 4096$ sequence instances. The sum of all integer stoichiometry ratios among different target sequences must sum to 4096 (or less). For example, it is possible to purify a library of 2097 target sequences, in which 2096 target sequences are at equal stoichiometry to one another, and the last target sequence is present at 1000x excess.

Importantly, degenerate randomer sequences can also be incorporated in region 4 of the precursor sequences, in order to reduce the cost of precursor synthesis. For example, in Fig. 5, Target 2 occupies two sequence instances of the barcode through the use of a degenerate W in region 7.

5 In some instances, to yield uniform concentrations of target oligonucleotides in the final pool, the capture probe library should be at a significantly lower concentration than the input target oligonucleotide sample. For example, the full-length product of Target 1 is initially at 5 μ M and the full-length product of Target 2 is initially at 8 μ M, each member of the capture probe library should be kept below 5 μ M, such as 1 μ M. In such an instance, the
10 purification yield may be lower than for HPLC and PAGE methods for single targets but will provide a uniform final concentration of target molecules. In instances where a uniform final target concentration is not needed, the yield will not be reduced in such a way.

IX. Barcode Assignment Based on Target Sequence

To simultaneously achieve high sequence specificity and high hybridization yield, the
15 standard free energies of hybridization ($\Delta G^\circ\text{Hyb}$) between the different precursor and their respective matched probe sequence instances must be similar. Naive design of the validation region sequence (region 3) and assignment of barcodes (region 4) may result in precursor oligonucleotides with significant secondary structure between region 5 and regions 3 and 4, resulting in $\Delta G^\circ\text{Hyb}$ significantly more positive than expected, in turn leading to lower
20 capture yields. Consequently, it is suggested that the sequences of regions 3 and 4 be rationally designed given desired target sequences, so that similar secondary structure is observed for all precursor sequences.

X. Examples

The following examples are included to demonstrate preferred embodiments. It
25 should be appreciated by those of skill in the art that the techniques disclosed in the examples that follow represent techniques discovered by the inventors to function well in the practice of embodiments, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a
30 like or similar result without departing from the spirit and scope of the disclosure.

Example 1 - Stoichiometric SNAP Purification

Fig. 6 shows data demonstrating proof-of-concept of the SNAP purification technique. Denaturing polyacrylamide gel electrophoresis is used to visualize and quantitate the purity and concentration of different oligonucleotide species. Lane 1 shows a chemically synthesized precursor oligonucleotide, and lane 4 shows the corresponding chemically synthesized target oligonucleotide. Both the precursor and the target oligonucleotides were synthesized with a 3' FAM fluorophore functionalization to allow easy visualization. Lanes 2 shows the captured precursor molecules before USER enzyme treatment to remove regions 3 and 4, and Lane 3 shows the final purified product after USER enzyme treatment. The relative lack of truncation bands in Lanes 2 and 3 indicate that truncation products have been removed.

Lane 5 shows a mixture of 3 precursor oligonucleotides of different lengths (100nt, 90nt, and 80nt), prepared at a nominal stoichiometric ratio of 1:1:1. Lane 6 shows the output of the SNAP purification protocol. The stoichiometric ratio of the purified target oligonucleotides was quantitated to be 1.2:1:1.5. Lane 7 and 8 show a similar set of experiments, except the 3 precursor oligonucleotides were nominally prepared at 1:5:25 and the SNAP-purified products were observed to be at 1.2:1:1.7, and is closer to the designed 1:1:1 stoichiometric than the precursors.. Lane 9 and 10 show a similar set of experiments, except the 3 precursor oligonucleotides were nominally prepared at 5:25:1 and the SNAP-purified products were observed to be at 1.2:1:0.5.

Figs. 7A-B show data demonstrating purity and stoichiometry attainable by means of SNAP purification as measured by Next Generation Sequence and digital droplet PCR. Fig. 7A shows that the median of reads perfectly matching the desired sequence is about the 80%, in case of SNAP purification, which is greater than the median from PAGE purification which is about 60% and the one from unpurified oligos, which is about 55%. Fig 7B shows the concentration of the individual sequences as measured through digital droplet PCR, using target specific primers. After the SNAP purification the concentration of the oligo is within a factor of two for the 95% of the sequences of the pool.

Fig. 11 shows data demonstrating purity for a 256-plex, as measured by NGS. Also in this case the fraction of perfect reads for SNAP purified oligo is close the 80%.

Example 2 -Purification of Enzymatically-produced Precursors

Fig. 10 shows an example embodiment of purifying enzymatically produced precursors, in this case a transcribed RNA species. Chemical synthesis of RNA oligonucleotides is significantly (8-fold) more expensive than DNA synthesis, and limited to shorter lengths (typically ≤ 50 nt). In vitro transcribed RNA using RNA polymerase and a corresponding DNA template sequence can produce economically produce large quantity of desired RNA target sequences that is also significantly longer than chemically synthesized RNA, but requires labor- intensive Polyacrylamide Gel electrophoresis (PAGE) purification. The present SNAP purification methods can significantly reduce the labor needed to produce RNA molecules, especially in highly multiplexed settings.

Because enzymatically produced precursors disproportionately exhibit truncations and errors at the 3' end rather than the 5' end, the DNA template sequence is designed so that the validation and barcode regions (3 and 4, respectively) will be positioned at the 3' end of the transcript. The stoichiometric capture of full-length precursor RNA transcripts occurs similarly to that of DNA oligonucleotides described previously. An RNase H enzyme may be used to remove regions 6 and 7 from the precursor to leave only the desired target RNA sequence, because RNase H will selectively cleave RNA at regions where it is hybridized to DNA.

Example 3 - Probe Design Variations

Figs. 9A-D show a few possible variations in design of the probe and protector sequences. The relative ordering of the regions may be altered, as long as the complementarity relationship between the regions are preserved (*e.g.*, region 2 is complementary to region 3).

Fig. 9B In those instances when the number capture probes is higher than the number or target sequences, as well as in those case where multiple validation regions (region 2) are used simultaneously, the purpose of the protector oligonucleotide and of regions 7 and 8, is to ensure sequence-specific hybridization of each precursor to its matched probe oligonucleotide. Zhang, D. Y., Chen, S. X., & Yin, P. (2012). Optimizing the specificity of nucleic acid hybridization. *Nature chemistry*, 4(3), 208-214. and Wu, L. R., Wang, J. S., Fang, J. Z., Evans, E. R., Pinto, A., Pekker, I., ... & Zhang, D. Y. (2015). Continuously tunable nucleic acid hybridization probes. *Nature methods*, 12(12), 1191-1196. shows that the competitive

hybridization between precursors and the protectors is specific to even single- nucleotide changes in sequence when the sequences of the probe and protector are designed appropriately.

5 This specificity is useful for 2 purposes: First, it limits the off-target hybridization of precursors to non-cognate probe sequence instances that are not perfectly complementary. Second, it prevents the hybridization of imperfectly synthesized precursors that lack any nucleotide in regions 3 or 4.

10 Unless explicitly stated otherwise, "complementary" in this document refers to "partially or fully complementary". Two sequences are defined to be "partially complementary" when over 80% of the aligned nucleotides of one sequence is complementary to corresponding nucleotides of the other sequence.

15 Tables 1-4, below, shows a hypothetical set of sequences for the precursors, capture probe and protector that could be used in methods of the present disclosure. In the sequence of the capture probe, S or W indicates that all variants are included in the capture probe library. For example, both A and T would be present in the mixture of capture probes at each W as part of the randomer library of capture probes.

TABLE 1 – List of 64 precursor for the 64plex described in Figs. 7A-B

Name	Region 3	Region 4	Region 5
NGSPle x64_1	TAGGTGC GGTCGCA SEQ ID NO: 1	TGTGAC TCACTG SEQ ID NO:2	GCCCGTCGGCATGTATTAGCTCTAGAATTACCACAGTGCA ACCTTTCGAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 3
NGSPle x64_2	TAGGTGC GGTCGCA SEQ ID NO: 1	TGTCTG TGTCTG SEQ ID NO:4	GGGGCCGGAGAGGGGCTGACCGGGTTGGTTTTGATGCG GTGCTCGAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 5
NGSPle x64_3	TAGGTGC GGTCGCA SEQ ID NO: 1	TGTCAC ACTCTG SEQ ID NO:6	CCCTGATTCCCCGTCACCCGTGGTCACCATGGTAGTGGC CATAGCAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 7
NGSPle x64_4	TAGGTGC GGTCGCA SEQ ID NO: 1	TGTGAC ACTCTG SEQ ID NO:8	TTTTTCGTCACTACCTCCCCGGGTCGGGAGTGGGTGAATT ATGCTGAACGAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 9
NGSPle x64_5	TAGGTGC GGTCGCA SEQ ID NO: 1	TGTGAG AGACTG SEQ ID NO:10	GCCCGCCCCGCTCCCAAGATCCAACTACGAGCTTTTAGGT CAGTGGAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 11
NGSPle x64_6	TAGGTGC GGTCGCA SEQ ID NO: 1	TGTGTC TGTCTG SEQ ID NO:12	GGCCGTCCCTCTTAATCATGGCCTCAGTTCCGAAACCTAC ACATTATCTGAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 13
NGSPle x64_7	TAGGTGC GGTCGCA SEQ ID NO: 1	TGTGAG TCACTG SEQ ID NO:14	GGTATCTGATCGTCTTCAACCTCCGACTTTTCTGTTCTGGA CATGCCAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 15
NGSPle x64_8	TAGGTGC GGTCGCA SEQ ID NO: 1	TGTCTC AGACTG SEQ ID NO:16	TGGTGGTGCCCTTCCGTCAATTCCTTTAAGTTTCATGCTTC TACTCCTAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 17
NGSPle x64_9	TAGGTGC GGTCGCA SEQ ID NO: 1	TGTCAC TCACTG SEQ ID NO:18	CCTGTCCGTGTCCGGGCCGGGTGAGGTTTCCCGTGCACT AGGGCTGAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 19
NGSPle x64_10	TAGGTGC GGTCGCA SEQ ID NO: 1	TGTGAC TGTCTG SEQ ID NO:20	GTAAGTAGTTAGCATGCCAGAGTCTCGTTTCGTTATCGGAT GGCCTAGTATAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 21
NGSPle x64_11	TAGGTGC GGTCGCA SEQ ID NO: 1	TGTCAC TCTCTG SEQ ID NO:22	GCCCCGGACATCTAAGGGCATCACAGACCTGTTATTCCTT GTTGAAGAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 23
NGSPle x64_12	TAGGTGC GGTCGCA SEQ ID NO: 1	TGTCTC AGTCTG SEQ ID NO:24	GGTAGTAGCGACGGGCGGTGTGTACAAAGGGCAGGTGA GTATTTGATTCAAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 25
NGSPle x64_13	TAGGTGC GGTCGCA SEQ ID NO: 1	TGTCTC TGA CTG SEQ ID NO:26	GGCGCTGGGCTCTTCCCTGTTCACTCGCCGTTACTATGTT CGGCCTTTTATGCTCTCGTACGGTTAAGAGCC SEQ ID NO: 27
NGSPle x64_14	TAGGTGC GGTCGCA SEQ ID NO:	TGTGTG TGA CTG SEQ ID	TACCACCCGCTTTGGGCTGCATTCCCAAGCAACCCCCC GAAAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 29

NGSPle x64_15	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 28 TGTCTG ACACTG SEQ ID	CTTCCCTTACGGTACTTGTTGACTATCGGTCTCGTAAAC GGTTAGATCGAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 31
NGSPle x64_16	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 30 TGTCTC TCTCTG SEQ ID	GGCGGACTGCGCGGACCCACCCGTTTACCTCTTAGGTA TATAACGCCAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 33
NGSPle x64_17	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 32 TGTGAC TCTCTG SEQ ID	GGTGGAAATGCGCCCGGCGGCGGCGGTCGCCGGTACA CAGTTGAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 35
NGSPle x64_18	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 34 TGTGTG ACTCTG SEQ ID	CCTTCCCCGCCGGGCCTTCCAGCCGTCCCGGAGCAAGA TTGTTTACAGAAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 37
NGSPle x64_19	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 36 TGTCAG TGTCTG SEQ ID	GGGATTCGGCGAGTGCTGCTGCCGGGGGGGCTGTAGGA CAACGTACAACAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 39
NGSPle x64_20	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 38 TGTGAC TGA CTG SEQ ID	GCCGTGGGAGGGGTGGCCCGGCCCGCCCGACGAGGACTA CTCAAGAATTGCAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 41
NGSPle x64_21	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 40 TGTGAG ACTCTG SEQ ID	GCCGACCCCGTGCGCTCGCTCCGCCGTCCCCCTCTGCAC GCGGACAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 43
NGSPle x64_22	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 42 TGTGAG TGTCTG SEQ ID	GTGTTAGACTCCTTGGTCCGTGTTTCAAGACGGGTGGTCA TTTAGCGAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 45
NGSPle x64_23	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 44 TGTCTG TCACTG SEQ ID	CCAGGCATAGTTCACCATCTTTCGGGTCTTAACACGGAGC CCATTACAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 47
NGSPle x64_24	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 46 TGTGTG AGTCTG SEQ ID	GGGTGCGTCCGGTCTGCGAGAGCGCCAGCTATCCTATAA GCGCCGTCCAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 49
NGSPle x64_25	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 48 TGTGTC TCACTG SEQ ID	GTTCCGTTTCATCCCGCAGCGCCAGTTCTGCTTACCGTGC CACAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 51
NGSPle x64_26	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 50 TGTGTG AGACTG SEQ ID	GGATTCCGACTTCCATGGCCACCGTCCTGCTGTCTGAAAA AATTTCTGCAAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 53
NGSPle x64_27	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 52 TGTCAC AGACTG SEQ ID	ACGCTCCAGCGCCATCCATTTTCAGGGCTAGTTGACGCTA TGGCATCAAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 55
NGSPle x64_28	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 54 TGTGTC TCTCTG SEQ ID	GCAGCGGCCCTCCTACTCGTCGCGGCGTAGCGTCCCATT GAGCAGTTGAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 57
NGSPle x64_29	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 56 TGTCAG ACACTG SEQ ID	ACCCTTCTCCACTTCGGCCTTCAAAGTTCTCGTTTCTAGA GCCAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 59

NGSPle x64_30	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 58 TGTCAC ACACTG SEQ ID	ACTCTCCCCGGGGCTCCCCGCCGGCTTCTCCGGGATGTGA TGGGAGTACCAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 61
NGSPle x64_31	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 60 TGTCAC TGTCTG SEQ ID	GCCAGAGGCTGTTCACCTTGGAGACCTGCTGCGGACCGC TCACAAAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 63
NGSPle x64_32	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 62 TGTCTG AGACTG SEQ ID	CCCAGCCCTTAGAGCCAATCCTTATCCCGAAGTTATCAAT CAGTTGCAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 65
NGSPle x64_33	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 64 TGTGTC AGTCTG SEQ ID	GCTCCCCCGGGGAGGGGGGAGGACGGGGAGCGGGGTT GAGAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 67
NGSPle x64_34	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 66 TGTGAC AGTCTG SEQ ID	CCCCTGC CGCCCCGACCCTTCTCCCCCGCCGCCGTATCTAAGGTC CC GTCTCGTACGGTTAAGAGCC SEQ ID NO: 69
NGSPle x64_35	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 68 TGTCTG AGTCTG SEQ ID	GGCGGGGGGGACCGGCCCGCGGCCCTCCGCCGCCGT CATGTCCAAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 71
NGSPle x64_36	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 70 TGTGAG AGTCTG SEQ ID	GGATTCCCCTGGTCCGCACCAAGTTCTAAGTCGGCTAGGG AAGGCAAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 73
NGSPle x64_37	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 72 TGTCAG TCTCTG SEQ ID	GGCTACCTTAAGAGAGTCATAGTTACTCCCGCCGTGGCC CTTACCAAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 75
NGSPle x64_38	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 74 TGTGAG TGA CTG SEQ ID	CACCTCTCATGTCTCTTCACCGTGCCAGACTAGAGGCGG ATCCGAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 77
NGSPle x64_39	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 76 TGTCTC ACACTG SEQ ID	GCCCTCGGGGCTCGCCCCCGCCTCACCGGGTCCGG AACGCAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 79
NGSPle x64_40	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 78 TGTGTG ACACTG SEQ ID	GCCCTTCTGCTCCACGGGAGGTTTCTGTCCTCCCTAGTTT GCCAGACCAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 81
NGSPle x64_41	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 80 TGTCTC TGTCTG SEQ ID	GCTTGGCCGCCACAAGCCAGTTATCCCTGTGGTAATGATC TCTCTGAGTAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 83
NGSPle x64_42	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 82 TGTGTG TGTCTG SEQ ID	GCGGTTCTCTCGTACTGAGCAGGATTACCATGGCGGCC AAATATAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 85
NGSPle x64_43	1 TAGGTGC GGTCGCA SEQ ID NO:	NO: 84 TGTCTG TCTCTG SEQ ID	CCGAGGCTCCGCGGCGCTGCCGTATCGTTCCGCCTATGG AGGAGGACAGTCTCGTACGGTTAAGAGCC SEQ ID NO: 87
NGSPle x64_44	1 TAGGTGC GGTCGCA	NO: 86 TGTGAC ACACTG	AGGTCGTCTACGAATGGTTTAGCGCCAGGTTCCCCAGCC TCAGGCAGTCTCGTACGGTTAAGAGCC

	SEQ ID NO:	SEQ ID	SEQ ID NO: 89
NGSPle	1	NO: 88	
x64_45	TAGGTGC	TGTCAC	CATCTTTCCCTTGCGGTACTATATCTATTGCGCCAGCCTC
	GGTCGCA	TGACTG	CCCTCAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 91
	1	NO: 90	
NGSPle	TAGGTGC	TGTCAG	GACGGGTGTGCTCTTTTAGCTGTTCTTAGGTAGCTAGTAT
x64_46	GGTCGCA	AGTCTG	CTCGTCCCCAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 93
	1	NO: 92	
NGSPle	TAGGTGC	TGTGTC	GCTTTTAGGCCTACTATGGGTGTTAAATTTTTACTCTCTC
x64_47	GGTCGCA	ACACTG	TACAAGTCGAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 95
	1	NO: 94	
NGSPle	TAGGTGC	TGTCTG	AGGGTGATAGATTGGTCCAATTGGGTGTGAGGAGTTCAA
x64_48	GGTCGCA	TGACTG	CGTGTATTGTAAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 97
	1	NO: 96	
NGSPle	TAGGTGC	TGTGTC	GACTTGTTGGTTGATTGTAGATATTGGGCTGTTAATTGTCA
x64_49	GGTCGCA	ACTCTG	GTTTGTTGTAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 99
	1	NO: 98	
NGSPle	TAGGTGC	TGTCTG	GTAAGATTTGCCGAGTTCCTTTTACTTTTTTAAACCTTTCCT
x64_50	GGTCGCA	ACTCTG	TATGCCAAAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 101
	1	NO: 100	
NGSPle	TAGGTGC	TGTCAG	GCTGAACCCTCGTGGAGCCATTTCATACAGGTCCCTGTCC
x64_51	GGTCGCA	AGACTG	ACCGGCTAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 103
	1	NO: 102	
NGSPle	TAGGTGC	TGTCAG	GCTCGGAGGTTGGGTTCTGCTCCGAGGTCGCCCCACTTG
x64_52	GGTCGCA	TGACTG	CATCCTTTGGAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 105
	1	NO: 104	
NGSPle	TAGGTGC	TGTGTC	GGATTGCGCTGTTATCCCTAGGGTAACTTGTTCCGCAGAC
x64_53	GGTCGCA	AGACTG	GTTGTGCAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 107
	1	NO: 106	
NGSPle	TAGGTGC	TGTCAG	GCCTTATTTCTCTTGTCCTTTCGTACAGGGAGGAATTTGAA
x64_54	GGTCGCA	ACTCTG	TATCTGTTTAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 109
	1	NO: 108	
NGSPle	TAGGTGC	TGTCAG	TTTCCCGTGGGGGTGTGGCTAGGCTAAGCGTTTTGCATT
x64_55	GGTCGCA	TCACTG	CATAGACCTTAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 111
	1	NO: 110	
NGSPle	TAGGTGC	TGTCTC	CAGGTGAGTTTTAGCTTTATTGGGGAGGGGGTGATCTACT
x64_56	GGTCGCA	ACTCTG	GCATCGTTAGAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 113
	1	NO: 112	
NGSPle	TAGGTGC	TGTGTC	GGCTCGTAGTGTTCTGGCGAGCAGTTTTGTTGATTTGAGT
x64_57	GGTCGCA	TGACTG	CTAAGGAGTAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 115
	1	NO: 114	
NGSPle	TAGGTGC	TGTGTG	GTAAGGAGTAGTCTCGTACGGTTAAGAGCC
x64_58	GGTCGCA	TCACTG	TTACCTGCAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 117
	1	NO: 116	
NGSPle	TAGGTGC	TGTGAG	GTGACGGGCGGTGTGTACGCGCTTCAGGGCCCTGTACAA
x64_59	GGTCGCA	TCTCTG	TCCTGGTAAGAAGTCTCGTACGGTTAAGAGCC

