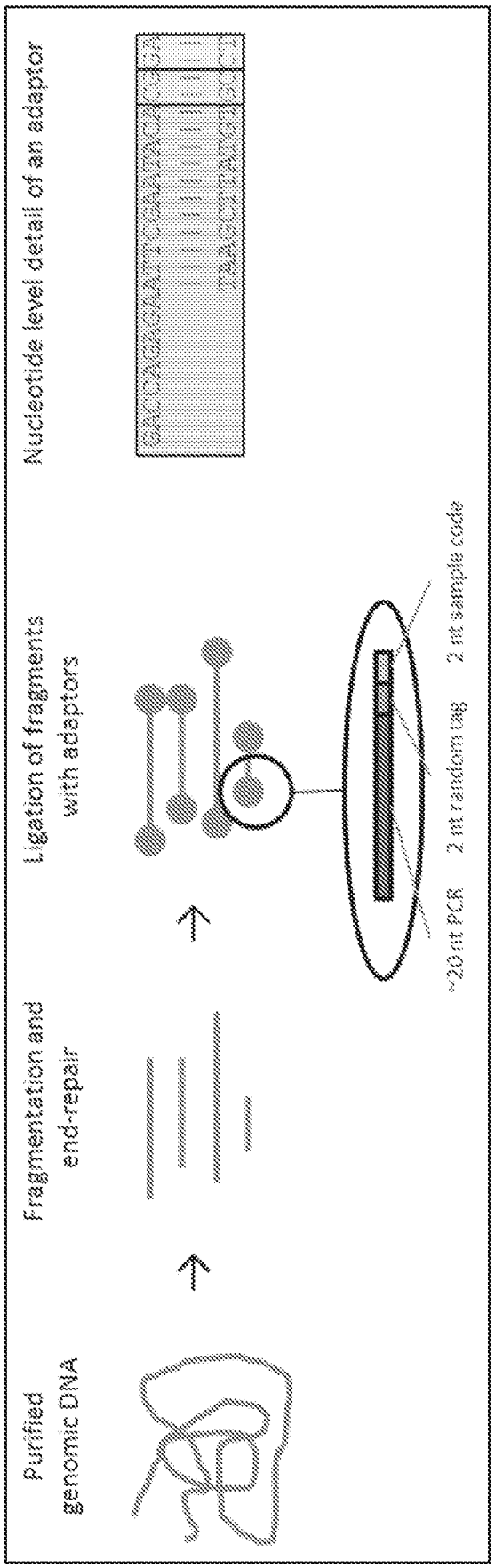

Purified genomic DNA
Fragmentation and end-repair
Ligation of fragments with adaptors
Nucleotide level detail of an adaptor
~20 nt PCR
2 nt random tag
2 nt sample code