	SEQ ID NO:	SEQ ID	
	1	NO: 118	SEQ ID NO: 119
NGSPle	TAGGTGC	TGTCAC	TCTTCATCGACGCACGAGCCGAGTGATCCACCGCTGTTTC
x64_60	GGTCGCA	AGTCTG	CCTTCCAAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 121
	1	NO: 120	
NGSPle	TAGGTGC	TGTGAC	GATCAATGTGTCCTGCAATTCACATTAATTCTCGCAGCTA
x64_61	GGTCGCA	AGACTG	GCGTTCAAAAAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 123
	1	NO: 122	
NGSPle	TAGGTGC	TGTGTG	GCTCAGACAGGCGTAGCCCCGGGAGGAACCCGGGGTCA
x64_62	GGTCGCA	TCTCTG	CGAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 125
	1	NO: 124	
NGSPle	TAGGTGC	TGTCTC	GCCTACAGCACCCGGTATTCCCAGGCGGTCTCCCAAGGA
x64_63	GGTCGCA	TCACTG	GCATATATAACAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 127
	1	NO: 126	
NGSPle	TAGGTGC	TGTGAG	GAGATCGGGCGCGTTCAGGGTGGTATGGCCGTAGAGCG
x64_64	GGTCGCA	ACACTG	TTATCCAGTCTCGTACGGTTAAGAGCC
	SEQ ID NO:	SEQ ID	SEQ ID NO: 129
	1	NO: 128	
Name	Region 1		Region 2
Capture			
Probe64			
plex	CAGWSWSWS		ACATGCGACCGCACCTA TTTTTTTTTT\Biotin
			(SEQ ID NO: 130)

Table 2 - List of 256 precursor for the 256plex of in Fig. 8

Name	Region 3	Region 4	Region 5
NGSPLEx 256_1	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	CTCAG ACAG	TCGCCGCGTAAACGACGCGGCGCGCGTGTGCCGCACT GGTCGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_2	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	GACAG ACAG	AGCCCGCGCGCACGCGCCCTGCGCCGCGCGCCCC CAGGTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_3	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	CTGTC TCAG	SEQ ID NO: 133 CTGCTCCCGGCTGGGCCACCGCCAAAGCAGCGGCCCC ACAGGAGCTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_4	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	GTCAC ACAG	SEQ ID NO: 134 CTCCTCCGGCCCCGCGCGCCCACTCCGCGCCCGGCCTG GCCGCTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_5	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	CAGAC TCAG	SEQ ID NO: 135 GCCGCCGCGCGCCGCGCCGCGCCGCCCCGCTGCCTT CTCAGCCTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_6	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	CAGTG TGAG	SEQ ID NO: 136 GCGGGTGGGCGGCCCCGCGTTCCTTAGCCGCGGCTCCG CGGCTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_7	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	GAGTC TCAG	SEQ ID NO: 137 CCATTCTGCCAGACCCCCGGCTATCCCGGTGGCCAGG CTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_8	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	GA CTC TCAG	SEQ ID NO: 138 GGCCTCGCGTGCCTGGACAGCCCCGCGGGCCAGCAAG CCTATGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_9	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	CTCTC TCAG	SEQ ID NO: 139 CCATCTCAGGGTGAGGGGCTTCGGCAGCCCCCTCATGCT GTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_10	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	CACAC AGTG	SEQ ID NO: 140 GCGACCAAAGGCCGCGCACGGCCTGGCCGCTCAGCG ACTCCCGGCTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_11	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	GTGTG TCAG	SEQ ID NO: 141 CGCGGCCTCAAAAGGCCTCCTAGGCCGCGGCGGGCAAA GCACTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_12	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	GTGTG AGAG	SEQ ID NO: 142 TACGCTCTCGCGCACCAGGTACGCCTGGTGTTTCTTTGT GGTTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_13	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	GAGAG TGTG	SEQ ID NO: 143 CTCTGCCCAATCCCGGCTCCGGGCGACCCGGGCCCCCTG GTCGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_14	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	CA CTC TGAG	SEQ ID NO: 144 TCTTGTAGGAGGCCCATTCCTCCACACGGGGCCACC CACCCCGTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_15	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	CTGAC TGTG	SEQ ID NO: 145 CAGCCCCCAAACCCGACTGGTCAAGGGGGACATCAAG TCCCCCTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_16	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	GTCTC ACTG	SEQ ID NO: 146 CGCAGCCGACCGCGCGGCGAGAGCAGGGGCGGGGCCG GCGCGGTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_17	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	CAGTC AGAG	SEQ ID NO: 147 GGCCCGCTCGGCAGGCCCAACTGGCCCTCCCCCTTGG CGGCGATGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx 256_18	TCGCGAAAT TCGGTTGT SEQ ID NO: 131	GAGTG TCAG	SEQ ID NO: 148 CGGGCCGAATGCCAGCCCGCCGAGCTCAGGGCAGCGG GGAGCTGGTTGGTGGTGCTCGCAGGCTCGGCA
NGSPLEx	TCGCGAAAT SEQ ID NO: 131	CACAG	SEQ ID NO: 149 ACCTCCGCTGCGTCTCTCCGCGCCGCCCGCGCTGCTCG

256_19	TCGGTTGT	TGAG	CCGTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 150
NGSPLEx	TCGCGAAAT	CTCAC	CGGCCAGGTCTCGGTCAGGGCCAGGGCCGCCGAGAG
256_20	TCGGTTGT	AGAG	CAGCAGGATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 151
NGSPLEx	TCGCGAAAT	GTGAG	CTGCCTCACGCATCACAGCACCCCCACCCGAGCGCGGG
256_21	TCGGTTGT	AGTG	CGGGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 152
NGSPLEx	TCGCGAAAT	CAGAC	GGCCACGCTGCCACCAGCAGCAGGCCCATGGGGTGGCA
256_22	TCGGTTGT	ACTG	GGTGGTGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 153
NGSPLEx	TCGCGAAAT	CTGTG	CCAGGGGTAGCCCCCTGGATTATGGTCTGACTCAGGACT
256_23	TCGGTTGT	TCAG	GGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 154
NGSPLEx	TCGCGAAAT	GTGTC	CGAAGTTGCCAGGGTGGCAGTGCAGCCCCGGGCTGAG
256_24	TCGGTTGT	ACTG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 155
NGSPLEx	TCGCGAAAT	GACAC	GTTACCTCCCCGCACACGGACTGTGTGGATGCGGCGGG
256_25	TCGGTTGT	TCTG	GTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 156
NGSPLEx	TCGCGAAAT	CAGAG	AGAGCTCATGTATGGGTTAATCCGACCATGAGCTCTGTG
256_26	TCGGTTGT	AGAG	GTGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 157
NGSPLEx	TCGCGAAAT	CTGAG	GGCCGGGCCACGGCCAGCATCCGGACCCGGGGCAGCG
256_27	TCGGTTGT	ACAG	GCGTGGTGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 158
NGSPLEx	TCGCGAAAT	GA CTC	GCCATTACGGCTTCCCCGGCCAATAGACGCCCGGCTGC
256_28	TCGGTTGT	ACAG	CCTTACATGGTGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 159
NGSPLEx	TCGCGAAAT	CTCTG	TTCAGCATTTCTGCTGAAATCTAGGGTGGAAATGCGTTCC
256_29	TCGGTTGT	ACAG	TAGTGTGGTGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 160
NGSPLEx	TCGCGAAAT	CTCAC	CTCAGCGGAAATCCGGCGATCTGGCCGGAAGTGCGGCA
256_30	TCGGTTGT	TCAG	CAC TCTGGTGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 161
NGSPLEx	TCGCGAAAT	GAGAC	TGGCAGGAAGCTGCAGCCTTTCTCAAGAGCAGCCAGGAT
256_31	TCGGTTGT	TCAG	CTCCTGCTGGTGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 162
NGSPLEx	TCGCGAAAT	GAGTC	GGTAGCGAGGAGAGCGGCTGAGGCTCAGTGCGCCTGC
256_32	TCGGTTGT	TCTG	GCGGCGCTGGTGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 163
NGSPLEx	TCGCGAAAT	GAGAG	GCCATGGCAGCTTTCATGGCGTCTGGGGTTTTACCCAC
256_33	TCGGTTGT	TCAG	TCATCTTTGGTGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 164
NGSPLEx	TCGCGAAAT	CA CTC	TGCATCTCAGGAGACGCTCGTAGCCCTCGCGCTTCTCC
256_34	TCGGTTGT	ACAG	TTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 165
NGSPLEx	TCGCGAAAT	CAGAG	CCGTTGGCCACTTGTGGCCATTCTACTCCCATGCCGGC
256_35	TCGGTTGT	ACTG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 166
NGSPLEx	TCGCGAAAT	CTCAG	AACTGGGTCTACGGCTTGGACTTTCCAACCCTGACAGA
256_36	TCGGTTGT	TGTG	CCCGCATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 167
NGSPLEx	TCGCGAAAT	GAGAC	GCCGCTGCTGCAGCAGCTGCCTTATCCACCCGGAGCTT
256_37	TCGGTTGT	TGAG	GTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 168
NGSPLEx	TCGCGAAAT	GTCTC	ACGCGGCCTCTCCCGGCCCTTCCGTTTGTAGTAGGAGCC
256_38	TCGGTTGT	TGAG	GCACTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 169
NGSPLEx	TCGCGAAAT	CTCTG	TGAGGGGCTTGGGCAGACCCTCATGCTGCACATGGCAG

256_39	TCGGTTGT	AGAG	GTGTATTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 170
NGSPLEx	TCGCGAAAT	GTGTG	CCCGGCCCAACAACTACCTACGTCCGGGAGTCGCCAA
256_40	TCGGTTGT	TGAG	CCGACGTGGTCCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 171
NGSPLEx	TCGCGAAAT	GAGAC	AGGCAGACTGCTGCAGGACGGGACTGGGCCGGGAACC
256_41	TCGGTTGT	AGAG	GGCTGGTCCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 172
NGSPLEx	TCGCGAAAT	CACTG	TCTAAACCGTTTATTTCTCCCCACCAGAAGGTTGGGGTG
256_42	TCGGTTGT	TGAG	GTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 173
NGSPLEx	TCGCGAAAT	GAGTG	AACACTGCCTTCTTGGCCTTTAAAGCCTTCGCTTTGGCTT
256_43	TCGGTTGT	TGTG	CATGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 174
NGSPLEx	TCGCGAAAT	GTCAC	CGGTAGCCGAAGGAGTTCAAAAGACCTCTAGTGCGCCCA
256_44	TCGGTTGT	TCTG	CCGCATGGTCCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 175
NGSPLEx	TCGCGAAAT	CTCTC	GGTGACAGGGTGGCCCAGGAGCGGCCACTGAGATGAGA
256_45	TCGGTTGT	TCTG	CCTTGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 176
NGSPLEx	TCGCGAAAT	GAGTG	GATCACTCCCCAGGCGCTGAGGACGATGCCGCAGGCGG
256_46	TCGGTTGT	TCTG	CTGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 177
NGSPLEx	TCGCGAAAT	GAGAG	CTTAAGACCAGTCAGTGTTGCTCCTACCCATTCACTGG
256_47	TCGGTTGT	TGAG	CCTGAGCTGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 178
NGSPLEx	TCGCGAAAT	GTCTG	GAAGCTCCTAGAAGCTTCACAAGTTGGGGCACAACCTCCT
256_48	TCGGTTGT	TCTG	GTTGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 179
NGSPLEx	TCGCGAAAT	CACTC	TGGCGCGCGGCACTGGGAGCCGCCGGGCCGAGCCTGT
256_49	TCGGTTGT	AGTG	CAATGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 180
NGSPLEx	TCGCGAAAT	GAGTC	TCATAGAAAGAGAGGGAAGTTTTGGCGATCACAACAGCG
256_50	TCGGTTGT	ACAG	CCAAATGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 181
NGSPLEx	TCGCGAAAT	CACTC	CAAAGAATGCAAACATCATGTTTGAGCCCTGGGGATCAG
256_51	TCGGTTGT	AGAG	GGTGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 182
NGSPLEx	TCGCGAAAT	GTCTG	GATAGCGCTCCTGTCTATTGGCTGCGCCATCGCCCGTCA
256_52	TCGGTTGT	AGTG	GACTGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 183
NGSPLEx	TCGCGAAAT	CTCAG	CATATGCAGGTCCCCTGTTGGCCATTCCAATGGGTGGCG
256_53	TCGGTTGT	TGAG	GTGGCTGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 184
NGSPLEx	TCGCGAAAT	CTCAG	ACTCCCTGCTCCTTGGAATACGGACCACGCAGTCTATA
256_54	TCGGTTGT	TCTG	ATGCCCTGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 185
NGSPLEx	TCGCGAAAT	CACAC	CGGGGTAACCGTGGAGGGCGACGCGCAGAGGCTGCGG
256_55	TCGGTTGT	TGTG	CTATTTATGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 186
NGSPLEx	TCGCGAAAT	CAGTC	GGACTGGTCCCATGAGGCAGAAGGAGCACCAGCGCCTG
256_56	TCGGTTGT	TGTG	CTGGGTGGTGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 187
NGSPLEx	TCGCGAAAT	GTCTC	CGAGAAACAGCGCCCGACACCTGGCCCTTCGCAGCTCT
256_57	TCGGTTGT	TGTG	CGCCTTGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 188
NGSPLEx	TCGCGAAAT	CAGAG	CTGCATGGCCTTCATGACATGAAGGTTGGGCACATTCTT
256_58	TCGGTTGT	TGTG	GTCTGGTCCGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 189
NGSPLEx	TCGCGAAAT	CTGAC	CCGAAACCCATGGTGTCCGGCTGTATCCGAGAGCTGGGG

256_59	TCGGTTGT	TGAG	AGCAGCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 190
NGSPLEx	TCGCGAAAT	GACAC	TTCATGCAGATCACCTGCACCCCGCTTGTGTTCAAGTGGG
256_60	TCGGTTGT	TCAG	GTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 191
NGSPLEx	TCGCGAAAT	GAGAC	AGCTGCCGGGGTCCGGTTCCTCAGCTCCAGGTGGATCC
256_61	TCGGTTGT	ACAG	TTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 192
NGSPLEx	TCGCGAAAT	CTCTG	CTCGGAAGTAGCCCCCGTAGGTGCCCTGCTTGTGGTCAA
256_62	TCGGTTGT	TGAG	ACTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 193
NGSPLEx	TCGCGAAAT	CACAC	CGGGGGTAGCGGTCAATTCCAGCCACCAGAGCATGGCT
256_63	TCGGTTGT	AGAG	GTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 194
NGSPLEx	TCGCGAAAT	GTCAC	GAAACATGATCGCTTATAAGCCAGCGGTCCCAATTCGGT
256_64	TCGGTTGT	ACTG	CCACCGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 195
NGSPLEx	TCGCGAAAT	GTGTG	CCCACACGTCCATGACTGGTCGTCCTAGATTTTAGGTGT
256_65	TCGGTTGT	ACAG	CTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 196
NGSPLEx	TCGCGAAAT	CACAC	CTTTAGCTCGAGATTGTCCCTCTCTGTCCAGCAGATAGG
256_66	TCGGTTGT	ACAG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 197
NGSPLEx	TCGCGAAAT	CTCTC	GCCGTCCTGCGCAAGCGCTTTTCAACCCCACTCCTTTCT
256_67	TCGGTTGT	AGAG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 198
NGSPLEx	TCGCGAAAT	GA CTG	CAGTCTCTGGGAGAATGGGCAGTTCCCAATCTTGGCCCC
256_68	TCGGTTGT	TGTG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 199
NGSPLEx	TCGCGAAAT	GTCTG	GGTTGGTGCTTGCCACACTTCTTACAGAAAGTCCGGCGG
256_69	TCGGTTGT	TGTG	GTTTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 200
NGSPLEx	TCGCGAAAT	GTGTC	CATGTAGTTGAGGTCAATGAAGGGGTCAATTGATGGCAAC
256_70	TCGGTTGT	ACAG	AATATTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 201
NGSPLEx	TCGCGAAAT	CTCAC	CGGAACTGGAGGTTTTCTTTTCCGCCATAGTTTGTCTG
256_71	TCGGTTGT	TGTG	GTGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 202
NGSPLEx	TCGCGAAAT	GTGTC	TTGTAGTCTGAGAGAGTGCGGCCATCCTCCAGCTGTTTG
256_72	TCGGTTGT	TCAG	CCGGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 203
NGSPLEx	TCGCGAAAT	GTCAG	CTGCAGTGGCTTTAAACCCACAGTAGTAACCTGCAGGAT
256_73	TCGGTTGT	TGTG	CACACTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 204
NGSPLEx	TCGCGAAAT	CA CTG	CGAAGGACAGGTGGTCTCTTCGTTGGGACGTCCCTTTTG
256_74	TCGGTTGT	TCTG	CCAGCATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 205
NGSPLEx	TCGCGAAAT	CAGTG	AGCTTGGCTCCCTTCTTGCGGCCCAAGGGCAGCGCATG
256_75	TCGGTTGT	AGAG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 206
NGSPLEx	TCGCGAAAT	GTCTG	TTCCGACATGTCCGCATTTTTGATCACGGCCTTTTCGGTG
256_76	TCGGTTGT	TGAG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 207
NGSPLEx	TCGCGAAAT	CAGAG	CTTGGGAAGACCAAGTCCTCAAGGATGGCATCGTGCACA
256_77	TCGGTTGT	TCAG	GCTGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 208
NGSPLEx	TCGCGAAAT	CTCTG	CGGCCGCTCCAGGAACGCCGACCACTCCACTTTAGGT
256_78	TCGGTTGT	TCAG	ATCATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 209
NGSPLEx	TCGCGAAAT	GTCTC	AGGCGAAGTTCCGTCTACGGCTATTTAATGGAGCGCCTG

256_79	TCGGTTGT	TCTG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 210
NGSPLEx	TCGCGAAAT	CTCTC	TGTATGTTCCATCCATGTGAGCAGCAAATGTGTATTTCCC
256_80	TCGGTTGT	TGAG	ACTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 211
NGSPLEx	TCGCGAAAT	CAGTC	GGTCTCATCCGAACCCTGCGGATATATTTTTACCCCAAGA
256_81	TCGGTTGT	TGAG	AATTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 212
NGSPLEx	TCGCGAAAT	GACAC	GGCCTTCACGCGGCCAGGAGTTTCTTATTGTTGCGGCA
256_82	TCGGTTGT	AGAG	GTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 213
NGSPLEx	TCGCGAAAT	GTGAG	CCTTTTCCAAGGATTTTACGTTGCGGCTTGTTAGGGTGAT
256_83	TCGGTTGT	ACAG	TTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 214
NGSPLEx	TCGCGAAAT	GTGTG	ACTGTTCTCTCTTGGCAAAGTAATCAGGATACATTGCCTG
256_84	TCGGTTGT	AGAG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 215
NGSPLEx	TCGCGAAAT	GTCAC	ACAGTAGCATGCAGTCCCACAACCTTGTAACAGCATCCCC
256_85	TCGGTTGT	TGTG	AGCGTCTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 216
NGSPLEx	TCGCGAAAT	GTCAG	ACTTGGCTCCAGCATGTTGTCACCATTTCCAACCAGAAATT
256_86	TCGGTTGT	TCAG	GGCATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 217
NGSPLEx	TCGCGAAAT	GA CTC	ACAATGCAAAGATGGCTTTTCAGAGCAGCCAGTG GGGGT
256_87	TCGGTTGT	AGAG	GGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 218
NGSPLEx	TCGCGAAAT	CTGAG	CGACGTAGCCCGGCCCTCTTCGACCTGCACCTCCGCGGC
256_88	TCGGTTGT	ACTG	TCCCTCTGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 219
NGSPLEx	TCGCGAAAT	CTCTC	AGTGGAACAGGATTACTATGATACAAAACCTTCCACTACT
256_89	TCGGTTGT	ACTG	GGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 220
NGSPLEx	TCGCGAAAT	GTGAC	GCAAAGGCAATCTTCAAATAGAAGCTGGCAACACAAGAC
256_90	TCGGTTGT	ACAG	CTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 221
NGSPLEx	TCGCGAAAT	GTGTG	AATCACGCACTGTCCCCAACAGCCCCAGTTAACACAGGG
256_91	TCGGTTGT	TGTG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 222
NGSPLEx	TCGCGAAAT	CTCAG	CAGGGTTTCTGGTCCAAATAGGCTTGGTCTTGTATGGT
256_92	TCGGTTGT	ACTG	CGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 223
NGSPLEx	TCGCGAAAT	GTGTC	CCCGAATCCGCGGCCCTTCTACCAAGAACATTCTGTT
256_93	TCGGTTGT	TCTG	GGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 224
NGSPLEx	TCGCGAAAT	GTCTC	GGAGATCCATCATCTCTCCCTTCAATTTGTCTTCGATGAC
256_94	TCGGTTGT	TCAG	ATTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 225
NGSPLEx	TCGCGAAAT	GA CTC	TGGCATTAGCAGTAGGTTCTTGTATTTGAGTCTGCTTGGT
256_95	TCGGTTGT	ACTG	CGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 226
NGSPLEx	TCGCGAAAT	CTCAC	GAAAACTGGTCAGATGAATATTATTGCTTCCCATTTTCAA
256_96	TCGGTTGT	AGTG	CCAGTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 227
NGSPLEx	TCGCGAAAT	GA CTC	TCCAGAGGGTCCGGATCGCTCTCTTCTGCACTGAGGTTG
256_97	TCGGTTGT	TGTG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 228
NGSPLEx	TCGCGAAAT	CTCTG	CCATCCTGGCAGGCGGCTGTGGTGGTTTGAAGAGTTTG
256_98	TCGGTTGT	ACTG	GACTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 229
NGSPLEx	TCGCGAAAT	GTCAG	GCTGGCAAGGCTGAGCAATTCATGTTTATCTGCAACAGC

256_99	TCGGTTGT	TGAG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 230
NGSPLEx	TCGCGAAAT	GAGAG	AGCATCAGCTACTGCCAGCGGTTTCATGGGCTTCTTTTAC
256_100	TCGGTTGT	ACTG	TATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 231
NGSPLEx	TCGCGAAAT	CAGTC	TGAGTGAGCCCTCCTGCCACGTCTCCACGGTCACCACT
256_101	TCGGTTGT	TCTG	CCTCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 232
NGSPLEx	TCGCGAAAT	GTGAC	CTTTGGGTCCCAAGGTGCTCTTTACCAAGTCTCCAATGG
256_102	TCGGTTGT	AGAG	CGATTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 233
NGSPLEx	TCGCGAAAT	CAGTG	AACCCAATTAGTTCCCAGAAGTCACAACCTCAGCTCATGG
256_103	TCGGTTGT	ACTG	CCACCTGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 234
NGSPLEx	TCGCGAAAT	CACAC	CTCCTTGGTTCCATCTCCCGTGGCATCGCTTCCCTCTCG
256_104	TCGGTTGT	TGAG	GTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 235
NGSPLEx	TCGCGAAAT	GAGTG	TATTTACCAGGCCGGCAAAGAATGGACGGTCCTTCAGG
256_105	TCGGTTGT	AGAG	TCAACGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 236
NGSPLEx	TCGCGAAAT	CTGTG	CCCTCAGAGGACAGGGCGCGGTTGCTGGGTCATGAGCA
256_106	TCGGTTGT	TCTG	CCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 237
NGSPLEx	TCGCGAAAT	CTGTC	CAGGCACCAGACCAAAGACCTCCTGCCCCACAGCAAATG
256_107	TCGGTTGT	TGTG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 238
NGSPLEx	TCGCGAAAT	CAGTG	AGTTTCCTCTTCACTCAGCAGCATGTTGGGGATCCCGCG
256_108	TCGGTTGT	AGTG	GTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 239
NGSPLEx	TCGCGAAAT	GTCAC	TCTGTGAGAAAACCTTGAGAAATCAATAATGGTGGATTCA
256_109	TCGGTTGT	TGAG	TTGATTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 240
NGSPLEx	TCGCGAAAT	GAGAC	CTGTATCTGGTCCTGGCGGCCGGCTGTGATGTTTGACA
256_110	TCGGTTGT	ACTG	TTGTCCATGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 241
NGSPLEx	TCGCGAAAT	CACTC	GATGGCGGCGGGGGGCGAGGGGGCGCACGTAGCCTGGC
256_111	TCGGTTGT	TCTG	CATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 242
NGSPLEx	TCGCGAAAT	GA CTG	CTCTCTCTTCAGCAATGGTGAGGCGGATACCCCTTTCCTC
256_112	TCGGTTGT	TGAG	GGGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 243
NGSPLEx	TCGCGAAAT	GTGTG	TCTGGGACAAGACAGTCGAGGGAGCTTCTTCCTCAGGGA
256_113	TCGGTTGT	ACTG	ACTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 244
NGSPLEx	TCGCGAAAT	CAGAG	GCTGCCAAAGCTGGGTCCATGACAACCTTCTGGTGGGGC
256_114	TCGGTTGT	AGTG	GTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 245
NGSPLEx	TCGCGAAAT	GTCAG	AGACTGCCAGCGAAGCCCCCTTATGAGCAAAAGAGCAA
256_115	TCGGTTGT	TCTG	CCCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 246
NGSPLEx	TCGCGAAAT	CACTC	CTTCAGGTGTCCTTGAAGCAATAATTTCTGTCACTACTTT
256_116	TCGGTTGT	TCAG	TTCAATTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 247
NGSPLEx	TCGCGAAAT	CTGTC	AACGATCAA AATTAGACATGTCTTCATCTGAATCATCTTC
256_117	TCGGTTGT	ACAG	CCAGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 248
NGSPLEx	TCGCGAAAT	GTGAG	CGGCCACACCATCTTTGTCAGCAGTCACATTGCCCAAGT
256_118	TCGGTTGT	TGTG	CTCCAACCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 249
NGSPLEx	TCGCGAAAT	GAGTC	TGACCTCTCACTTTTCCAGCACGGGCCAGGGAACCATGG

256_119	TCGGTTGT	AGTG	ACTTTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 250
NGSPLEx	TCGCGAAAT	GTGAC	GTCAGCATAATCTTTATTTCAAATAACATTTTATTATGG
256_120	TCGGTTGT	TGAG	TCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 252
NGSPLEx	TCGCGAAAT	CTCAC	AGCCACAGGATGTTCTCGTCACACTTTTCCATGTAGGCG
256_121	TCGGTTGT	ACAG	TTATTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 253
NGSPLEx	TCGCGAAAT	GTCTG	TATCCGTTCCCTTACATTGAACCATTTTACTGTTCCCAAAAC
256_122	TCGGTTGT	ACAG	CTTCGTGGTTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 254
NGSPLEx	TCGCGAAAT	CTGTC	GACTTTTACAATCGATTCCCCAAACCCCTTTATGGCAGCA
256_123	TCGGTTGT	AGTG	ACACTGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 255
NGSPLEx	TCGCGAAAT	GACAC	TTCTTGTAATTTGCATAATCCTCAAGAATGGAATCCACTG
256_124	TCGGTTGT	TGAG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 256
NGSPLEx	TCGCGAAAT	GAGAG	TGGTCCCAGGGGAAAGGAAGAGGCCAGTTGGTCCAGTT
256_125	TCGGTTGT	AGAG	TTGATTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 257
NGSPLEx	TCGCGAAAT	GACTG	CAGGGGACGGTACTCCACATCCTCTCTGAGCAGGCGGT
256_126	TCGGTTGT	ACTG	GGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 258
NGSPLEx	TCGCGAAAT	CTCTG	CTTCTACATCATCAGCTGCCATACGAAGAAGGGACTCCG
256_127	TCGGTTGT	AGTG	TTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 259
NGSPLEx	TCGCGAAAT	GACTG	GCCACCAGCATCAACCTTCTTGGCTTCGGGTTTCTTCTG
256_128	TCGGTTGT	ACAG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 260
NGSPLEx	TCGCGAAAT	GTCTG	CCCCACCGGTGCTCTTGGTACGAAGATCCATGCTAAATT
256_129	TCGGTTGT	AGAG	CCCCATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 261
NGSPLEx	TCGCGAAAT	CAGAG	TTGTATATAAGATTACTTTATTCCTGCATCTTCTCAATGGT
256_130	TCGGTTGT	TGAG	TTCTTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 262
NGSPLEx	TCGCGAAAT	GACAG	CATTTTCATGGTTTTGTAGAGCTTCAATTTTGTCTAAGCCT
256_131	TCGGTTGT	ACTG	CCATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 263
NGSPLEx	TCGCGAAAT	GTGAC	CATAGTTGTCAACAAGCACAGTGAAGCGCCATTCTCTTT
256_132	TCGGTTGT	TGTG	ACATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 264
NGSPLEx	TCGCGAAAT	GTCTC	GCCAACAGCATGCTGGGTAACATTGTAGACTCTTCCTGG
256_133	TCGGTTGT	AGTG	TCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 265
NGSPLEx	TCGCGAAAT	CTGAC	AGGAGATCTCCACAGGGGCTGGACGGTTCATTATGGCAA
256_134	TCGGTTGT	AGAG	ATTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 266
NGSPLEx	TCGCGAAAT	CTGAG	TTGAGCATCTCGTAGTTGGGAGGCTGGCCGCTGTTGACA
256_135	TCGGTTGT	TCAG	GGAGAGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 267
NGSPLEx	TCGCGAAAT	CAGAC	CTCCCTTTCCCAGTAGTTTCGGTTTCTCAACAGTTTCCT
256_136	TCGGTTGT	TGTG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 268
NGSPLEx	TCGCGAAAT	GTGAG	TGGTTTTTAACAGGTTTAACCAATCATCTACTATCTGATTG
256_137	TCGGTTGT	ACTG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 269
NGSPLEx	TCGCGAAAT	CTGAC	CCTTCTGTCCTCATGTTGGCAGAGATATCTACTCTGTGGT
256_138	TCGGTTGT	AGTG	CGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 270
NGSPLEx	TCGCGAAAT	GACAG	GGATTCCAGTAGCCAGGTTGGTACGGGACGGCATCATAA

256_139	TCGGTTGT	AGAG	CATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 271
NGSPLEx	TCGCGAAAT	CTGAG	CTTCAGCGGAGGCATTTCCACCAATGAGCGAGTCATTGG
256_140	TCGGTTGT	AGTG	TCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 272
NGSPLEx	TCGCGAAAT	GTCAG	GTTTATGGTAAAGCTTAGCCTTCAGACCAATCATTTTCTTT
256_141	TCGGTTGT	ACAG	GCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 273
NGSPLEx	TCGCGAAAT	CTCAC	GGCCGCAAAAGGGAAGAGAACTACACGCTGCTTCCGGT
256_142	TCGGTTGT	TCTG	TCTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 274
NGSPLEx	TCGCGAAAT	GAGTG	TTGTTCACTGGGTCTTTGTCTTTCTTGCCGACTTTCCAG
256_143	TCGGTTGT	TGAG	CGTCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 275
NGSPLEx	TCGCGAAAT	GAGAC	CTTCCAGTTAAGGCTCTTTATTTTATTTTGAACACTTTTT
256_144	TCGGTTGT	AGTG	TGCATTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 276
NGSPLEx	TCGCGAAAT	CAGAC	ATCAACAAGCCACGGTTTTAGCTCTTCAGGAATCTTTACT
256_145	TCGGTTGT	ACAG	TTAACTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 277
NGSPLEx	TCGCGAAAT	CACAC	GGTGGTTCCTTGAGGGCTTTGATGATCAGGGCAGAGGC
256_146	TCGGTTGT	TCTG	AGAAGGCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 278
NGSPLEx	TCGCGAAAT	GTGTC	AATTCACACACCTCACAGTAAACATCAGACTTTGCTGGGA
256_147	TCGGTTGT	AGTG	CCTCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 279
NGSPLEx	TCGCGAAAT	GTCTG	TGTCATCCTTCTTGCCACCTCCAGGACCATGACCACCAC
256_148	TCGGTTGT	ACTG	TCTGACTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 280
NGSPLEx	TCGCGAAAT	GTGAG	GAGCAAGGAGGGCTGGAAGCTGTTAGTCAGAGTGTTGA
256_149	TCGGTTGT	TCTG	AGCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 281
NGSPLEx	TCGCGAAAT	CACAC	AAAATTGTGCGGATGTGGCTTCTGGAAGACCTTCATTCTA
256_150	TCGGTTGT	TCAG	AAGCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 282
NGSPLEx	TCGCGAAAT	CACTG	GCAGTTTTCTAATTGAGAATGTAATCTTGGTCTTTAAAGA
256_151	TCGGTTGT	TGTG	ACATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 283
NGSPLEx	TCGCGAAAT	CACAC	GTTTCTGCATCAGCCCGCTCATCAAATCCAGGGAAGTTG
256_152	TCGGTTGT	ACTG	GTCCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 284
NGSPLEx	TCGCGAAAT	CTGAG	CCAGCGGCAACCTCAGCCAAGTAACGGTAGTAATCTCCT
256_153	TCGGTTGT	TCTG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 285
NGSPLEx	TCGCGAAAT	GACAC	GTGATCGGGGTTTCTTGATACCATTCTGTGCCATTTTCG
256_154	TCGGTTGT	TGTG	GGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 286
NGSPLEx	TCGCGAAAT	GTCAC	AGGAGGTCCTGCTGAGTTGGTGAATCTCTGGTAACGGTG
256_155	TCGGTTGT	AGAG	GTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 287
NGSPLEx	TCGCGAAAT	CTGAG	CCTGGTTTTCTAAAATTCTTCAGGTCAATAGTCAAGCCTT
256_156	TCGGTTGT	TGAG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 288
NGSPLEx	TCGCGAAAT	CTGAC	ACAATATCACCTTTCTTATAGATTGCGATATATGTGGCCA
256_157	TCGGTTGT	ACAG	AAGGATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 289
NGSPLEx	TCGCGAAAT	CTCTG	AAACAAAACAAGAAAAAGTAATCTGCTAAAACTATAGGG
256_158	TCGGTTGT	TGTG	TCCCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 290
NGSPLEx	TCGCGAAAT	CAGTG	TAGAATCTTTTTATTTCAGAAAAAAAACCCCAAAAAACA

256_159	TCGGTTGT	ACAG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 291
NGSPLEx	TCGCGAAAT	CTGTG	AGTTTAATAAAATACAAATACTCGTTTTCTTTTTGATTAGTGT
256_160	TCGGTTGT	AGAG	GGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 292
NGSPLEx	TCGCGAAAT	GTGAC	ACTTTGAGATTCTTTTCTTTTGCGCCTCTTATCAAGTCAG
256_161	TCGGTTGT	TCAG	CTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 293
NGSPLEx	TCGCGAAAT	CACAG	AGCCTGGTTGGAGGATTCTAGTTTTATACATGAGAAATA
256_162	TCGGTTGT	AGAG	GTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 294
NGSPLEx	TCGCGAAAT	GAGTC	GTTCCCAAGATAGAAGAGTAGGTATGAAGCAATTCTGAC
256_163	TCGGTTGT	AGAG	TCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 295
NGSPLEx	TCGCGAAAT	GA CTC	TCTTTGTATGAGTCTTCATTTCAGTGTATCAAGTTCATGGT
256_164	TCGGTTGT	AGTG	CGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 296
NGSPLEx	TCGCGAAAT	CTGTG	ACAGATGAATGTAGGATTGATGCAAGTCACTTCCAGGAA
256_165	TCGGTTGT	ACTG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 297
NGSPLEx	TCGCGAAAT	GTGAG	TCGCTTTAGTCTCCTCGAGTTTCTTCTGCTCCTCTTTTTGT
256_166	TCGGTTGT	TGAG	TTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 298
NGSPLEx	TCGCGAAAT	CTCAG	TACTTCTGGGCGCTCACAGGGGAGGGCAGGTGGATGGT
256_167	TCGGTTGT	AGTG	GATCATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 299
NGSPLEx	TCGCGAAAT	CA CTC	TGGTGCCGGATGAAC TTCTTGTTCTCTTTTTGACGATCT
256_168	TCGGTTGT	ACTG	TGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 300
NGSPLEx	TCGCGAAAT	GAGTC	TATGCC TAGA ACTTTACGCCAATTATTTACCTCTTGCA
256_169	TCGGTTGT	TGAG	CATATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 301
NGSPLEx	TCGCGAAAT	CACAG	CCCAGGTCCTGTGATGTTTATTGAAGGAAGCAAGGGCAG
256_170	TCGGTTGT	TCTG	GGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 302
NGSPLEx	TCGCGAAAT	CAGAC	GTCCCTTTGTTTTCTTCTTTTCCCCACTCTAGTTGGT
256_171	TCGGTTGT	AGTG	ACAGGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 303
NGSPLEx	TCGCGAAAT	CAGAC	CCATCGATGTTGGTGTTGAGTACTCGCAAATATGCTGG
256_172	TCGGTTGT	TCTG	TCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 304
NGSPLEx	TCGCGAAAT	GTGTG	GATACCACAGAATCAGCAGGGTGAGAAACAATTGCACAA
256_173	TCGGTTGT	TCTG	AAGACTTGCTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 305
NGSPLEx	TCGCGAAAT	GA CTC	AAAGCTTGAGATCACTTGAGGCCAGAGTTTTTCAGACCTG
256_174	TCGGTTGT	TCTG	GTGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 306
NGSPLEx	TCGCGAAAT	GACAG	CGGGTGTTGACGGGCGGCGGATCGGCAAAGGCGAGGC
256_175	TCGGTTGT	TCTG	TCTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 307
NGSPLEx	TCGCGAAAT	CACAG	CAGCGCAGTGCAGAGCATGGTGGTAGATGTGGCAGAG
256_176	TCGGTTGT	ACAG	GATGGCATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 308
NGSPLEx	TCGCGAAAT	CTGAG	GACCCTCATAGACAGCAGACAGAAGAGGAGTAATATGAT
256_177	TCGGTTGT	TGTG	GTTTATTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 309
NGSPLEx	TCGCGAAAT	GTGTG	AAATACACTTTTAATTGATTTTCAATAAAAACTACTTGGTC
256_178	TCGGTTGT	AGTG	GGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 310
NGSPLEx	TCGCGAAAT	GTGAG	GCCGCGGAAGCCGGAGAGGAGAAGAAGAGAAGGAGG

256_179	TCGGTTGT	TCAG	GTTAGGCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 311
NGSPLEx	TCGCGAAAT	CAGTC	CAAATCTCAGGGAAGCAGTGATGGAGGACACAATCTGGC
256_180	TCGGTTGT	ACTG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 312
NGSPLEx	TCGCGAAAT	CACTC	GGTCATAGTGAGGGTAAGAGCTTTTACATCCCGCAGTG
256_181	TCGGTTGT	TGTG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 313
NGSPLEx	TCGCGAAAT	CACAG	GATCCATCATTTCTCCTTTAAGCTTATCTTCCAAAATGGTC
256_182	TCGGTTGT	TGTG	GGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 314
NGSPLEx	TCGCGAAAT	GTCTG	TGTTTACTGATTTCTGTCTGGTTAAACATCCAATACTGGT
256_183	TCGGTTGT	TCAG	CGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 315
NGSPLEx	TCGCGAAAT	GACAG	GCCTATCTCTTTCCATCAGACTCCAGTGATACCCAATGGT
256_184	TCGGTTGT	TCAG	CGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 316
NGSPLEx	TCGCGAAAT	GACAG	CCTGTAATCTCAGCACGTTGGGAGGCGAGGTGGGTGGA
256_185	TCGGTTGT	AGTG	TGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 317
NGSPLEx	TCGCGAAAT	CTCTC	CTTAAATACCAGATACATTTTTAGTCCTCTACATAATGGTC
256_186	TCGGTTGT	ACAG	GGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 318
NGSPLEx	TCGCGAAAT	GACAC	GTTTTTGGAAAGATTGCGGTTTCAGCACAGGATTCCATTTG
256_187	TCGGTTGT	AGTG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 319
NGSPLEx	TCGCGAAAT	GTCTC	TGATATCCTTGTTTTTAACTGTTGTGGCTTGCTGAATCAA
256_188	TCGGTTGT	AGAG	ATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 320
NGSPLEx	TCGCGAAAT	CAGAG	AAGACGGAGTAGTTAAGAGCCAGGCCTAATCGGATGGTG
256_189	TCGGTTGT	ACAG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 321
NGSPLEx	TCGCGAAAT	CTGTC	CGAGTTCAGAGACAATATCAAAATTACCCTCCTTTTGGT
256_190	TCGGTTGT	ACTG	CGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 322
NGSPLEx	TCGCGAAAT	CTCTG	GGTGAGAGAACTAATAGCAACCAGGCAACTGAGGACGAA
256_191	TCGGTTGT	TCTG	GTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 323
NGSPLEx	TCGCGAAAT	CTCAC	TTGGGGTGCTTTATCTTCTTTGAGTTTTCGCACAAAGATGG
256_192	TCGGTTGT	ACTG	TCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 324
NGSPLEx	TCGCGAAAT	CTGTG	GGGCATGGGCTCACATTCATTTCCTTTATAACTCCATCCT
256_193	TCGGTTGT	TGTG	GGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 325
NGSPLEx	TCGCGAAAT	CAGTC	GCCTGTTTTCCCTTTGCTCCCTTTTCCCTTTTGTGTTGCA
256_194	TCGGTTGT	TCAG	CTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 326
NGSPLEx	TCGCGAAAT	GAGAC	CATTTTTCCGATAGTTAATAGTAATGGAGTAATAATGTTG
256_195	TCGGTTGT	TCTG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 327
NGSPLEx	TCGCGAAAT	GTGAC	CACTTGCCCTTTCTCTTCTTATCTCCTCCCAGTTCTGGT
256_196	TCGGTTGT	TCTG	CGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 328
NGSPLEx	TCGCGAAAT	GA CTG	TGTCCTCTCTCTCTTTCATCCACACCATCCACCTCGGTG
256_197	TCGGTTGT	TCAG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 329
NGSPLEx	TCGCGAAAT	GACAG	CCTCCTCTTCTCCACCTTCTTCTTCTTCTTCTGCTAC
256_198	TCGGTTGT	TGAG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 330
NGSPLEx	TCGCGAAAT	GACTG	ACGATGGCGGAGAAAGGAAGAGGAGGGAAGCTGGCGG