FIG. 1

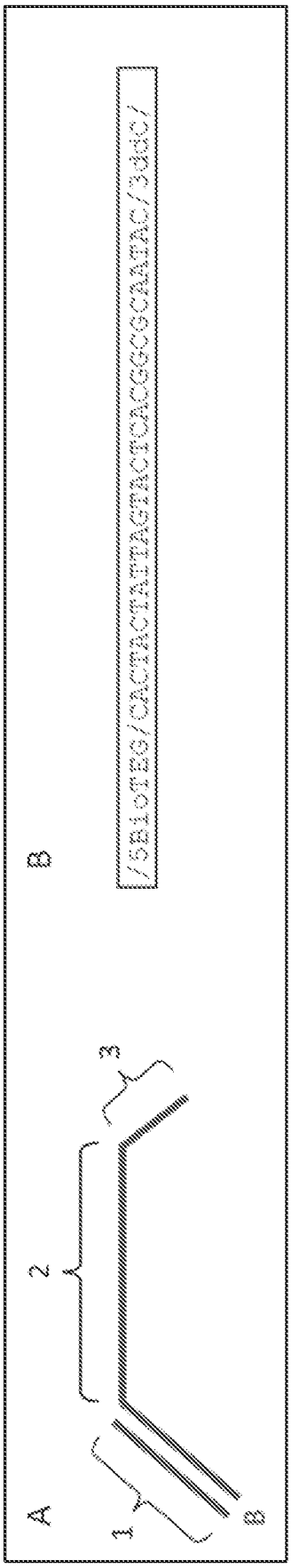


FIG. 2

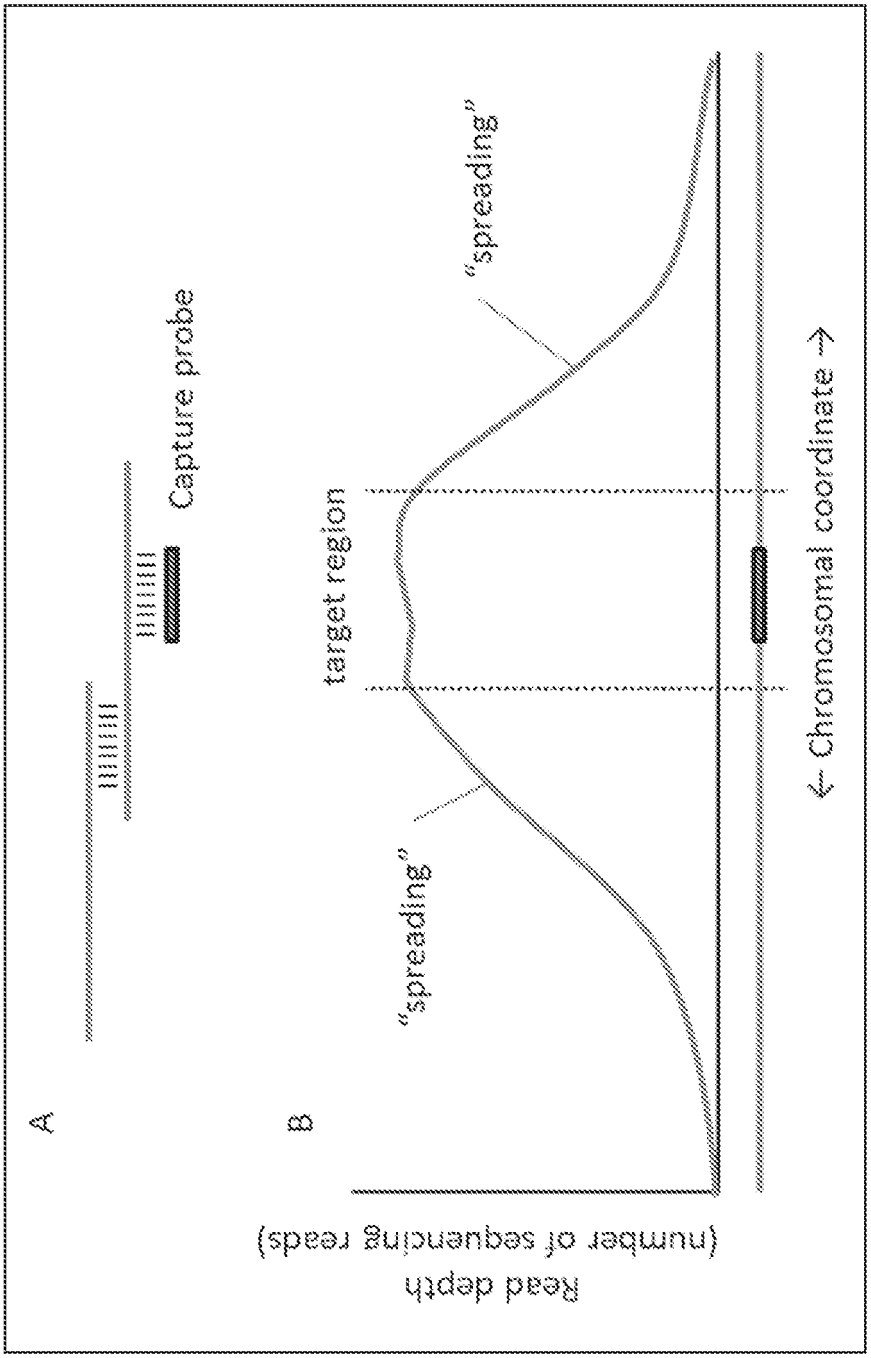

A
Capture probe
B
target region
"spreading"
"spreading"
Read depth
(number of sequencing reads)
← Chromosomal coordinate →