256_199	TCGGTTGT	AGAG	AATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 331
NGSPLEx	TCGCGAAAT	GAGAG	TATAATACAAAAAAGACCAAAAAACAAAACAAAA
256_200	TCGGTTGT	ACAG	CATCAATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 332
NGSPLEx	TCGCGAAAT	CTCAC	AACACAAGTGTGTTGTTGTCTTCTATCTTCTTCATGGCAT
256_201	TCGGTTGT	TGAG	GGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 333
NGSPLEx	TCGCGAAAT	GAGTG	ACCAAAACCACAATTTCTGCAGTTTAAATGTTTCACTTG
256_202	TCGGTTGT	ACAG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 334
NGSPLEx	TCGCGAAAT	CACAG	CCGCTGCTCGGTCCCCCAGGCCCGCCGTCTTGCTGT
256_203	TCGGTTGT	AGTG	TTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 335
NGSPLEx	TCGCGAAAT	GACTG	CGAGATCCTGGTGCTCCCACTCGCGTTGCTGCAGCAAGA
256_204	TCGGTTGT	AGTG	AATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 336
NGSPLEx	TCGCGAAAT	CACTG	GCCGGCCGGGGTGGGGAACGAGCGCCGGGTTCCGTCC
256_205	TCGGTTGT	TCAG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 337
NGSPLEx	TCGCGAAAT	GTCAG	TCTCTGCCACCGCTGGTGCTGCTGTCTCCCACTCGGTGG
256_206	TCGGTTGT	AGTG	TCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 338
NGSPLEx	TCGCGAAAT	CTGTG	CATCGAAGACGCTCGCTTCAGAAATGTCCCTGACTGCTG
256_207	TCGGTTGT	AGTG	CGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 339
NGSPLEx	TCGCGAAAT	CTGAC	GTCTTTCAGGTCAATGTAGTGCTGCTTCAGGTGTTCTTCA
256_208	TCGGTTGT	ACTG	GAGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 340
NGSPLEx	TCGCGAAAT	CTGTC	CATCAGCATAGCCTCCGATGACCATGGTGTTCCACAAAG
256_209	TCGGTTGT	TGAG	GGTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 341
NGSPLEx	TCGCGAAAT	CACTG	GATGCCCAGAATCAGGGCCCAGATGTTTCAGGCACTTGG
256_210	TCGGTTGT	ACTG	CGGTGGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 342
NGSPLEx	TCGCGAAAT	CAGTG	AGAACCGBAAGAGAAAGGGGCTGCGGTGCAGCACGGGA
256_211	TCGGTTGT	TCAG	AATAGGGTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 343
NGSPLEx	TCGCGAAAT	GTGAG	CCCCCAACCCCTCACTGTTTCCCGTTGCCATTGATGGTG
256_212	TCGGTTGT	AGAG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 344
NGSPLEx	TCGCGAAAT	CACTG	ACCTCATAGGTGCCTGCGTGGGCGCTCTTGTTGGTCCAG
256_213	TCGGTTGT	ACAG	GCTCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 345
NGSPLEx	TCGCGAAAT	GAGAC	ACAGGAGTCTTGCCCAAGCCCTGTCATGTCAGTGTGTGT
256_214	TCGGTTGT	TGTG	GTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 346
NGSPLEx	TCGCGAAAT	CAGTG	CTTCTTCAAGGTGATATAGACGCTGCCCGACGTCCGGTG
256_215	TCGGTTGT	TCTG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 347
NGSPLEx	TCGCGAAAT	GTGTC	GCCATCTGGGCCATCAGACCTGGCTGCCGGGGCGCATG
256_216	TCGGTTGT	TGAG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 348
NGSPLEx	TCGCGAAAT	GACTC	GCTTCTTGGGAAATGAAGCCACAGCCAGCTCATATATGT
256_217	TCGGTTGT	TGAG	GGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 349
NGSPLEx	TCGCGAAAT	CAGAG	CTCATCCACGATGGCTGCTATCGGTAAACAGTTAAACA
256_218	TCGGTTGT	TCTG	GTCTGTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 350
NGSPLEx	TCGCGAAAT	GTCAG	TGACCCGCTCGATCGGAGCCACGGCCGTCTTGAGATG

256_219	TCGGTTGT	AGAG	GTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 351
NGSPLEx	TCGCGAAAT	CAGTC	GGGTGATCAGCTGTGAGGCATTGAACTTGGCCACCACAC
256_220	TCGGTTGT	AGTG	TCTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 352
NGSPLEx	TCGCGAAAT	GTGTC	ACAGCACAGTAACAAAAGTTATTAGGAAAAACAGGACTACC
256_221	TCGGTTGT	TGTG	ACAAAGATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 353
NGSPLEx	TCGCGAAAT	CAGAC	ACCAATGTTTTTTAGAAATAGTGGCACCATCATTGGTTGGT
256_222	TCGGTTGT	TGAG	CGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 354
NGSPLEx	TCGCGAAAT	GA CTG	GCAGTTTACGCTGTCTAGCCAGAGTTTCACCGTAAATATG
256_223	TCGGTTGT	TCTG	ATTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 355
NGSPLEx	TCGCGAAAT	GTCTC	CACTCTTTCACCTTAAAGAGATATAGCTAGAAGGATTACACA
256_224	TCGGTTGT	ACAG	GTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 356
NGSPLEx	TCGCGAAAT	GACAC	ACCTTCAGGTCGTCCAGCTGTTTCAGCAGCTCCTCCTGG
256_225	TCGGTTGT	ACAG	TCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 357
NGSPLEx	TCGCGAAAT	CACAG	GCCGTCAACTTGCGTCGGAACATGGTCCCCGCTTCTCGC
256_226	TCGGTTGT	TCAG	TCTGGTCTGGTGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 358
NGSPLEx	TCGCGAAAT	GTGAC	CGATCCAAAAAGTGCGCGATGCGAGTAGTCAAGTCGTAC
256_227	TCGGTTGT	ACTG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 359
NGSPLEx	TCGCGAAAT	CTGTC	AGACAATGGTCCCTCTATTTCAACACCTTTTTTCGGTGACA
256_228	TCGGTTGT	TCTG	GTGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 360
NGSPLEx	TCGCGAAAT	GAGTC	GGTGATCTTGCTCTTGCTCCTTTTCGATGGTCACCACCCC
256_229	TCGGTTGT	ACTG	TCCATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 361
NGSPLEx	TCGCGAAAT	CTCTC	AACAGCCTTTAGTTCTACAGGAAATGGCACTGATGGACA
256_230	TCGGTTGT	TGTG	GAAGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 362
NGSPLEx	TCGCGAAAT	GAGAG	CCGGCTGTCTGTCTTGCTGCTCTCCACCTTCCGCACCAC
256_231	TCGGTTGT	TCTG	CTCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 363
NGSPLEx	TCGCGAAAT	CTGAG	GGGAGGTGAACCCAGAACCCAGTTCCCCACCAAAGCTG
256_232	TCGGTTGT	AGAG	TGGAAATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 364
NGSPLEx	TCGCGAAAT	GAGTG	TCACAACAGGGGAGGCCTTGGTGAAAGCTGGGTGGAAA
256_233	TCGGTTGT	AGTG	ACCCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 365
NGSPLEx	TCGCGAAAT	CTGTG	GTGGAGTCTAGAGGATCCACAGCTGGATAGATGCCAG
256_234	TCGGTTGT	ACAG	CTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 366
NGSPLEx	TCGCGAAAT	CACTG	AGGTAAAGGCCTGCAGCGATGAAACAGTTGTAGCTGACT
256_235	TCGGTTGT	AGAG	TGCTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 367
NGSPLEx	TCGCGAAAT	CTCAG	TCATTGATTGGTTGCCCGTCAAATCGGAATCTGATCTGCT
256_236	TCGGTTGT	TCAG	GGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 368
NGSPLEx	TCGCGAAAT	CTGTG	GCTGAAACTTTACAGGCTTACAAATCTTTTGCTTAGGTG
256_237	TCGGTTGT	TGAG	CTGCCCTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 369
NGSPLEx	TCGCGAAAT	GTGAC	CACATAGAAGTCCAGGCCGTAGATACCAATGCTTGGTGG
256_238	TCGGTTGT	AGTG	TCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 370
NGSPLEx	TCGCGAAAT	CTGAC	ACCATGCCCAGCACATCCTGCACATGCTGGCCCAGGTTG

256_239	TCGGTTGT	TCAG	GAGCCCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 371
NGSPLEx	TCGCGAAAT	GTCAG	GGTGATGGTAGCCTTTCTGCCAGCGCGTGCCACAGTG
256_240	TCGGTTGT	ACTG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 372
NGSPLEx	TCGCGAAAT	CACTG	GCCGCATCCGCGTCAGATTCCCAAACTCGCGGCCATTG
256_241	TCGGTTGT	AGTG	TGGCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 373
NGSPLEx	TCGCGAAAT	CTCAG	TTGCTGTACCAGCAACGTTGCCACGACGAACATCCTTG
256_242	TCGGTTGT	AGAG	ACAGACATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 374
NGSPLEx	TCGCGAAAT	CAGTC	GCTGGTATAAGGTGGTCTGGTTGACTTCTGGTGTCCCCA
256_243	TCGGTTGT	ACAG	CGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 375
NGSPLEx	TCGCGAAAT	GAGAG	CGGCATCCTCTCAGGAGGGCCGGTCCGGGTCTCAGCGC
256_244	TCGGTTGT	AGTG	GCCTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 376
NGSPLEx	TCGCGAAAT	CAGAC	AGGTTAACCATGTGCCCGTCGATGTCCTTGGCGGAAAAC
256_245	TCGGTTGT	AGAG	TCGTGCATGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 377
NGSPLEx	TCGCGAAAT	GAGTC	ACCACAAACTCTCCACCAGCCAGCATGGCAAATTTGAG
256_246	TCGGTTGT	TGTG	GTGCTTGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 378
NGSPLEx	TCGCGAAAT	CTGAC	TGGAGATTGCAGTGAGCTGAGATCACACCACTGGGCTCC
256_247	TCGGTTGT	TCTG	AGCCTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 379
NGSPLEx	TCGCGAAAT	CAGTG	GGCCAGTGGTCTTGGTGTGCTGGCCTCGGACACGAATG
256_248	TCGGTTGT	TGTG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 380
NGSPLEx	TCGCGAAAT	CACAG	GCAGCTGGAGCATCTCCACCCTTGGTATTTCTGGTGTA
256_249	TCGGTTGT	ACTG	TGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 381
NGSPLEx	TCGCGAAAT	CTCTC	GTAGCTGGGGGTGCTGGGGTTTCATTCTCGGCACGGCTG
256_250	TCGGTTGT	AGTG	CTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 382
NGSPLEx	TCGCGAAAT	GTCAC	GCTGTAACCACACCGACGCGGAGCTCTGCGCGGGCTT
256_251	TCGGTTGT	AGTG	CACTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 383
NGSPLEx	TCGCGAAAT	CTGTC	TCCAGGTGATCTCCAAGGACTGGACTGTACGTCTCAGC
256_252	TCGGTTGT	AGAG	TCTTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 384
NGSPLEx	TCGCGAAAT	GACAG	TTAACCTACCACTGTTTTGTTTAGAGCGAACACAGTGTGG
256_253	TCGGTTGT	TGTG	TCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 385
NGSPLEx	TCGCGAAAT	GACAC	TCTCCTCCAGGGTGGCTGTCACTGCCTGGTACTTCCATG
256_254	TCGGTTGT	ACTG	GTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 386
NGSPLEx	TCGCGAAAT	GAGTG	CACGACAGCAAATAGCACGGGTGTCAGATGCCCTTGGCTGA
256_255	TCGGTTGT	ACTG	AAAGTGGTCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 387
NGSPLEx	TCGCGAAAT	GTCAC	GGGACCAGCCGTCCTTATCAAAGTGCTCCCAGAAATTGG
256_256	TCGGTTGT	TCAG	TCGGTGCTCGCAGGCTCGGCA
	SEQ ID NO: 131		SEQ ID NO: 388
Name	Region 1	Region 2	
Capture			
Probe256	CWSWSWS	ACAACCGAATTTTCGCGAT TTTTTTTTTT\Biotin	
plex	WS	(SEQ ID NO: 389)	

Table 3 - List of sequences used for proof of concept experiment Fig. 6

Name	Region 3	Region 4	Region 5
Precursor	TAGCGCCT	CTCTCT	TAGCGCCTGCGGCCTGTCTCTCTCTGUAAACGCATCGG
Oligo 1	GCGGCCT	CTGU	TCGAATTATCTCCTGCTAGGCACTCGCTGTGCCCTGGA
	GT		CTATCGTAACCCATGCTGTTT/36-FAM-3'
	SEQ ID NO: 390	SEQ ID NO: 391	SEQ ID NO: 392
Precursor	TAGCGCCT	GTGTGA	CTTGCGGAACACGAATCGACCACTGACACAATTCGTAAT
Oligo 2	GCGGCCT	CTGU	CTCATTGCAAGCGTTT/36-FAM-3'
	GT		
	SEQ ID NO: 390	SEQ ID NO: 393	SEQ ID NO: 394
Precursor	TAGCGCCT	CAGAGA	ATGCCCATTTCAGCCTCACGTGGTGCTGATTTGGGGTGT
Oligo 3	GCGGCCT	CTGU	TT/36-FAM-3/
	GT		
	SEQ ID NO: 390	SEQ ID NO: 395	SEQ ID NO: 396
Name	Region 1	Region 2	
Capture			
Probe2	CAGWSWS	ACAGGCCGCGAGGCGCTA	/iBiodT/TTTTTTTT/iBiodT/TTTTTTTT/3B
56plex	WS	(SEQ ID NO: 397)	

5

Table 4 - List of sequences used for Fig. 10

Name/Region	6	7	8
Precursor	GTGGATGATC	ACACACAC	UCGCTTCCATACCGGGCGATGGACACAATTAAGAT
Oligo 4	AACGC		CGCATTTAGAGTGAAGTATCAATCGGAAATCGTGC
	SEQ ID NO: 398		AGCGACC/36-FAM-3'
			SEQ ID NO: 399
Precursor	GTGGATGATC	ACTGTGTC	UCAATCAACCAGATTAGGACTCGGTTCCCGTGAGA
Oligo 5	AACGC		AATAGAAGTCCGTATAAACGTTCAACGGGGTGC/36-FAM-3'
	SEQ ID NO: 398		SEQ ID NO: 400
Precursor	GTGGATGATC	AGTCTCTC	UCGCTTCCATACCGGGCGATGGACACAATTAAGAT
Oligo 6	AACGC		CGCATTTAGAGT/36-FAM-3'
	SEQ ID NO: 398		SEQ ID NO: 401
Name/Region	1	4	2
Capture	GWSWSWST	GCGTTGAT	ATAGACTCG/iBiodT/TTTTTTTT/iBiodT/TTTTTTTT/3B
Probe		CATCCAC	
		SEQ ID NO: 402	SEQ ID NO: 403
Name/Region	5	3	
Protect	AGTCTAT	GTGGATGA	
or		TCAACGC	
		SEQ ID NO: 398	

* * * * *

All of the compositions and methods disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this disclosure have been described in terms of preferred
5 embodiments, it will be apparent to those of skill in the art that variations may be applied to the compositions and methods and in the steps or in the sequence of steps of the method described herein without departing from the concept, spirit and scope of the disclosure. More specifically, it will be apparent that certain agents which are both chemically and physiologically related may be substituted for the agents described herein while the same or
10 similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the disclosure as defined by the appended claims.

XI. References

The following references, to the extent that they provide exemplary procedural or other details supplementary to those set forth herein, are specifically incorporated herein by reference.

5

SantaLucia Jr, J., & Hicks, D. (2004). The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, 33, 415-440.

10 Wu, L. R., Wang, J. S., Fang, J. Z., Evans, E. R., Pinto, A., Pekker, I., ... & Zhang, D. Y. (2015). Continuously tunable nucleic acid hybridization probes. *Nature methods*, 12(12), 1191-1196.

Zhang, D. Y., Chen, S. X., & Yin, P. (2012). Optimizing the specificity of nucleic acid hybridization. *Nature chemistry*, 4(3), 208-214.

CLAIMS**What is Claimed:**

1. A method for generating a set of precursor nucleotide sequences comprising a target oligonucleotide molecule, wherein the precursor nucleotide sequence comprises a
5 fifth region that comprises the nucleotide sequence of the target oligonucleotide molecule and a fourth region and a third region, wherein at least one of the fourth and third regions differs from any subsequence within the target oligonucleotide, the method comprising:
 - 10 (i) calculating for the precursor nucleotide sequence the standard free energy of hybridization between the precursor nucleotide sequence and (a) a first oligonucleotide comprising a second region that is complementary to the third region of the precursor oligonucleotide sequence, and (b) a first region that is complementary to the fourth region of the precursor oligonucleotide sequence;
 - 15 (ii) calculating the standard free energy of a capture reaction as the standard free energy of hybridization between the precursor nucleotide sequence and the first oligonucleotide and the standard free energy of hybridization between the first oligonucleotide and target oligonucleotide;
 - (iii) rejecting the precursor nucleotide sequence if the standard free energy of the capture reaction does not meet a certain criterion; and
 - 20 (iv) repeating steps (i) to (iii) until a set of precursor nucleotide sequences meets the criteria; and
 - (v) producing the set of precursor nucleotide sequences.
2. The method of claim 1, wherein the criteria is a negative free energy.
- 25 3. A method for producing a set pf precursor nucleotide sequences comprising a plurality of barcode sequences, comprising:
 - 30 (i) generating a set of precursor nucleotide sequences comprising each target oligonucleotide molecule, wherein each precursor nucleotide sequence comprises a third region that is conserved across all precursor nucleotide sequences, a fourth region that is unique for each target oligonucleotide molecule, and an fifth region that comprises the nucleotide sequence of a target oligonucleotide molecule;

(ii) calculating for each precursor nucleotide sequence the standard free energy of hybridization between the precursor nucleotide sequence and (a) a first oligonucleotide, comprising a second region that is complementary to the third region of the precursor oligonucleotide sequence, and (b) a first region that is complementary to the fourth region of the precursor oligonucleotide sequence;

(iii) calculating for each precursor nucleotide sequence the standard free energy of hybridization of folding;

(iv) calculating the standard free energy of a capture reaction as the standard free energy of hybridization between the precursor nucleotide sequence and the first oligonucleotide and the standard free energy of folding of the first oligonucleotide and the standard free energy of folding of the precursor oligonucleotide;

(v) rejecting the set of precursor nucleotide sequences if the standard free energy of the capture reaction for any precursor nucleotide sequence exceeds a certain criteria;

(vi) repeating steps (i) to (v) until a set of precursor nucleotide sequences meets the criteria; and

(vii) producing said set of precursor nucleotide sequences.

4. The method of claim 3, wherein the criteria is a selected from the group consisting of a maximum range of standard free energies of capture, a standard deviation of standard free energies of capture, and a difference between two ranks in a sorted list.

5. The method of claim 3, wherein the criteria is a maximum range of no more than 5 kcal/mol between a lowest standard free energy of capture and a highest standard free energy of capture for the set of precursor nucleotide sequences.

6. The method of claim 3, wherein the maximum range is no more than 2 kcal/mol.

7. A method for purifying one or multiple target nucleic acid molecules from a sample comprising one or a plurality of species of precursor molecules, wherein each species of precursor molecule comprises an fifth region comprising a target nucleic acid molecule sequence, a fourth region comprising a sequence unique to the species of precursor molecule in the plurality of species of precursor molecules, defined as a barcode sequence of length n wherein 2^n is greater than or equal to the number of unique target nucleic acid molecule sequences, a third region that is conserved across all precursor molecules, the method comprising:

contacting the sample with a capture probe library at temperature and buffer conditions conducive to hybridization;

wherein the capture probe library comprises a plurality of capture probe species, wherein each capture probe species comprises a first oligonucleotide comprising a first region comprising a nucleotide sequence of n nucleotides in length and a second region that is conserved across all capture probe species, wherein each nucleotide in the nucleotide sequence of n nucleotides in length is selected from two or more nucleotides and the first region is unique to each capture probe, and wherein the second region is complementary to the third region, and wherein the fourth region of each species of precursor molecule is complementary to the first region of a species of precursor molecule;

separating the plurality of species of precursor molecules hybridized to the plurality of capture probe species from the species of precursor molecules not hybridized to the plurality of capture probe species;

treating the plurality of species of precursor molecules hybridized to the plurality of capture probe species with a cleavage agent sufficient to site-specifically cleave the plurality;

of species of precursor molecules at a site to separate the fifth regions from at least a portion of the third and fourth regions;

recovering the fifth regions from the plurality of capture probe species and the at least a portion of the third and fourth regions, and

thereby producing a purified target nucleic acid molecule or molecules.

8. The method of claim 7, wherein each capture probe species further comprises a second oligonucleotide comprising a ninth region, wherein the ninth region is complementary to the second region.

9. The method of claim 8, wherein each first oligonucleotide further comprises a seventh region, each second oligonucleotide further comprises an eighth region, and wherein the seventh region is complementary to the eighth region.

10. The method of claim of any one of claims 7-9, wherein each first oligonucleotide further comprises a chemical moiety, and wherein the separating the plurality of species of precursor molecules hybridized to the plurality of capture probe species comprises surface capture of the chemical moiety.

11. The method of claim 10, wherein the chemical moiety is selected from the group consisting of biotin, a thiol, an azide, an alkyne, a primary amine and a lipid.
- 5 12. The method of claim of any one of claims 7-9, wherein the first nucleotide hybridized with the precursor oligonucleotide is the preferred ligand of an antibody or other receptor that mediate the surface capture of the complexes.
- 10 13. The method of any one of claims 7-12, wherein the recovering the fifth regions from the plurality of capture probe species and the at least a portion of the third and fourth regions comprises a treatment selected from the group consisting of heating, introducing denaturants, washing with low salinity buffers, and introducing a nuclease.
- 15 14. The method of any one of claims 7-13, wherein the site-specific cleavage comprises a treatment selected from the group consisting of changing the temperature, changing the pH, and illuminating the plurality of species of precursor molecules hybridized to the plurality of capture probe species at a specific wavelength.
- 20 15. The method of any one of claims 7-14, wherein the standard free energies of binding between each first oligonucleotide and a DNA sequence complementary to the entire sequence of the first oligonucleotide are within 5 kcal/mol of each other.
- 25 16. The method of any one of claims 7-15, wherein the two or more nucleotides at each nucleotide in the nucleotide sequence of n nucleotides in length are A or T.
- 30 17. The method of any one of claims 7-15, wherein the two or more nucleotides at each nucleotide in the nucleotide sequence of n nucleotides in length are G or C.
- 35 18. The method of any one of claims 7-15, wherein the two or more nucleotides at each nucleotide in the nucleotide sequence of n nucleotides in length are G or C for one or more nucleotides in the nucleotide sequence and A or T for one more nucleotides in the nucleotide sequence.
19. The method of any one of claims 7-18, wherein the first region comprises between 3 and 25 nucleotides.
20. The method of any one of claims 7-19, wherein n is between 3 and 60, 3 and 18, or 3 and 10 and not greater than the number of nucleotides in the first region.

21. The method of any one of claims 7-20, wherein the first region further comprises at least one nucleotide in addition to the nucleotide sequence of n nucleotides in length.
- 5 22. The method of any one of claims 7-20, wherein the second region comprises between 8 and 200 nucleotides.
23. The method of any one of claims 7-22, wherein the barcode sequence of each species of precursor molecule is assigned based on a method comprising:
- 10 generating a set of precursor nucleotide sequences comprising each target oligonucleotide molecule, wherein each precursor nucleotide sequence comprises a third region that is conserved across all precursor nucleotide sequences, a fourth region that is unique for each target oligonucleotide molecule, and an fifth region that comprises the nucleotide sequence of the target oligonucleotide molecule;
- 15 calculating for each precursor nucleotide sequence the standard free energy of hybridization between the precursor nucleotide sequence and a first oligonucleotide, comprising a second region that is complementary to the third region of the precursor oligonucleotide sequence, and a first region that is complementary to the fourth region of the precursor oligonucleotide sequence;
- 20 calculating for each precursor nucleotide sequence the standard free energy of hybridization of folding;
- calculating the standard free energy of a capture reaction as the standard free energy of hybridization between the precursor nucleotide sequence and the first oligonucleotide and the standard free energy of folding of the first oligonucleotide and the standard free energy of folding of the precursor oligonucleotide;
- 25 rejecting the set of precursor nucleotide sequences if the standard free energy of the capture reaction for any precursor nucleotide sequence exceeds a certain criteria; and
- repeating the method until a set of precursor nucleotide sequences meets the criteria.
24. A capture probe library comprising:
- 30 a plurality of oligonucleotides comprising a first plurality of oligonucleotides wherein each oligonucleotide of the first plurality of oligonucleotides comprises:

a first region comprising a first nucleotide sequence comprising at least 3 variable positions, wherein each variable position comprises a nucleotide selected from at least two possible nucleotides,

wherein the first nucleotide sequence comprising at least 3 variable positions is unique to each oligonucleotide, and

a second region comprising a second nucleotide sequence, wherein the second nucleotide sequence of the second region is conserved across each oligonucleotide in the first plurality of oligonucleotides,

25. A capture probe library comprising:

a plurality of oligonucleotides comprising a first plurality of oligonucleotides and a second plurality of oligonucleotides, wherein each oligonucleotide in the first plurality of species of oligonucleotides comprises

a first region comprising a nucleotide sequence comprising at least 3 variable positions, wherein each variable position comprises a nucleotide selected from at least two possible nucleotides,

wherein the nucleotide sequence comprising at least 3 variable positions is unique to each species of oligonucleotide, and

a second region comprising a nucleotide sequence, wherein the nucleotide sequence of the second region is conserved across each oligonucleotide in the first plurality of oligonucleotides,

wherein each oligonucleotide in the second plurality of oligonucleotides comprises a third region, wherein the third region is complementary to the second region.

26. The capture probe library of claim 24 or 25, wherein the standard free energies of binding between each oligonucleotide in the first plurality of oligonucleotides and a DNA sequence complementary to the entire sequence of the respective oligonucleotide in the first plurality of oligonucleotides are within 5 kcal/mol of each other.

27. The capture probe library of any one of claims 24-26, wherein the at least two possible nucleotides at each nucleotide in are A or T.

28. The capture probe library of any one of claims 24-27, wherein the at least two possible nucleotides at each nucleotide in are G or C.
- 5 29. The capture probe library of any one of claims 24-26, wherein the at least two possible nucleotides at each nucleotide in are G or C for one or more nucleotides in the nucleotide sequence and A or T for one more nucleotides in the nucleotide sequence.
- 10 30. The capture probe library of any one of claims 24-29, wherein a concentration of a second oligonucleotide is greater than a sum of the concentrations of each oligonucleotide of the first plurality of oligonucleotides.
- 15 31. The capture probe library of any one of claims 24-30, wherein the first region comprises between 3 and 25 nucleotides.
32. The capture probe library of any one of claims 24-31, wherein the number of variable positions in the first region is between 3 and 60, between 3 and 18, or between 3 and 10, and not greater than a total number of nucleotides in the first region.
- 20 33. The capture probe library of any one of claims 24-32, wherein the first region further comprises at least one nucleotide in addition to the at least 3 variable positions.
34. The capture probe library of any one of claims 24-33, wherein the second region comprises between 8 and 200 nucleotides.
- 25 35. The capture probe library of claims 24-34, wherein the at least three variable regions are contiguous.
- 30 36. The capture probe library of claims 24-34, wherein the at least three variable regions are non-contiguous.
37. An oligonucleotide library for the multiplexed capture of a set of desired precursor nucleic acid molecules comprising:
- 35 a plurality of species of precursor molecules, wherein each species of precursor molecule comprises
- a third region comprising a nucleotide sequence that is conserved across all species of precursor molecules,

a fourth region comprising a barcode sequence comprising 3-60 nucleotides,

an fifth region comprising a target nucleic acid molecule sequence that is unique to the species of precursor molecule in the plurality of species of precursor molecules,

5 wherein the barcode sequence of each species of precursor molecule is different and wherein 2^n is greater than or equal to the number of unique target nucleic acid molecule sequences; and

10 a capture probe library comprising a plurality of capture probe species, wherein each capture probe species comprises an oligonucleotide comprising a first region comprising a nucleotide sequence of n nucleotides in length, a second region that is conserved across all capture probe species, wherein each nucleotide in the nucleotide sequence of n nucleotides in length is selected from two or more nucleotides and the first region is unique to each capture probe, and wherein the second region is complementary to the third region, and wherein the fourth region of each species of precursor molecule is complementary to the first region of a species of precursor molecule.

38. The oligonucleotide library of claim 37, wherein the standard free energies of binding between each species of first oligonucleotide in the plurality of capture probe species and a DNA sequence complementary to the entire sequence of the respective species of first oligonucleotide are within 5 kcal/mol of each other.

39. The oligonucleotide library of claim 38, wherein the two or more nucleotides at each nucleotide in the nucleotide sequence of n nucleotides in length are A or T.

25 40. The oligonucleotide library of claims 38, wherein the two or more nucleotides at each nucleotide in the nucleotide sequence of n nucleotides in length are G or C.

41. The oligonucleotide library of claims 38, wherein the two or more nucleotides at each nucleotide in the nucleotide sequence of n nucleotides in length are G or C for one or more nucleotides in the nucleotide sequence and A or T for one more nucleotides in the nucleotide sequence.

42. The oligonucleotide library of any one of claims 38-41, wherein a concentration of a second oligonucleotide is greater than a sum of the concentrations of each species of first oligonucleotide.

43. The oligonucleotide library of any one of claims 38-42, wherein the first region comprises between 3 and 25 nucleotides.
44. The oligonucleotide library of any one of claims 38-43, wherein n is between 3 and 60, 3 and 18, or 3 and 10, and not greater than the number of nucleotides in the first region.
45. The oligonucleotide library of any one of claims 38-44, wherein the first region further comprises at least one nucleotide in addition to the nucleotide sequence of n nucleotides in length.
46. The oligonucleotide library of any one of claims 38-45, wherein the second region comprises between 8 and 200 nucleotides.
47. The oligonucleotide library of any one of claims 38-46, wherein the barcode sequence of each species of precursor molecule is selected based on the method of any one of claims 3-6.
48. The oligonucleotide library of any one of claims 38-47, wherein at least one species of precursor molecule is chemically synthesized or produced enzymatically.
49. The oligonucleotide library of claim 49, wherein an enzyme used to produce the at least one species of precursor molecule is a ligase or a polymerase.
50. A method for purifying multiple target nucleic acid molecules from a sample comprising a plurality of precursor molecules, wherein the method comprises the steps of:
- providing a plurality of nucleic acid probes, wherein each probe has a sequence complementary to a region of one of the precursor molecules and a first moiety sufficient to allow isolation of the probe;
- adding the plurality of nucleic acid probes to a sample comprising the plurality of precursor molecules under conditions sufficient to promote hybridization of each nucleic acid probe to the region of the precursor molecule complimentary to the sequence thereby forming a plurality of probe- precursor complexes, wherein each precursor molecule comprises a different target nucleic acid molecule, and wherein the region of each precursor molecule does not comprise the target nucleic acid molecule; and

isolating the plurality of probe-precursor complexes via interaction with the first separating the target nucleic acid molecules from the plurality of probe-precursor.

51. A method for producing a plurality of distinct target oligonucleotides each having a specified sequence, the method comprising:

- 5 (1) synthesizing a precursor oligonucleotide for each distinct target oligonucleotide, wherein the precursor comprises a third sequence, a fourth sequence, and a fifth sequence, wherein the third sequence is identical for all precursors, the fourth sequence comprises a barcode and is distinct for all precursors, and a fifth sequence corresponds to the target sequence,
- 10 (2) synthesizing a capture probe library of distinct oligonucleotides comprising a first sequence and a second sequence, wherein the first sequence comprises degenerate randomer nucleotides and wherein at least one first sequence is complementary to each fourth sequence, and the second sequence is complementary to the third sequence,
- 15 (3) mixing the precursors and the capture probe library in an aqueous hybridization buffer,
- (4) removing precursor molecules not bound to the capture probes,
- (5) enzymatically or chemically cleaving the fifth sequence from the remainder of the precursor molecules, and
- 20 (6) removing the capture probe library and the remainder of the precursor molecules.

52. The method of claim 51, wherein the first sequence comprises an S degenerate nucleotide at certain positions and/or a W degenerate nucleotide at certain positions,
25 but does not comprise an N degenerate nucleotide at any position, such that any degenerate variant of the first sequence is complementary to one or more fourth sequences, wherein S is a strong base, W is a weak base, and N is any base.

53. The method of claims 51-52, wherein capture probes are functionalized with a moiety
30 permitting for rapid binding to a ligand, such as a moiety selected from a biotin, a thiol, an azide, or an alkyne.

54. The method of claims 51-53, wherein in step (4), eliminating precursor molecules not bound to capture probes comprises adding particles that specifically bind the capture probe, followed by removal of supernatant solution.
35

55. The method of claim 54, wherein the particles are selected from streptavidin-coated magnetic beads or streptavidin-coated agarose beads.
56. The method of claims 51-53, wherein the precursors further comprise a deoxyuracil nucleotide or an RNA nucleotide, and cleavage of the fifth sequence comprises introduction of a uracil DNA glycosylase or an RNase enzyme.
57. The method of claims 51-53, wherein the precursors further comprise a photolabile or heat-labile moiety, and the cleaving of the fifth sequence comprises exposure of the solution to light of the appropriate wavelength or heating to the appropriate temperature.
58. The method of claims 51-57, wherein the fourth sequence further comprises the "S" degenerate nucleotide at certain positions and/or the "W" degenerate nucleotide at certain positions, but does not comprise the "N" degenerate nucleotide at any position.
59. The method of claims 51-58, wherein the length of the first sequence is between 5 and 50 nucleotides, and wherein the number of degenerate nucleotides is between 1 and 30, between 1 and 20, between 1 and 10, between 2 and 8, between 2 and 6, or between 3 and 5.
60. The method of claims 51-59, wherein the length of the second sequence is between 5 and 50 nucleotides, and/or wherein the length of each target oligonucleotide is between 5 and 500 nucleotides.
61. An oligonucleotide capture probe library comprising a first sequence and a second sequence, wherein the first sequence comprises degenerate randomer nucleotides comprising an "S" degenerate nucleotide at one or more positions and/or a "W" degenerate nucleotide at one or more positions, but does not comprise an "N" degenerate nucleotide at any position, and wherein the length of the first sequence is between 5 and 50 nucleotides, the number of degenerate nucleotides is between 1 and 30, and the length of the second sequence is between 5 and 50 nucleotides.
62. The oligonucleotide capture probe library of claim 61, wherein the second sequence comprises an "S" degenerate nucleotide at certain positions and/or a "W" degenerate nucleotide at certain positions, but does not comprise a "N" degenerate nucleotide at any position.

63. The oligonucleotide capture probe library of claims 61-62, wherein oligonucleotide capture probe library is functionalized with a chemical moiety for rapid binding, selected from a biotin, a thiol, an azide, or an alkyne.

5 64. The oligonucleotide capture probe library of claims 61-63, wherein one or more of the oligonucleotide capture probes further comprise a deoxyuracil nucleotide or an RNA nucleotide.

10 65. The oligonucleotide capture probe library of claims 61-63, wherein one or more of the oligonucleotide capture probes further comprise a photolabile or heat-labile moiety.

15 66. The oligonucleotide capture probe library of claims 61-65, wherein the length of the first sequence is between 5 and 50 nucleotides, and wherein the number of degenerate nucleotides is between 1 and 30.

67. The oligonucleotide capture probe library of claims 61-66, wherein the length of the second sequence is between 5 and 50 nucleotides.

20 68. The library of claims 61-67, wherein the library has at least 8, at least 32, or at least 256 members.

69. The library of claims 61-67, wherein the library has between 8 and 32 members, between 8 and 256 members, between 32 and 256 members, between 8 and 1024 members, between 32 and 1024 members, or between 256 and 1024 members.

25 70. The oligonucleotide capture probe library of claims 61-69, wherein the library is found on one or more substrates.

30 71. An aqueous solution comprising an oligonucleotide capture probe library and a plurality of precursor oligonucleotides, wherein:

the capture probe library comprises a first sequence and a second sequence, wherein the first sequence comprises degenerate randomer nucleotides comprising an "S" degenerate nucleotide at certain positions and/or an "W" degenerate nucleotide at certain positions, but does not comprise an "N" degenerate nucleotide at any position, and wherein the length of the first sequence is between 5 and 50 nucleotides, the number of degenerate nucleotides is

35

between 1 and 30, and the length of the second sequence is between 5 and 50 nucleotides,

each of the plurality of precursor oligonucleotides comprises a third sequence, a fourth sequence, and a fifth sequence, wherein the third sequence is identical for all precursors, the fourth sequence comprises a barcode and is distinct for all precursors, the second sequence is complementary to the third sequence, and at least one instance of a first sequence is complementary to each fourth sequence.

- 10 72. The aqueous solution of claim 71, wherein the second sequence of the oligonucleotide capture probe library comprises an "S" degenerate nucleotide at certain positions and/or a "W" degenerate nucleotide at certain positions, but does not comprise a "N" degenerate nucleotide at any position.
- 15 73. The aqueous solution of claims 71-72, wherein oligonucleotide capture probe library is functionalized with a chemical moiety for rapid binding, selected from a biotin, a thiol, an azide, or an alkyne.
- 20 74. The aqueous solution of claims 71-73, wherein one or more of the oligonucleotide capture probes further comprise a deoxyuracil nucleotide or an RNA nucleotide.
75. The aqueous solution of claims 71-73, wherein one or more of the oligonucleotide capture probes further comprise a photolabile or heat-labile moiety.
- 25 76. The aqueous solution of claims 71-75, wherein the length of the first sequence of the oligonucleotide capture probe library is between 5 and 50 nucleotides, and wherein the number of degenerate nucleotides is between 1 and 30.
- 30 77. The aqueous solution of of claims 71-76, wherein the length of the second sequence of the oligonucleotide capture probe library is between 5 and 50 nucleotides.
78. The aqueous solution of claims 71-77, wherein the oligonucleotide capture probe library has at least 8, at least 32, or at least 256 members.
- 35 79. The aqueous solution of claims 71-77, wherein the oligonucleotide capture probe library has between 8 and 32 members, between 8 and 256 members, between 32 and 256 members, between 8 and 1024 members, between 32 and 1024 members, or between 256 and 1024 members.

80. The aqueous solution of claims 71-79, wherein the oligonucleotide capture probe library is found on one or more substrates.
- 5 81. The aqueous solution of claims 71-80, wherein the precursors further comprise a deoxyuracil nucleotide or an RNA nucleotide, and cleavage of the fifth sequence comprises introduction of a uracil DNA glycosylase or an RNase enzyme.