FIG. 3

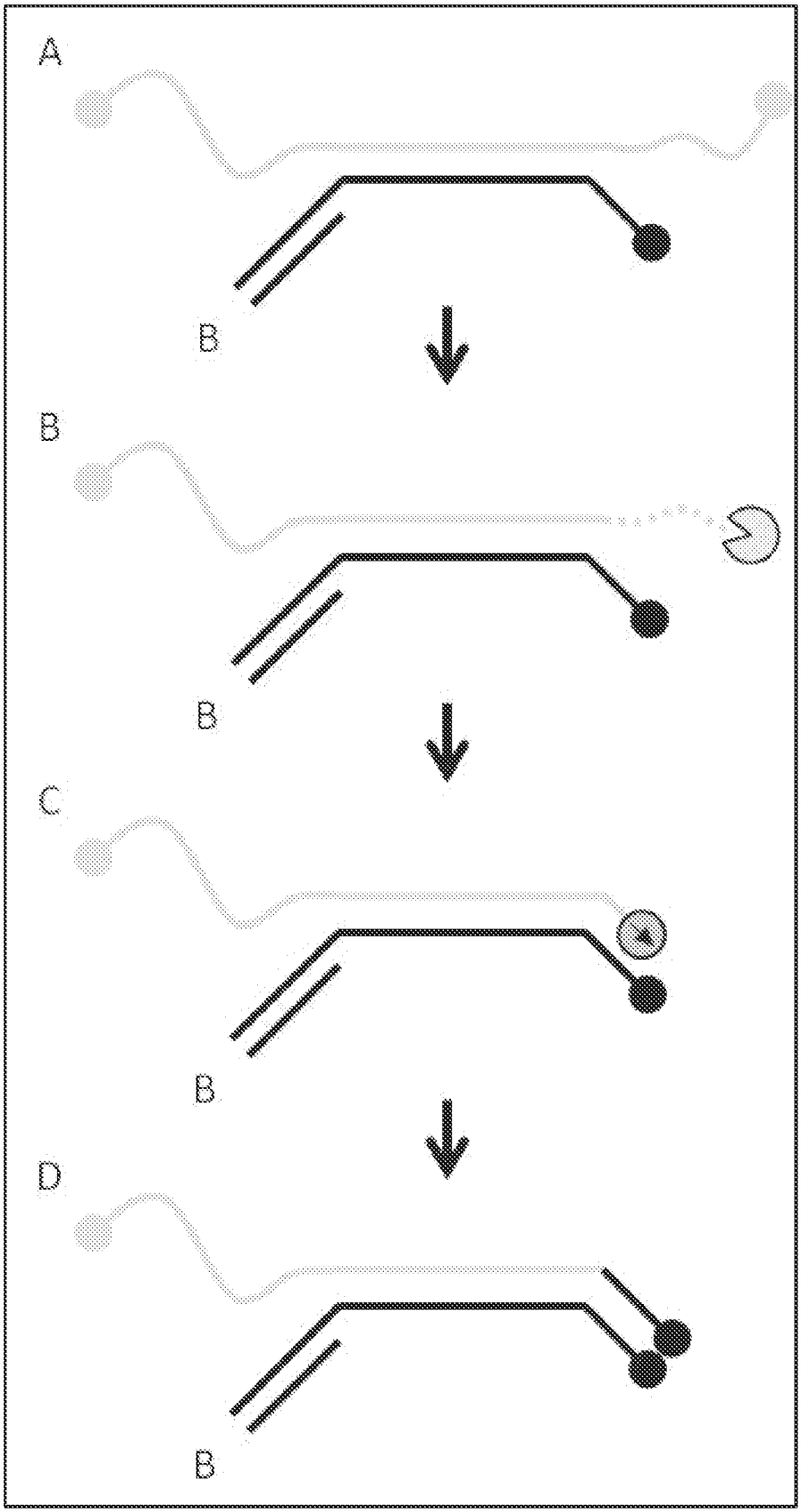


FIG. 4

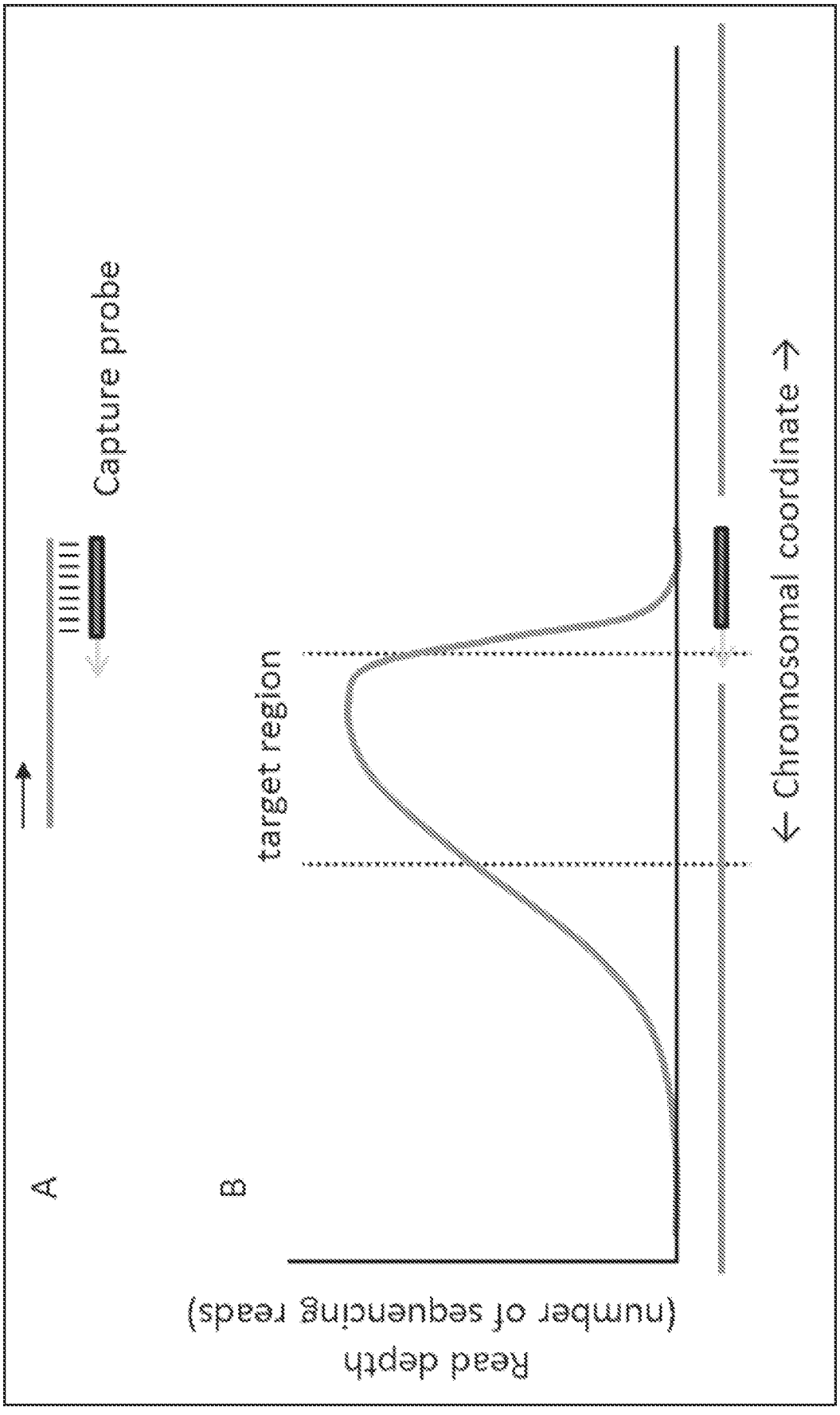

A
B
Capture probe
target region
Read depth
(number of sequencing reads)
← Chromosomal coordinate →