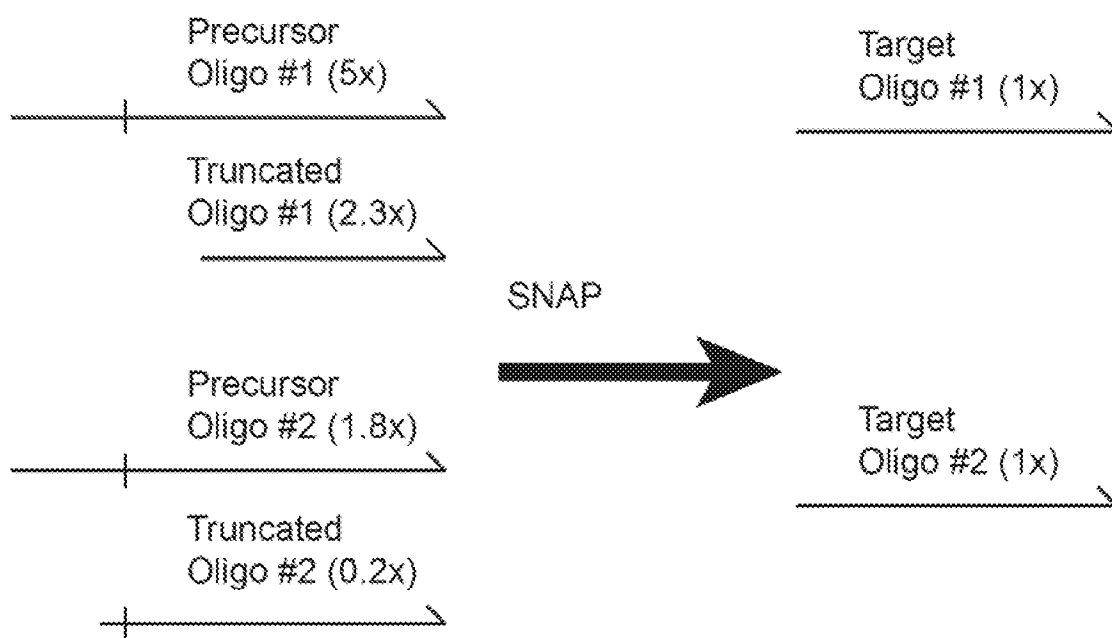


FIG. 1

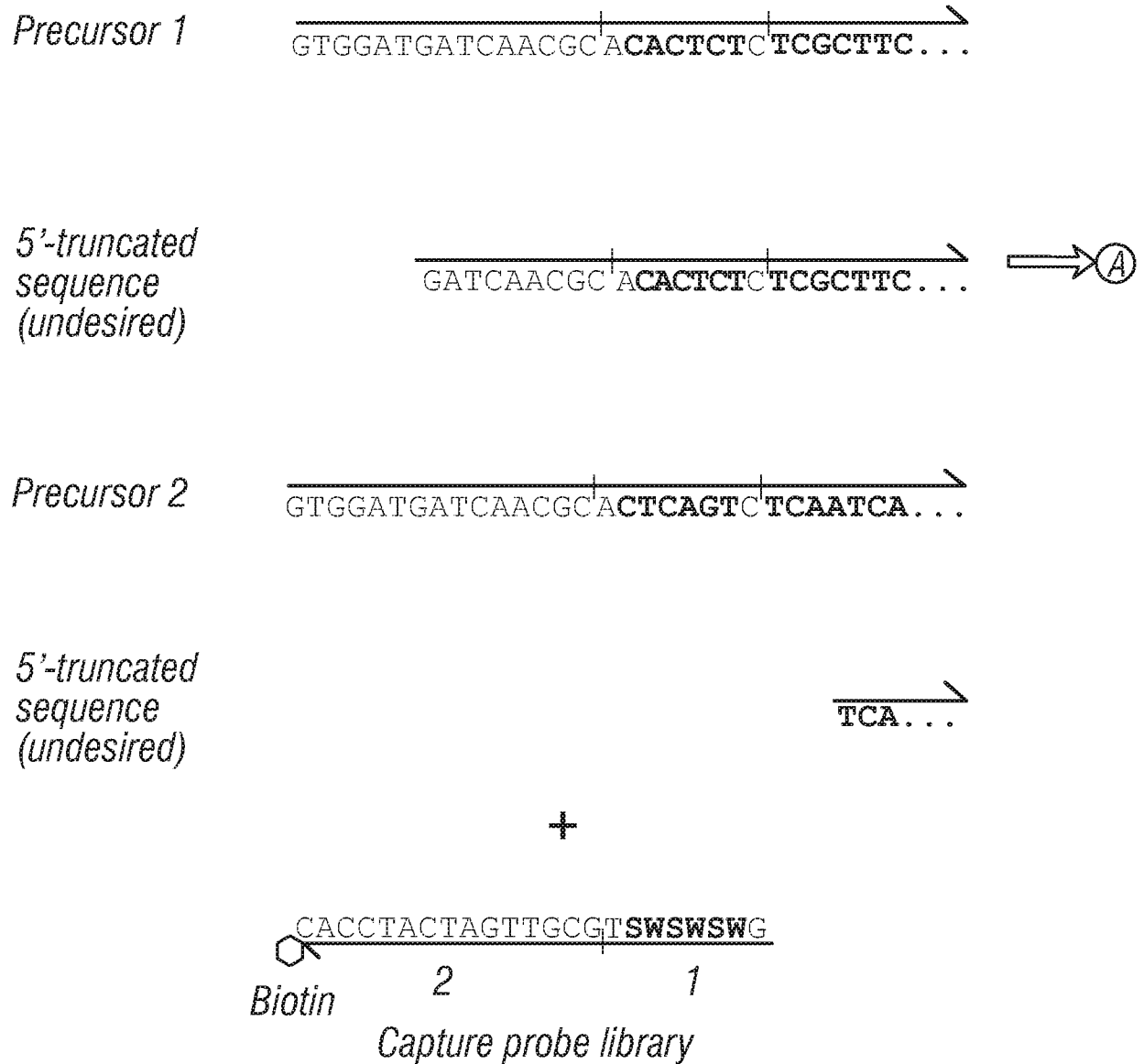


FIG. 2

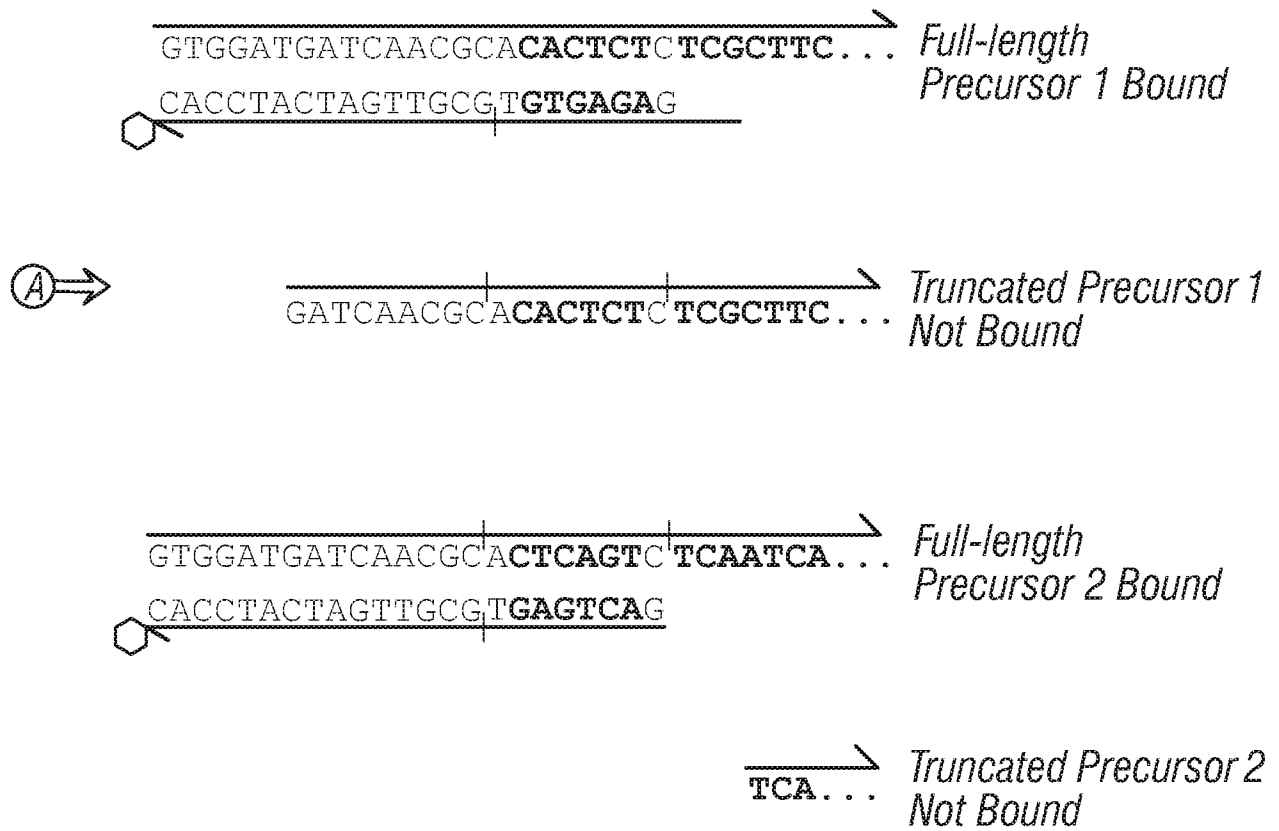


FIG. 2
(Cont'd)

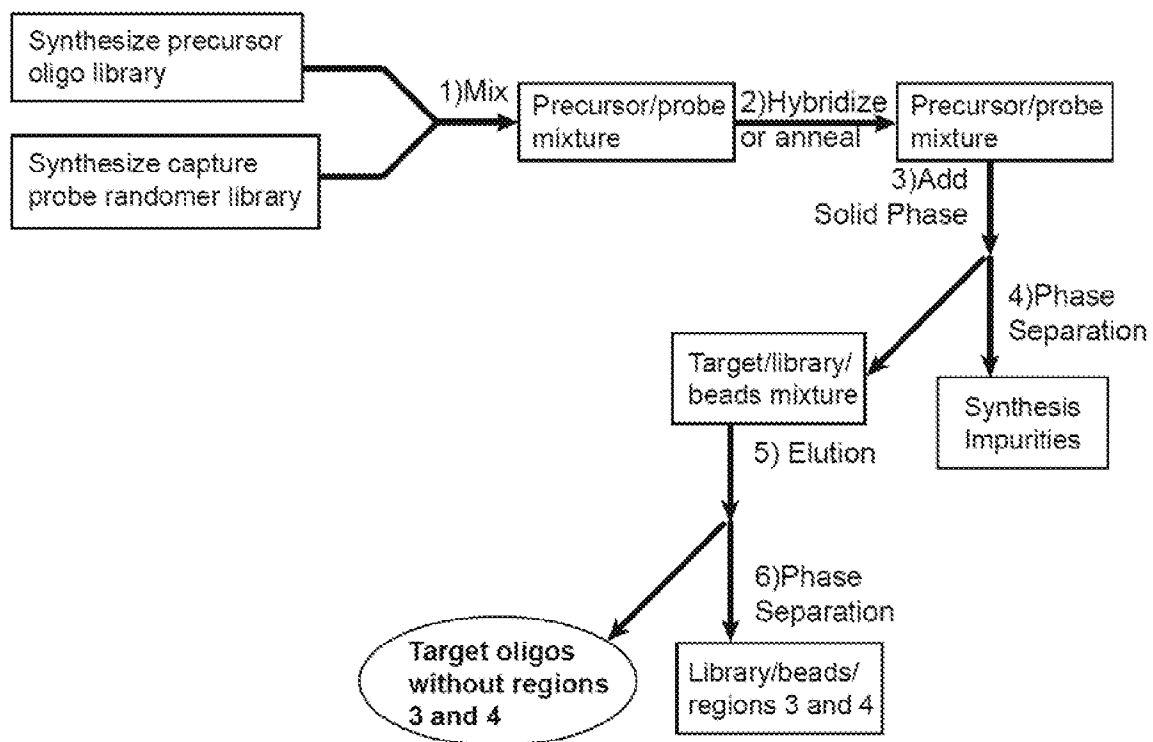


FIG. 3

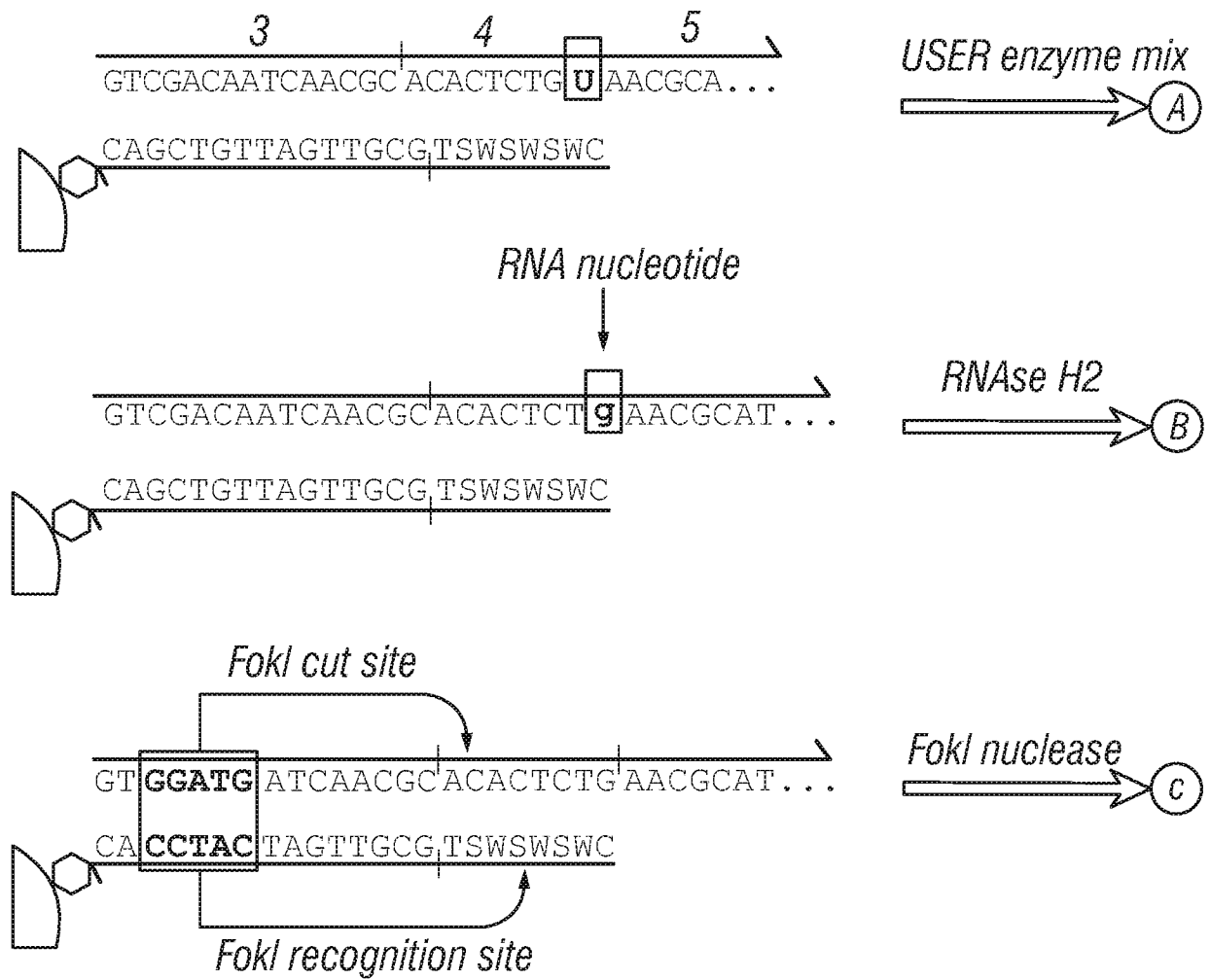


FIG. 4

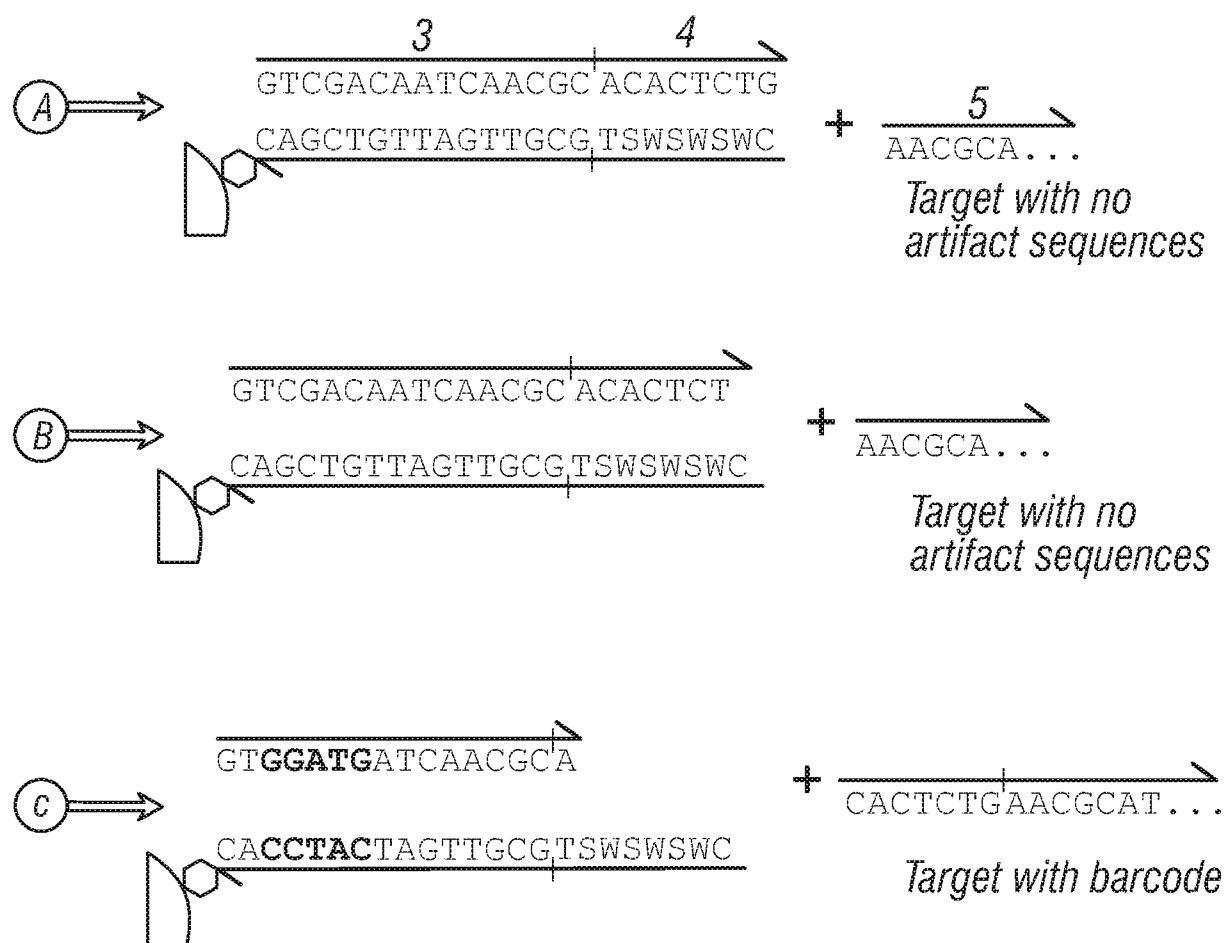
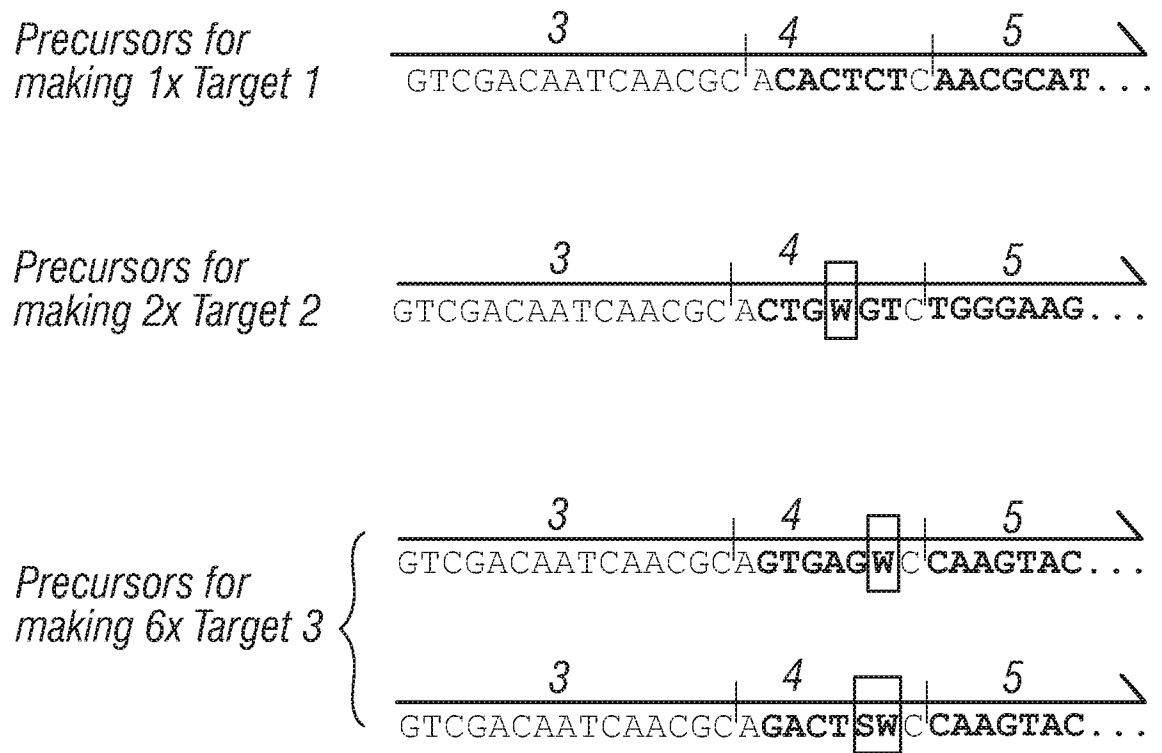


FIG. 4
(Cont'd)

**FIG. 5**

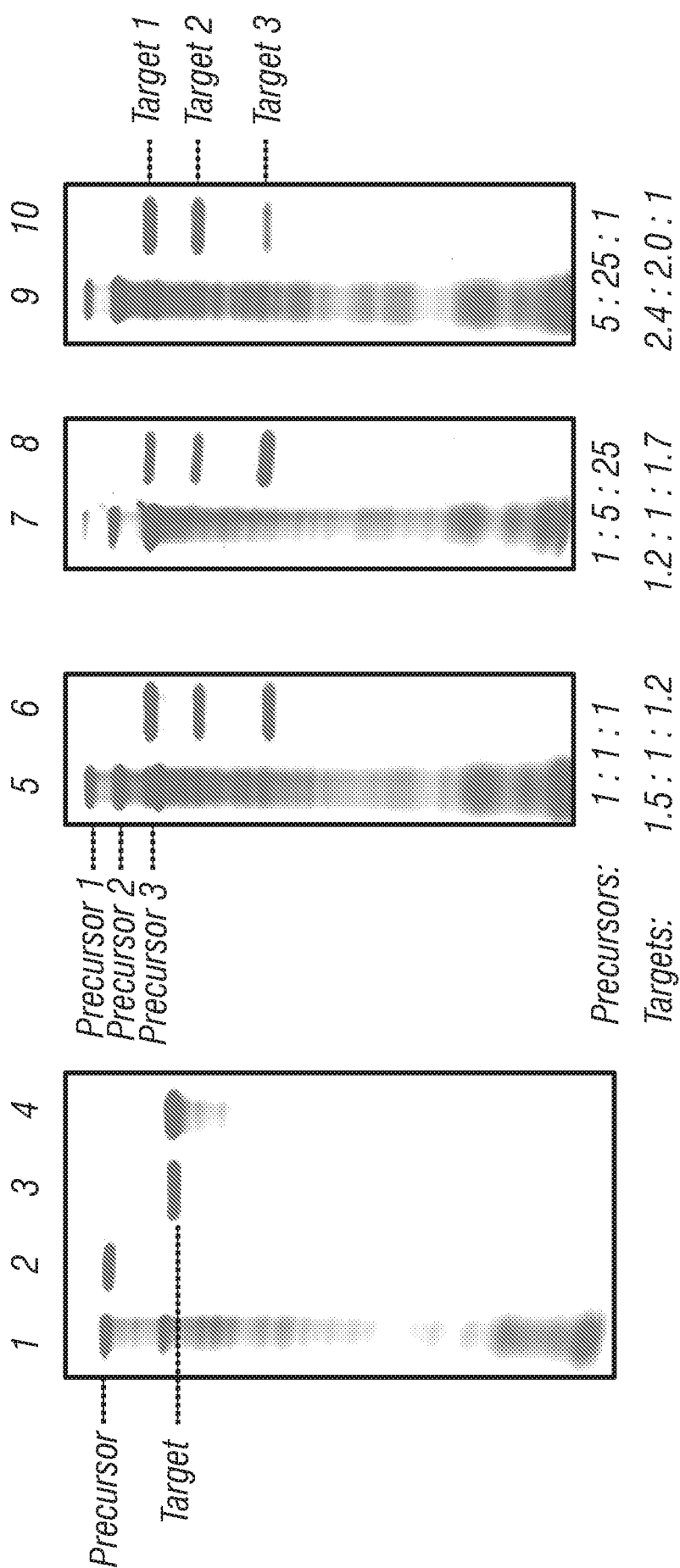


FIG. 6

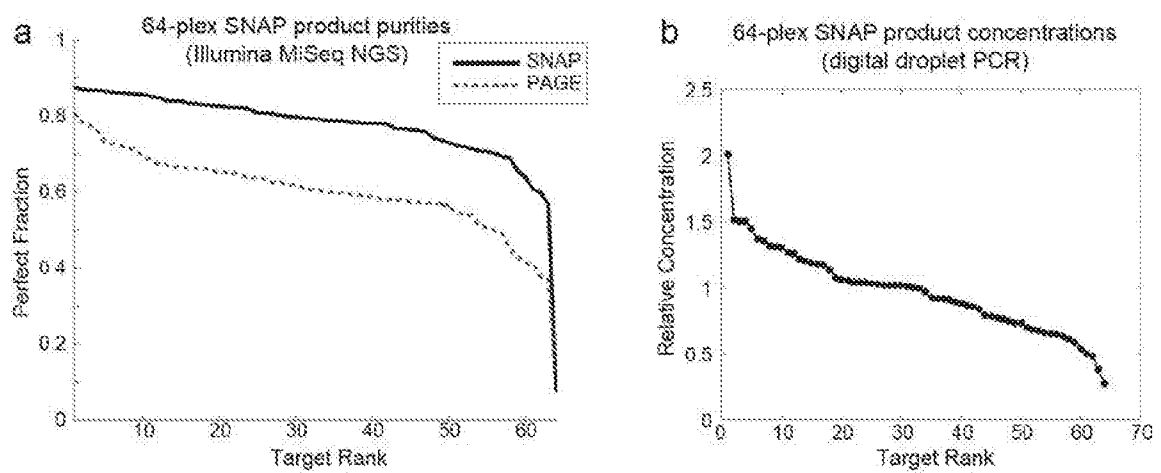


FIG. 7A-B

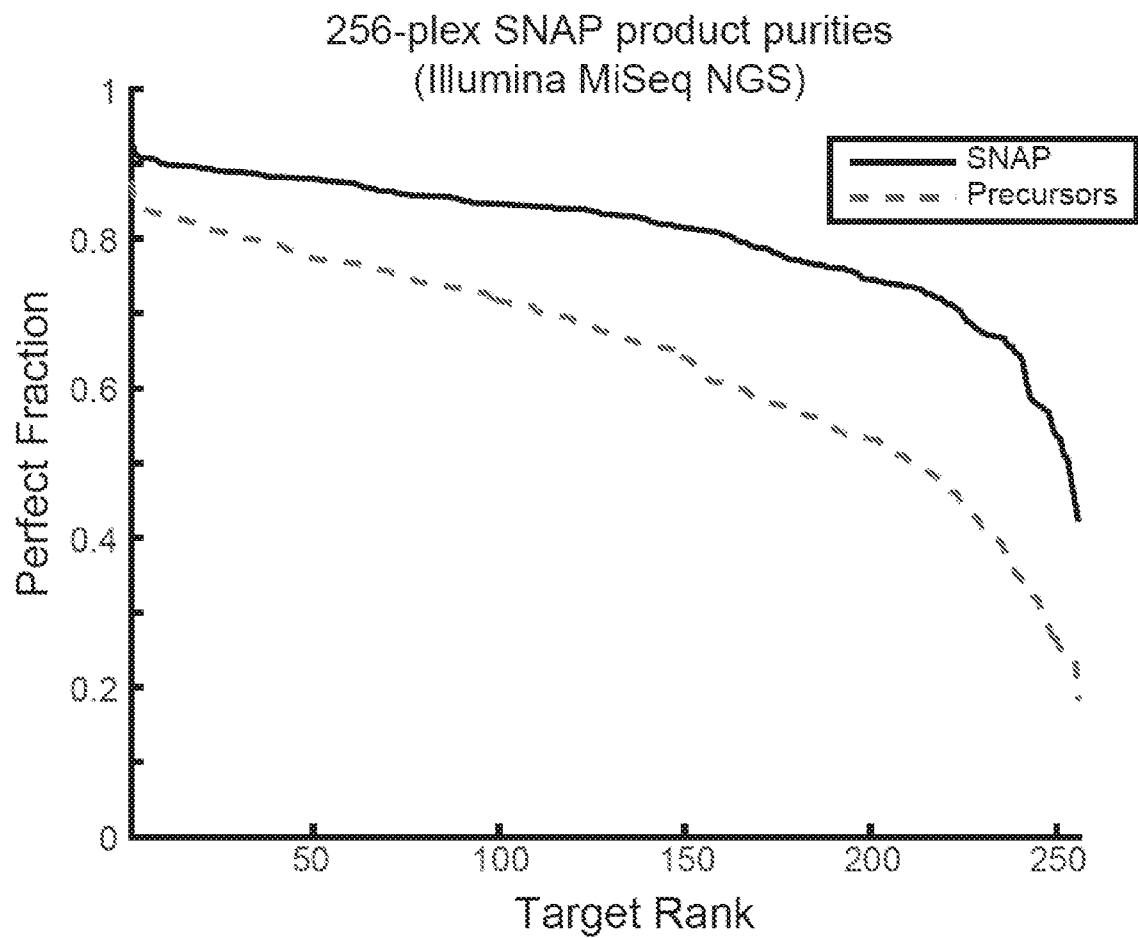
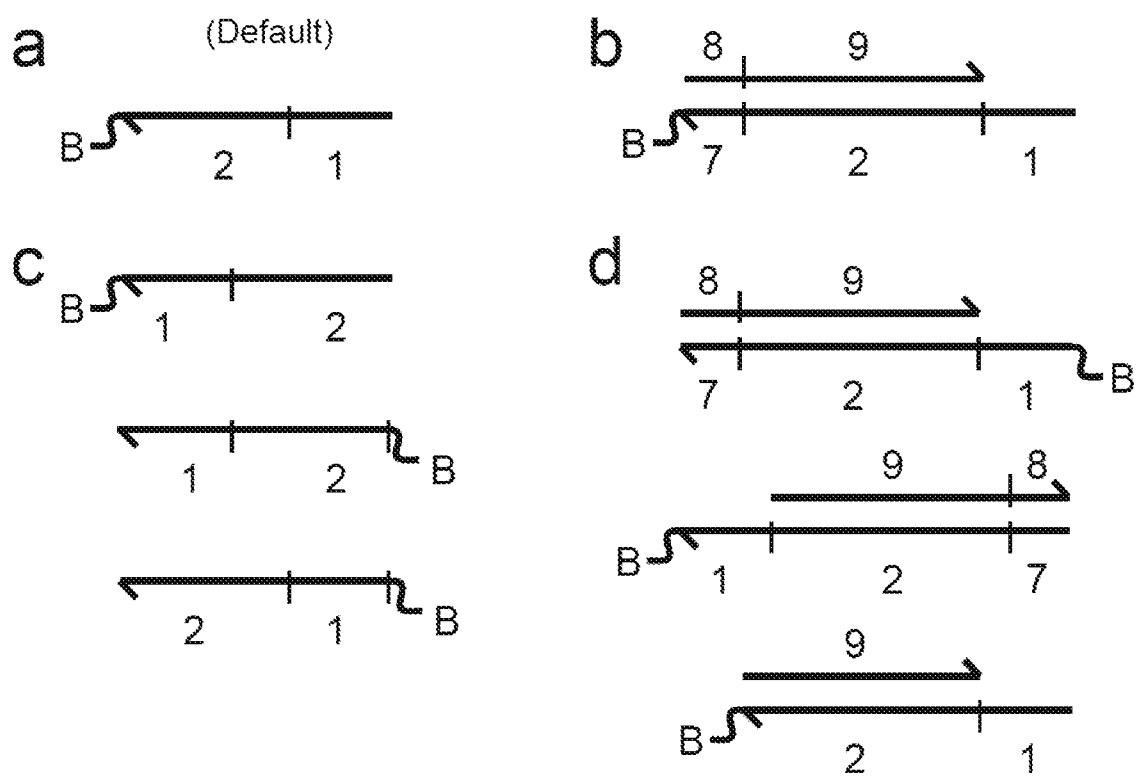