FIG. 5

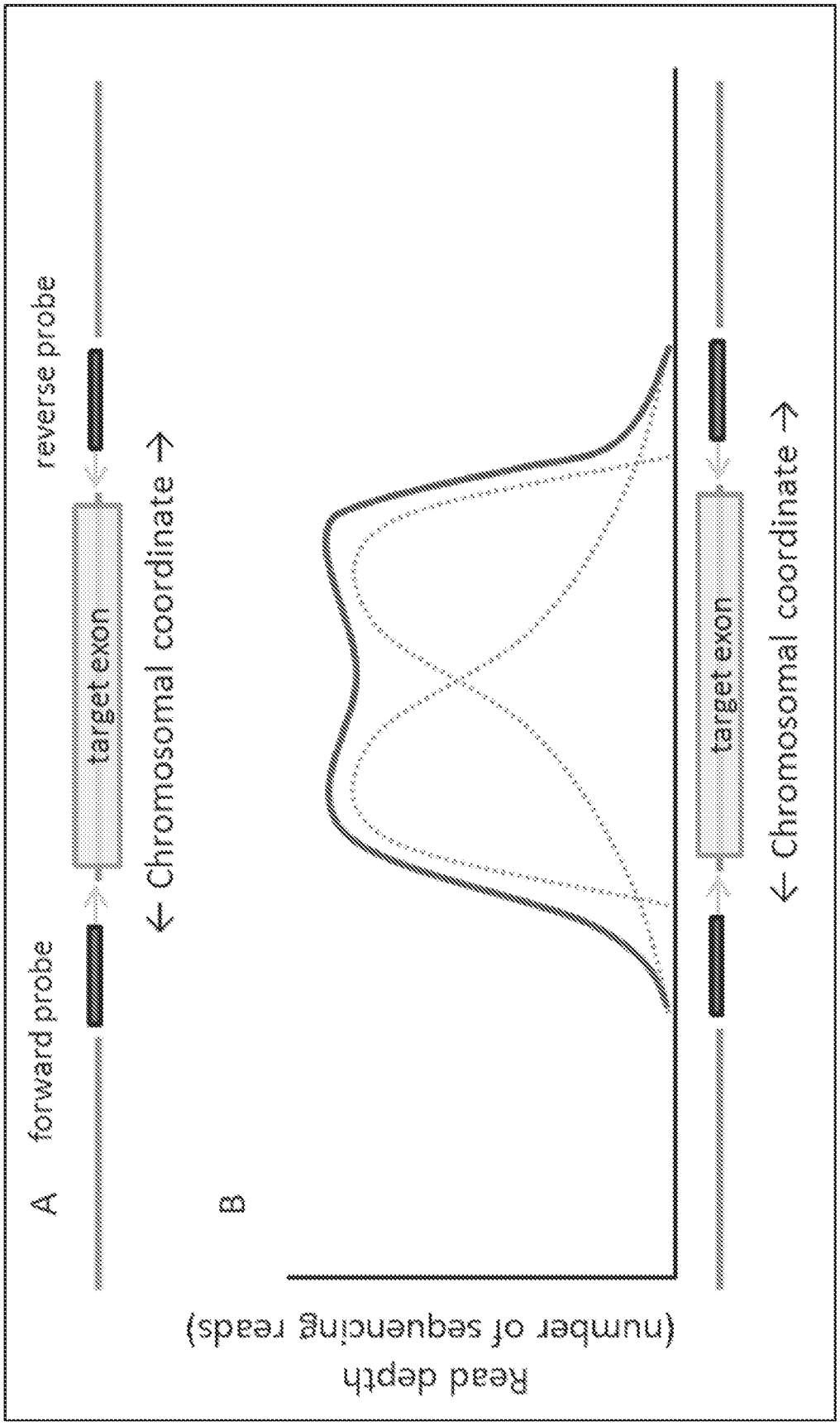

A
forward probe
reverse probe
target exon
← Chromosomal coordinate →
B
Read depth
(number of sequencing reads)
target exon
← Chromosomal coordinate →