FIG. 8



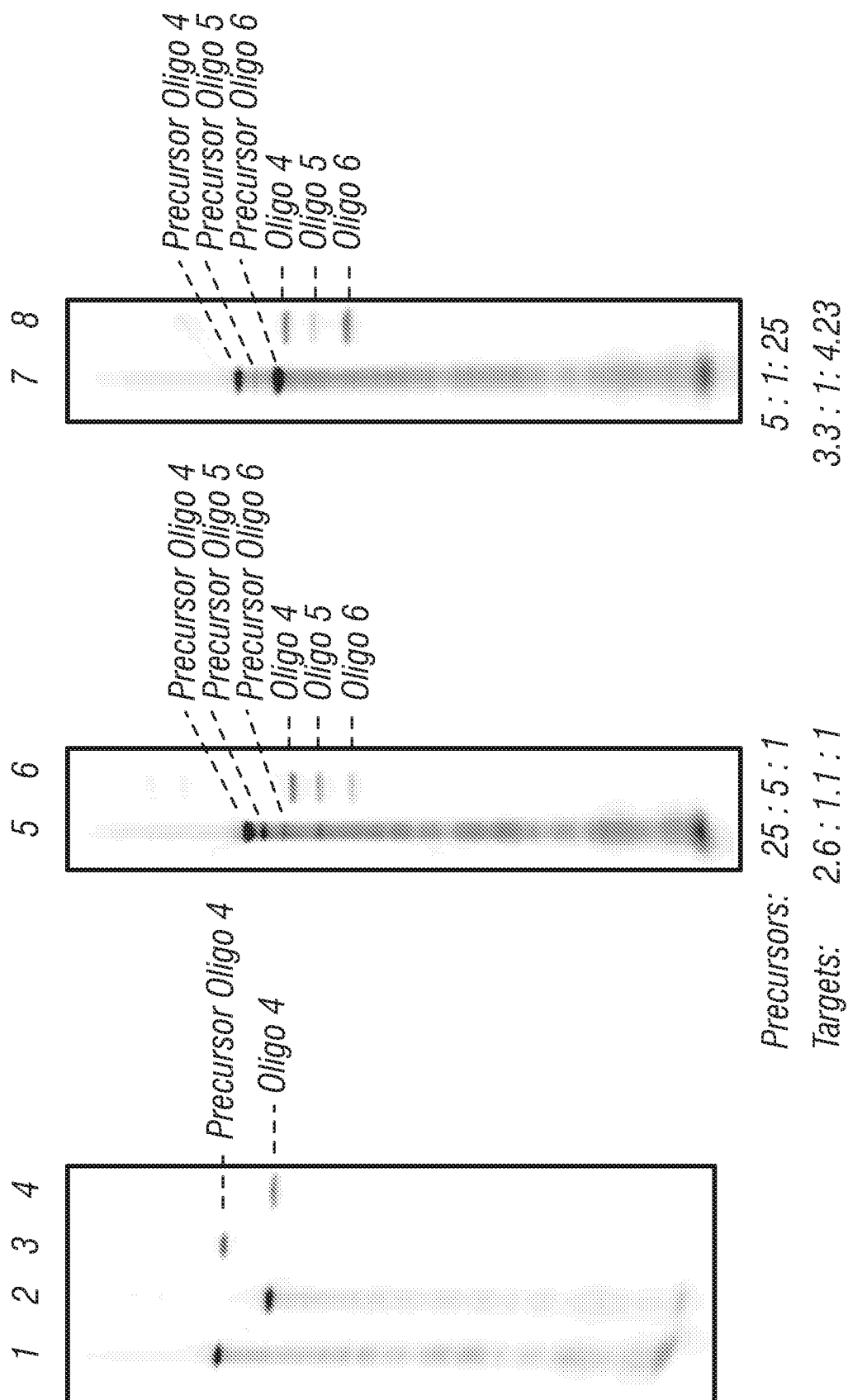


FIG. 10

RNA transcript purification

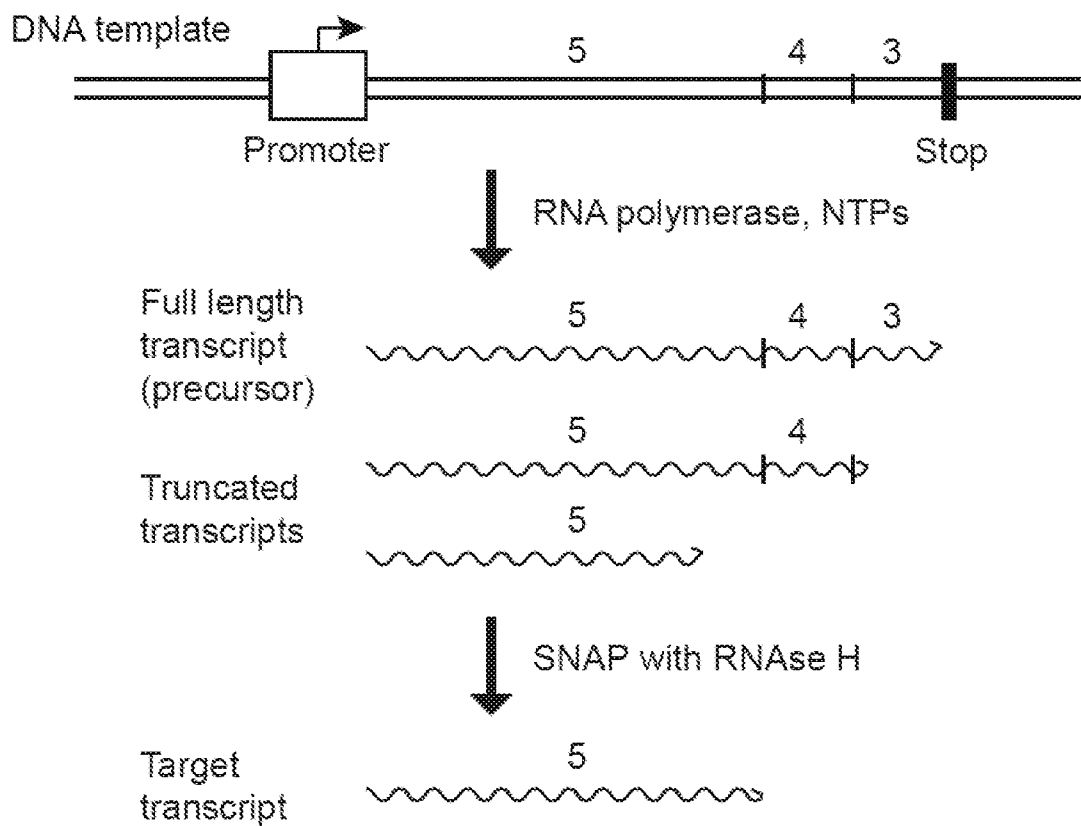


FIG. 11

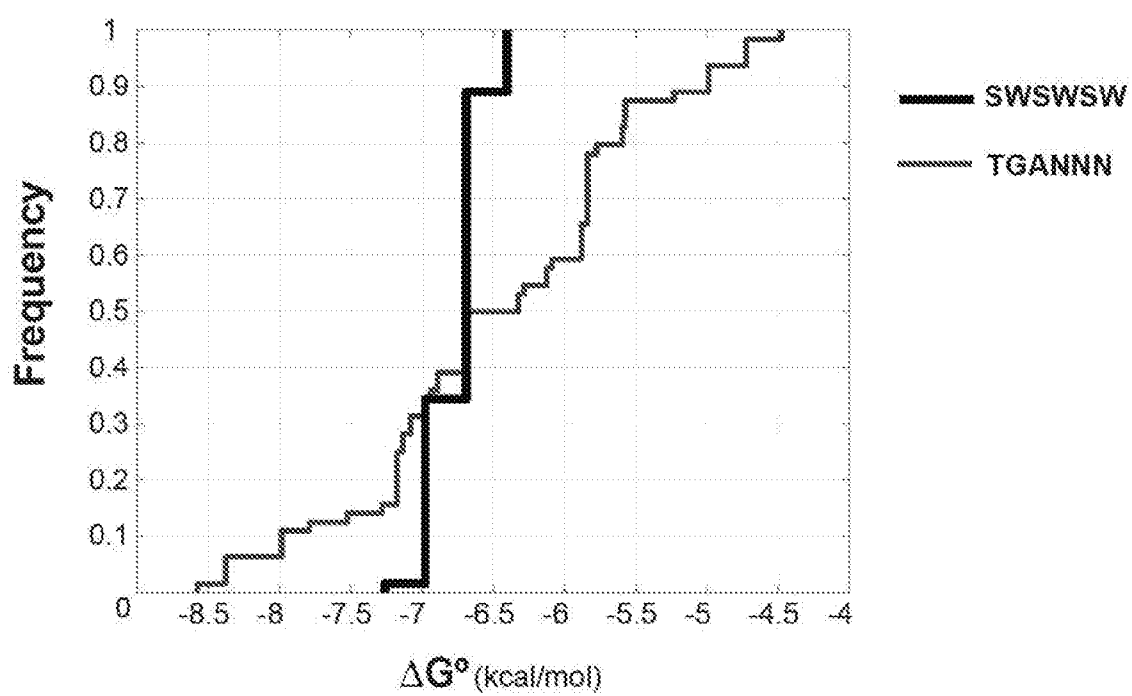


FIG. 12

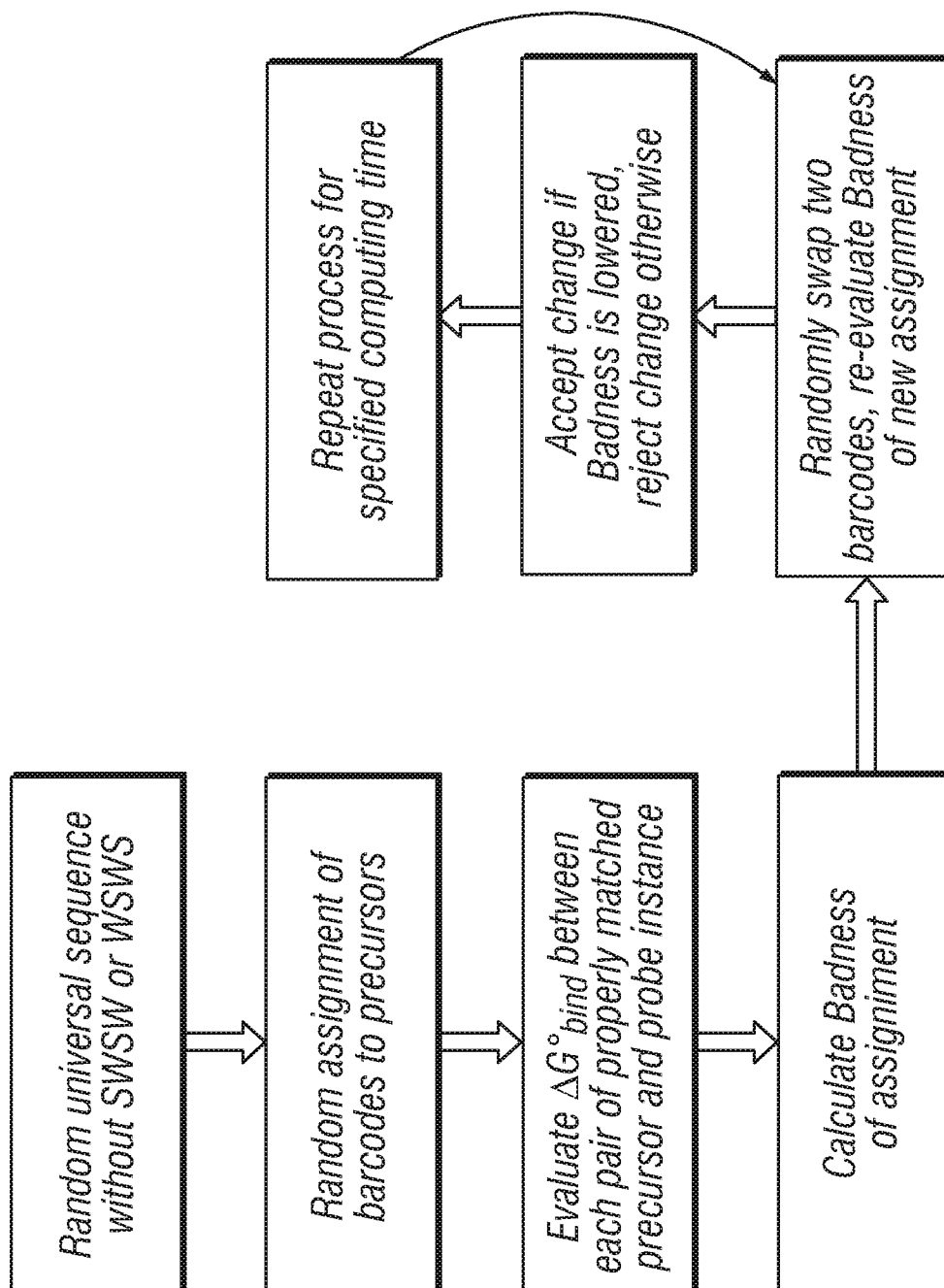


FIG. 13A

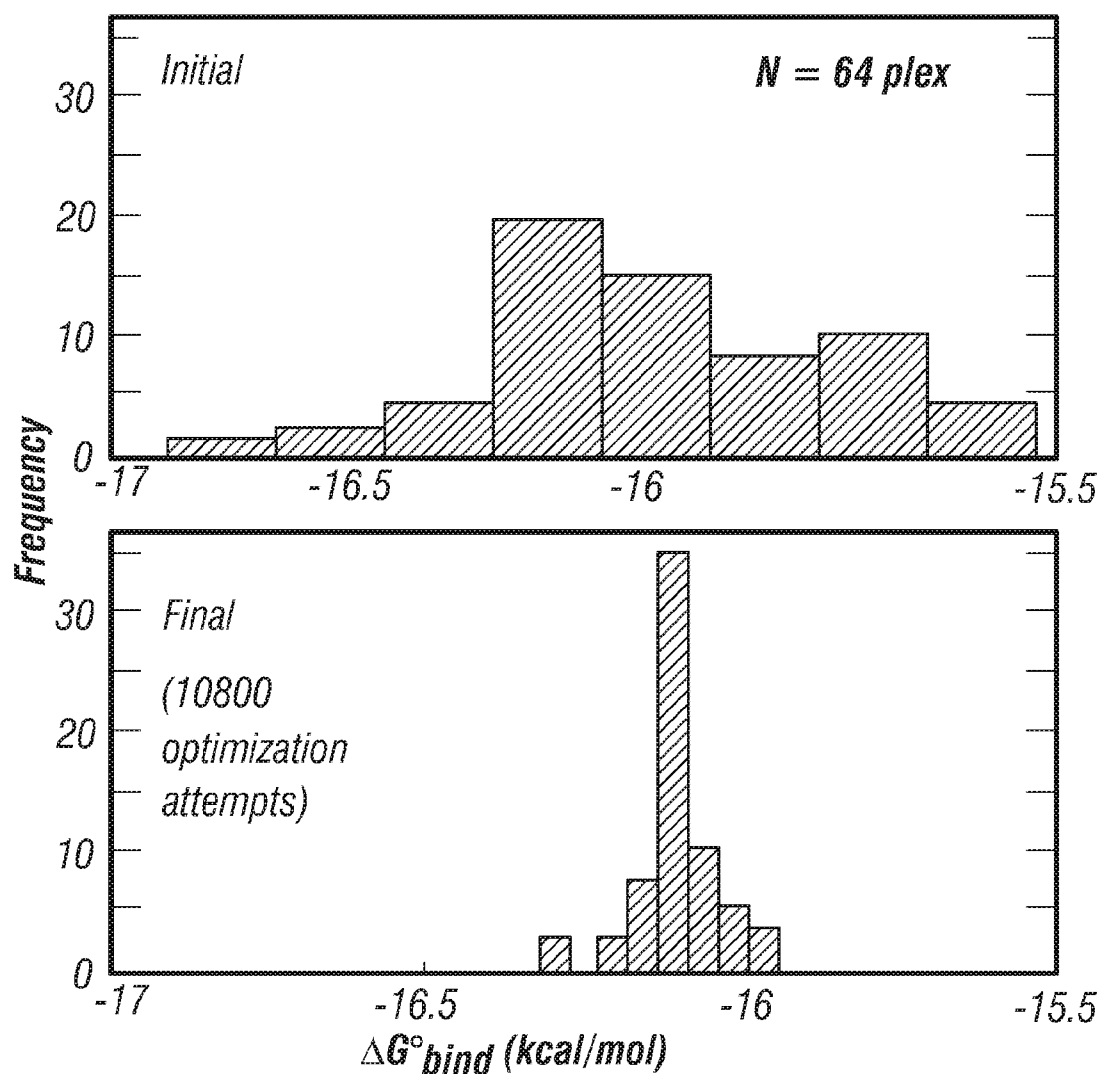


FIG. 13B

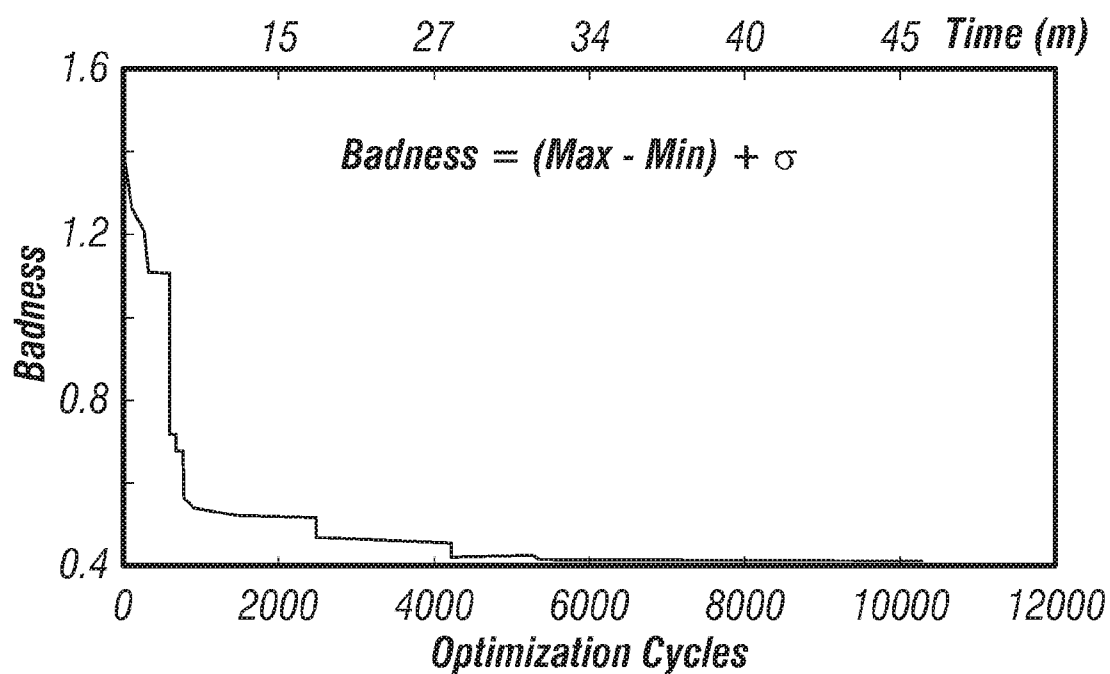
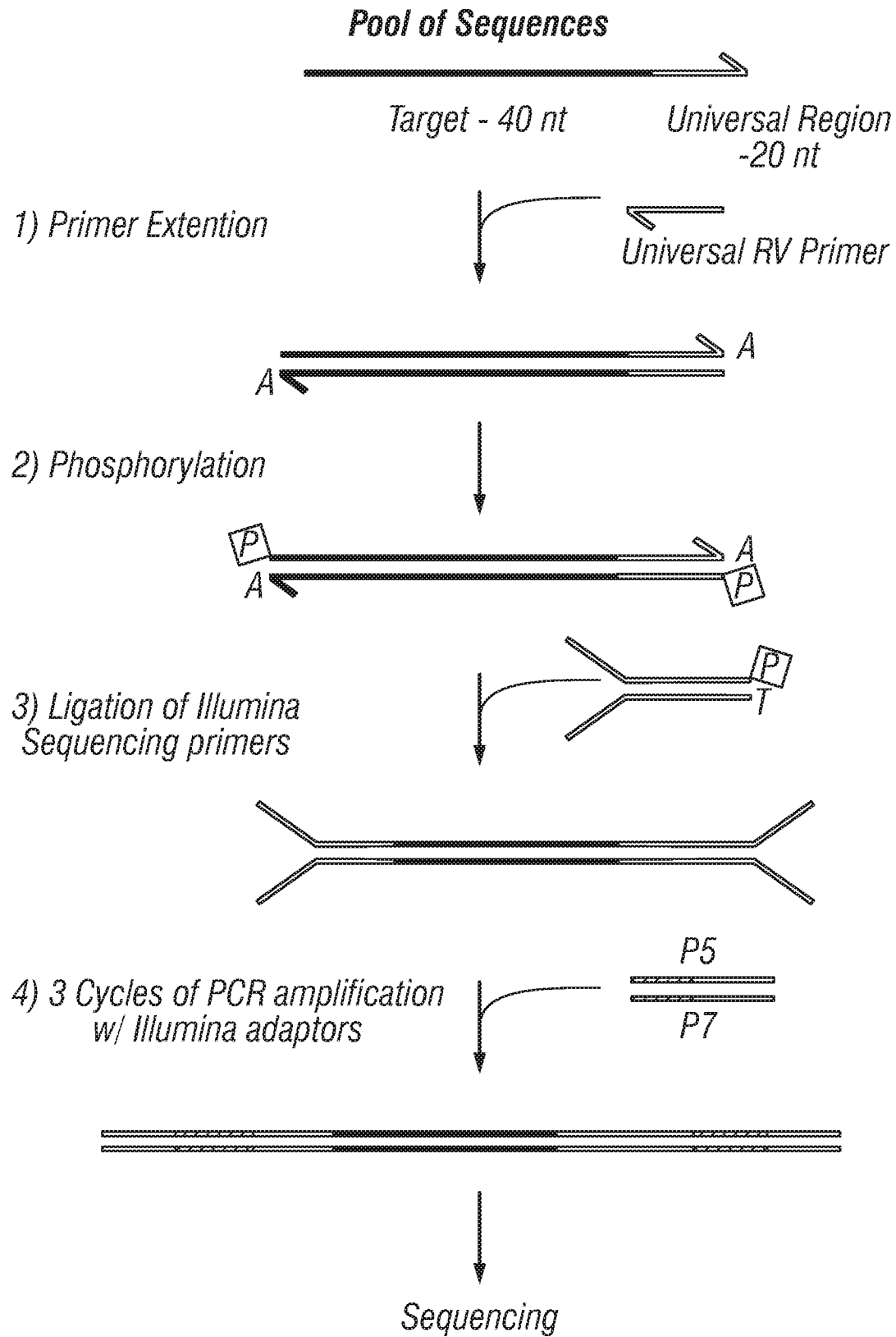
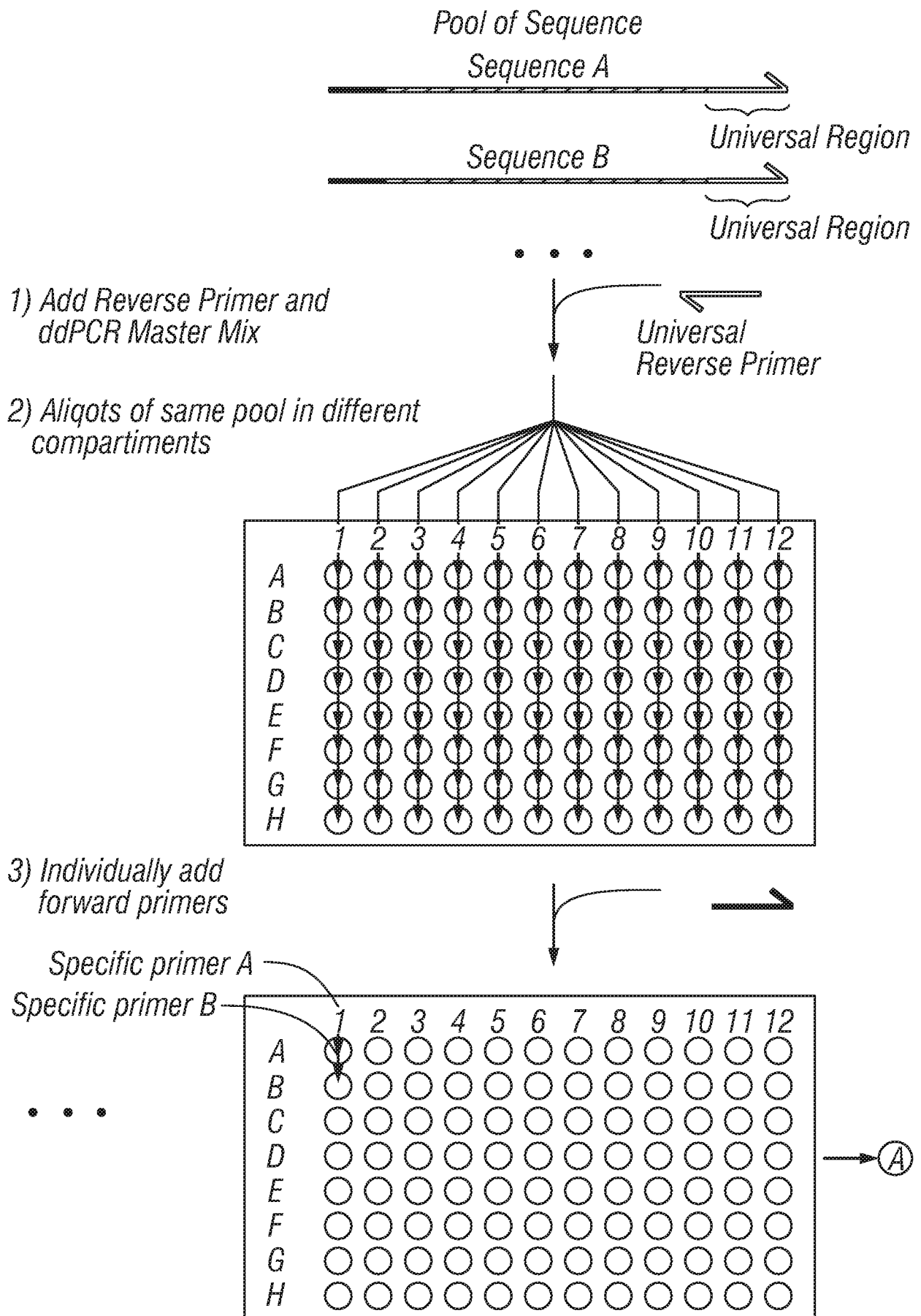
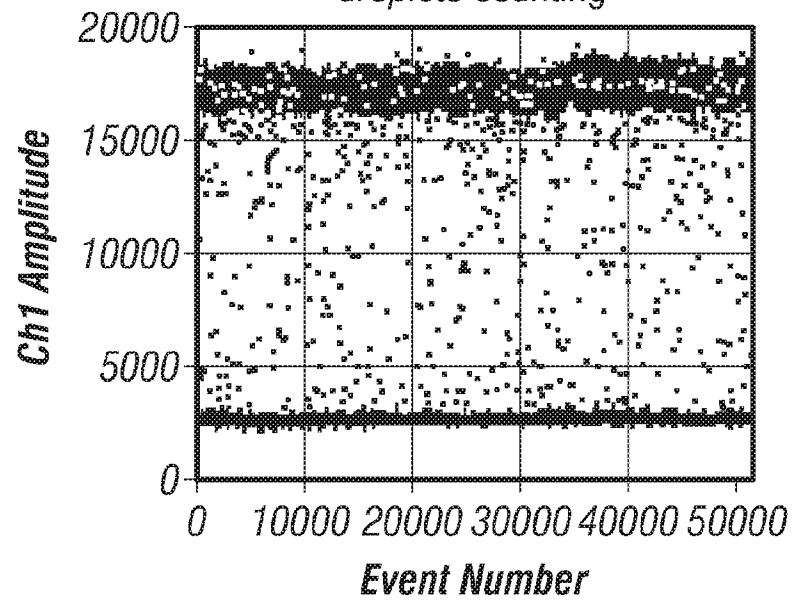


FIG. 13C

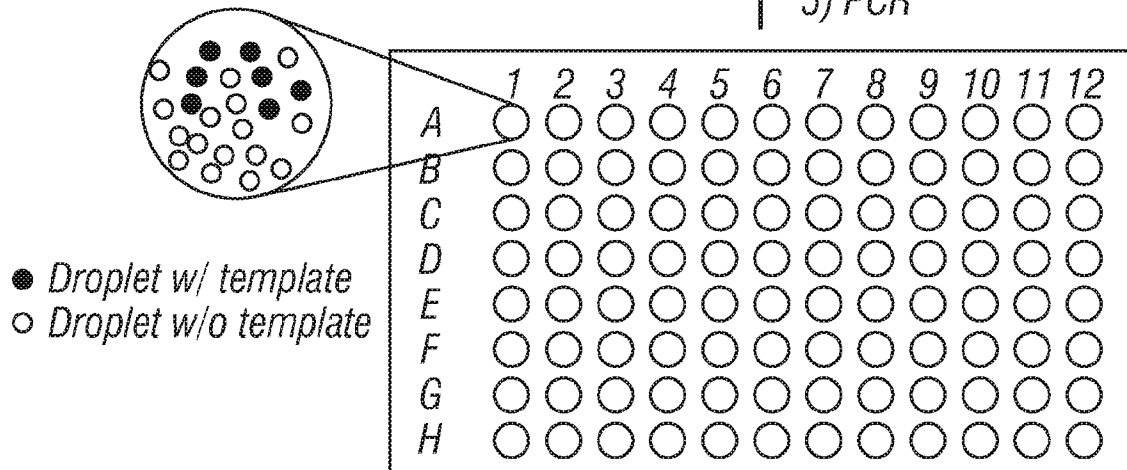
**FIG. 14**

**FIG. 15**

6) Fluorescence Reading and droplets counting



↑ 5) PCR



↑ 4) Droplet formation

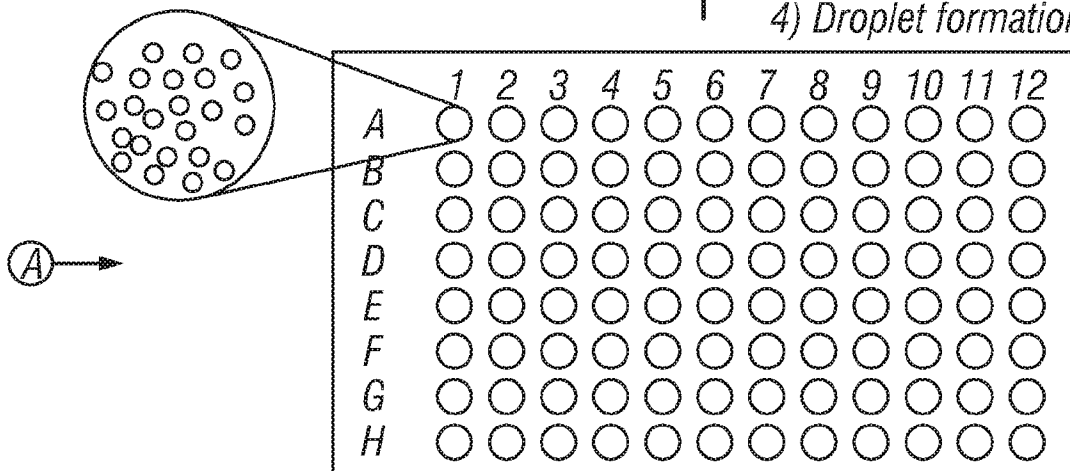


FIG. 15
(Cont'd)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2017/031346

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12N 15/10; C12P 19/34; C12Q 1/68 (2017.01)

CPC - C12N 15/1068; C12Q 1/6827; C12Q 1/6841; C12Q 1/6853; C12Q 1/6876; C12Q 2535/131;
C12Q 2600/156 (2017.02)

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

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

USPC - 435/6.1; 435/6.11; 435/6.12; 506/16; 702/20 (keyword delimited)

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

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	WO 2015/094429 A1 (WILLIAM MARSH RICE UNIVERSITY) 25 June 2015 (25.06.2015) entire document	1, 2 ----- 3-6, 37-42, 51-53, 71-73
X -- Y	WO 2015/161173 A1 (WILLIAM MARSH RICE UNIVERSITY) 22 October 2015 (22.10.2015) entire document	50 ----- 7-12, 24-26, 53, 63, 73
X -- Y	WO 2016/064856 A1 (GEN9, INC.) 28 April 2016 (28.04.2016) entire document	61, 62 ----- 52, 63, 71-73
Y	US 2013/0231253 A1 (AMORESE et al) 05 September 2013 (05.09.2013) entire document	3-12, 24-26, 37-42, 51-53
A	WO 2015/010020 A1 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 22 January 2015 (22.01.2015) entire document	1-12, 24-26, 37-42, 50-53, 61-63, 71-73
A	WO 2016/065192 A1 (ROKA BIOSCIENCE, INC.) 28 April 2016 (28.04.2016) entire document	1-12, 24-26, 37-42, 50-53, 61-63, 71-73

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

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

12 July 2017

Date of mailing of the international search report

07 AUG 2017

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Blaine R. Copenheaver

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2017/031346

Box No. 1 Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

a. ☒ forming part of the international application as filed:

☒ in the form of an Annex C/ST.25 text file.

☐ on paper or in the form of an image file.

b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

c. ☐ furnished subsequent to the international filing date for the purposes of international search only:

☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2017/031346

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 13-23, 27-36, 43-49, 54-60, 64-70, 74-81
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.