FIG. 6

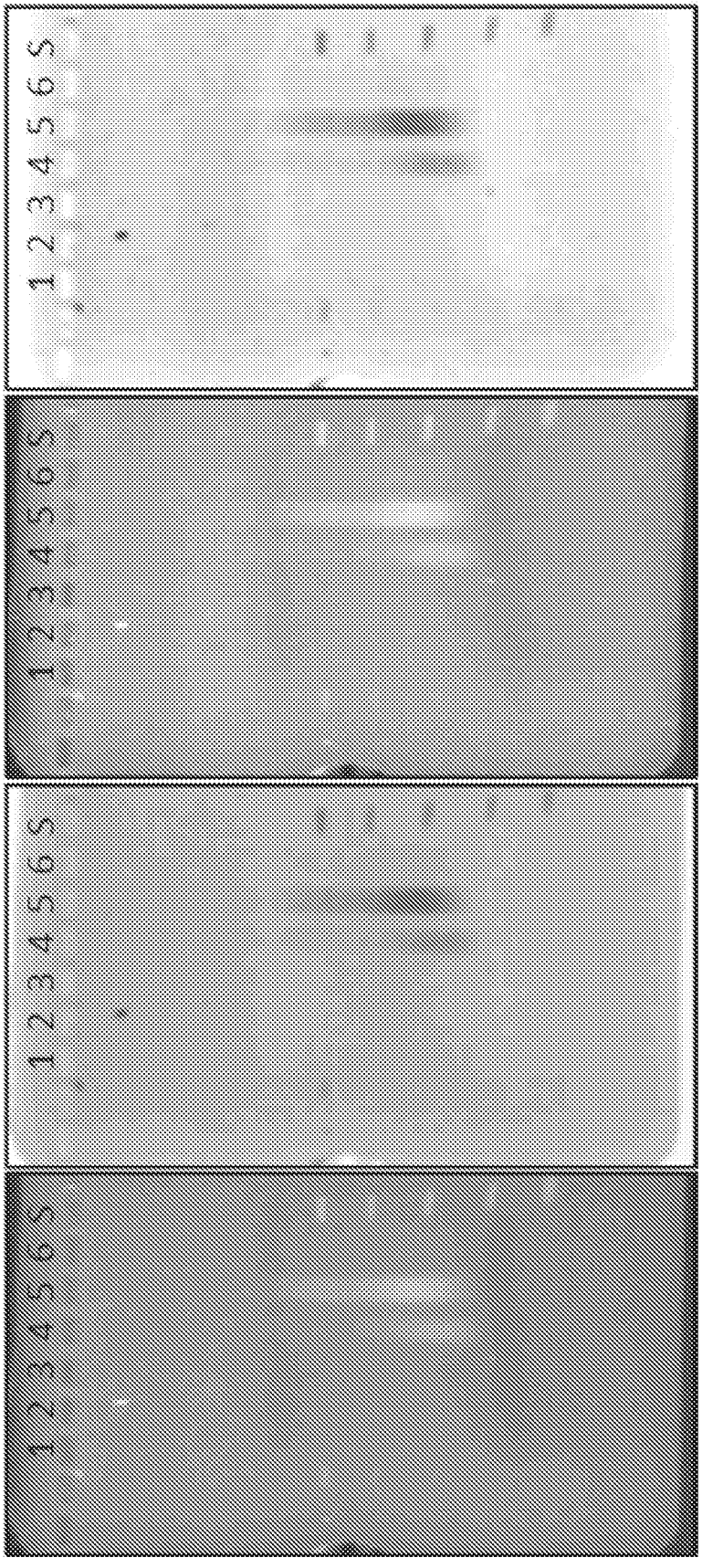


FIG. 7

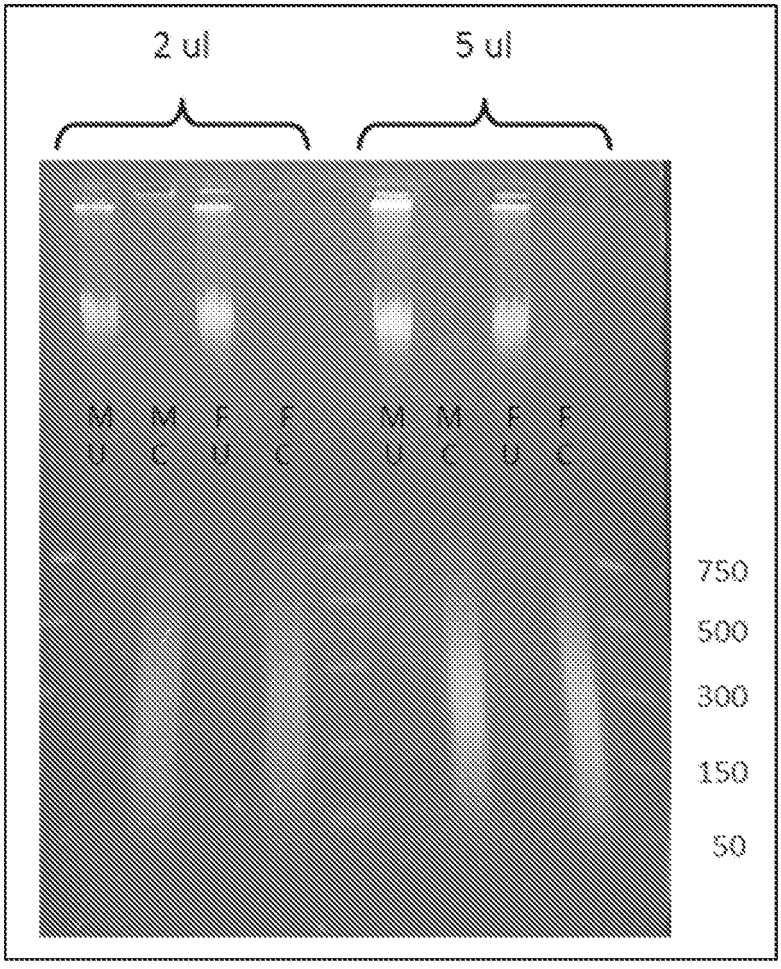


FIG. 8

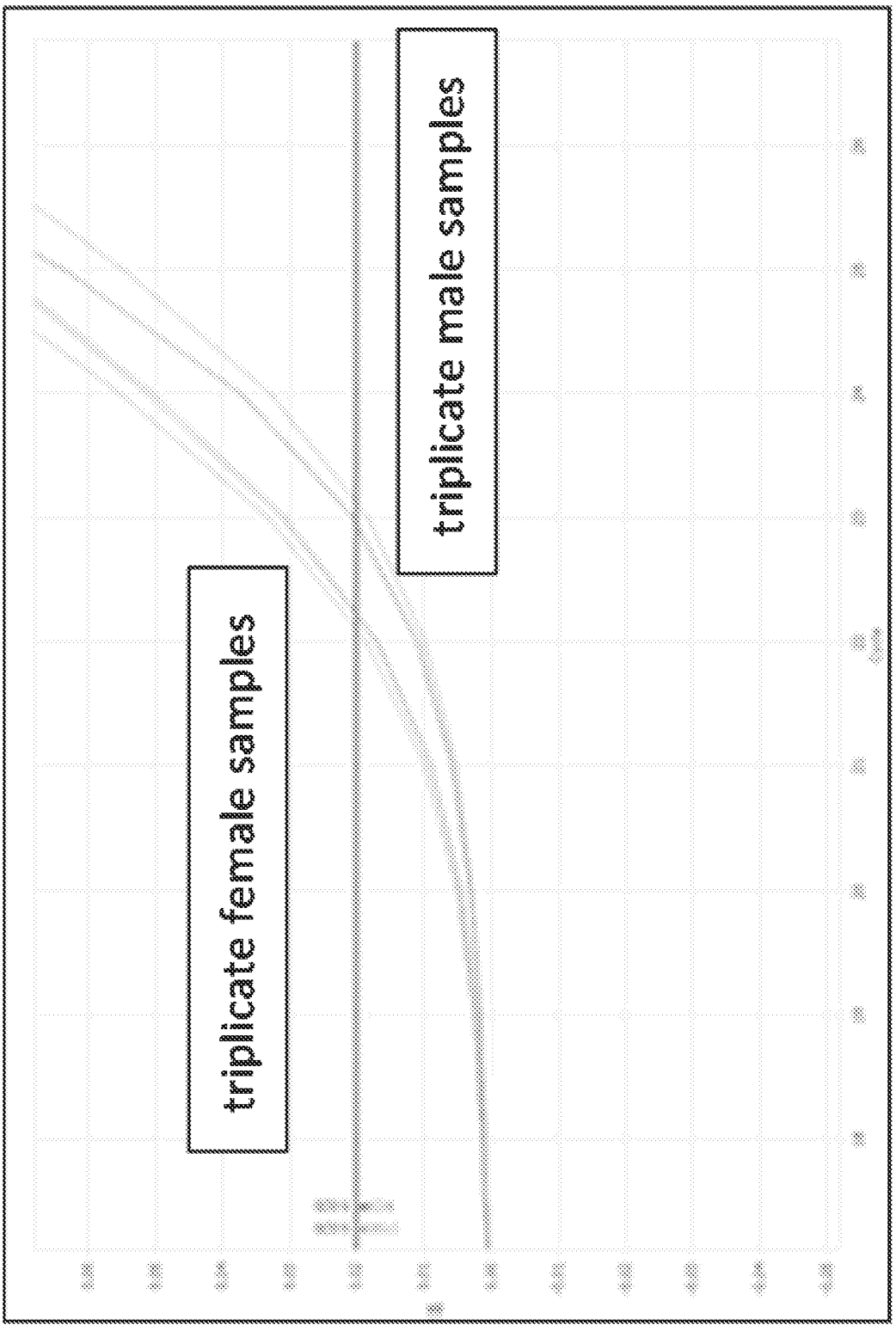


FIG. 9

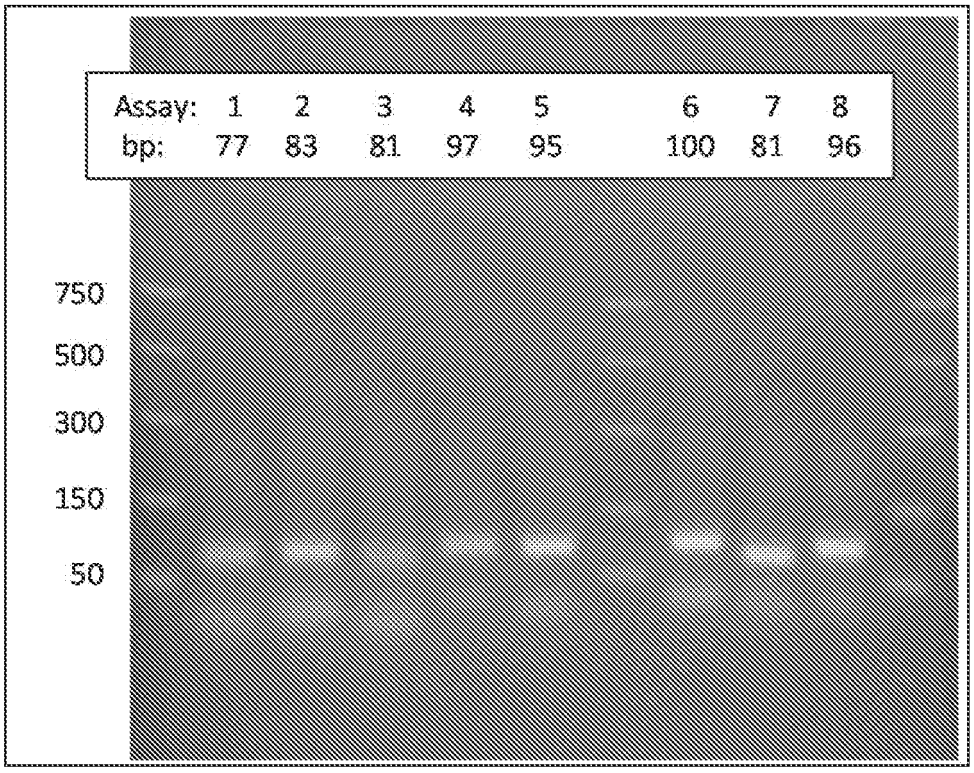


FIG. 10

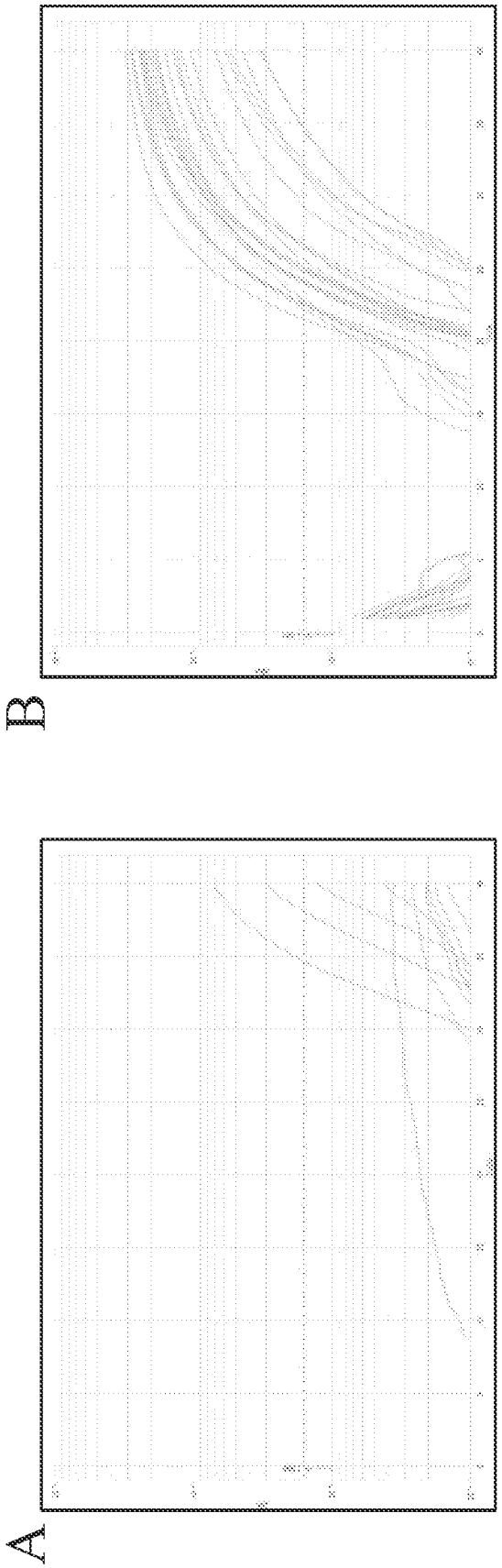


FIG. 11

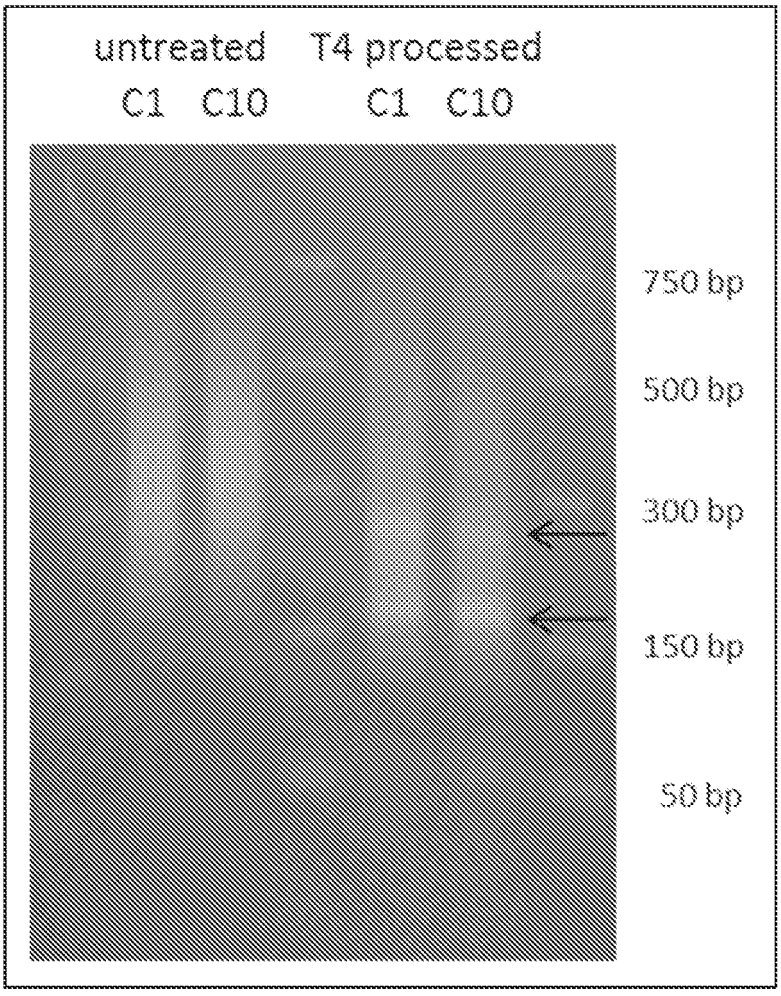

untreated
T4 processed
C1
C10
C1
C10
750 bp
500 bp
300 bp
150 bp
50 bp

FIG. 12

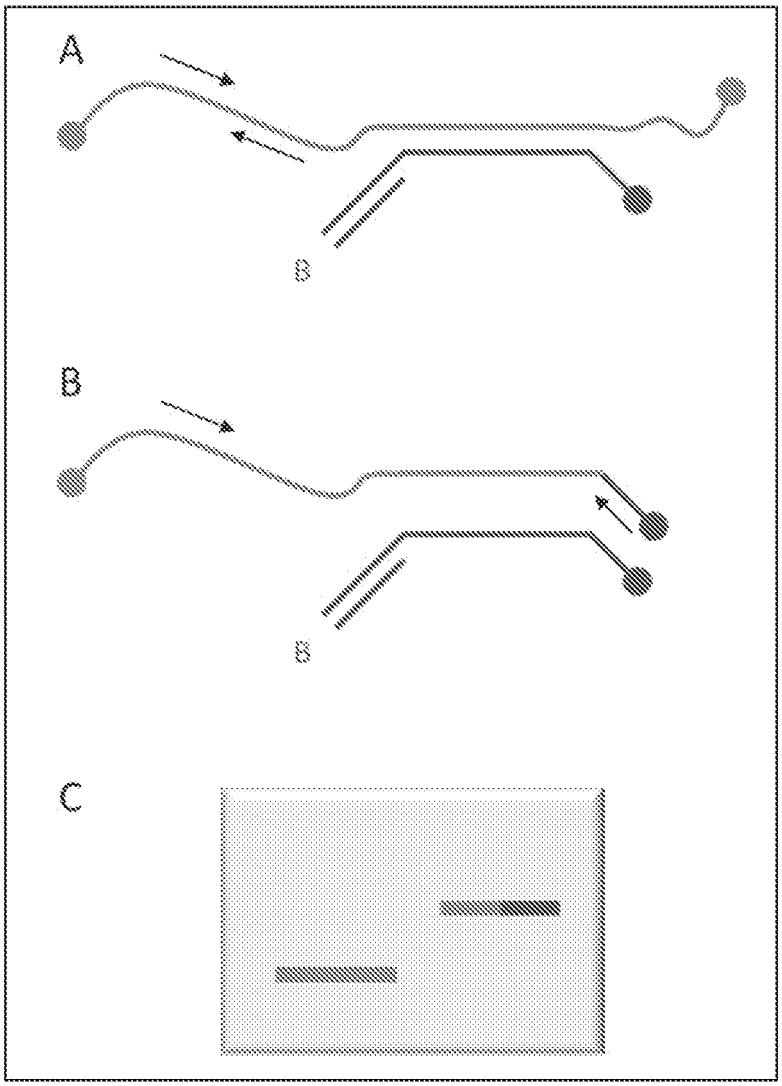


FIG. 13

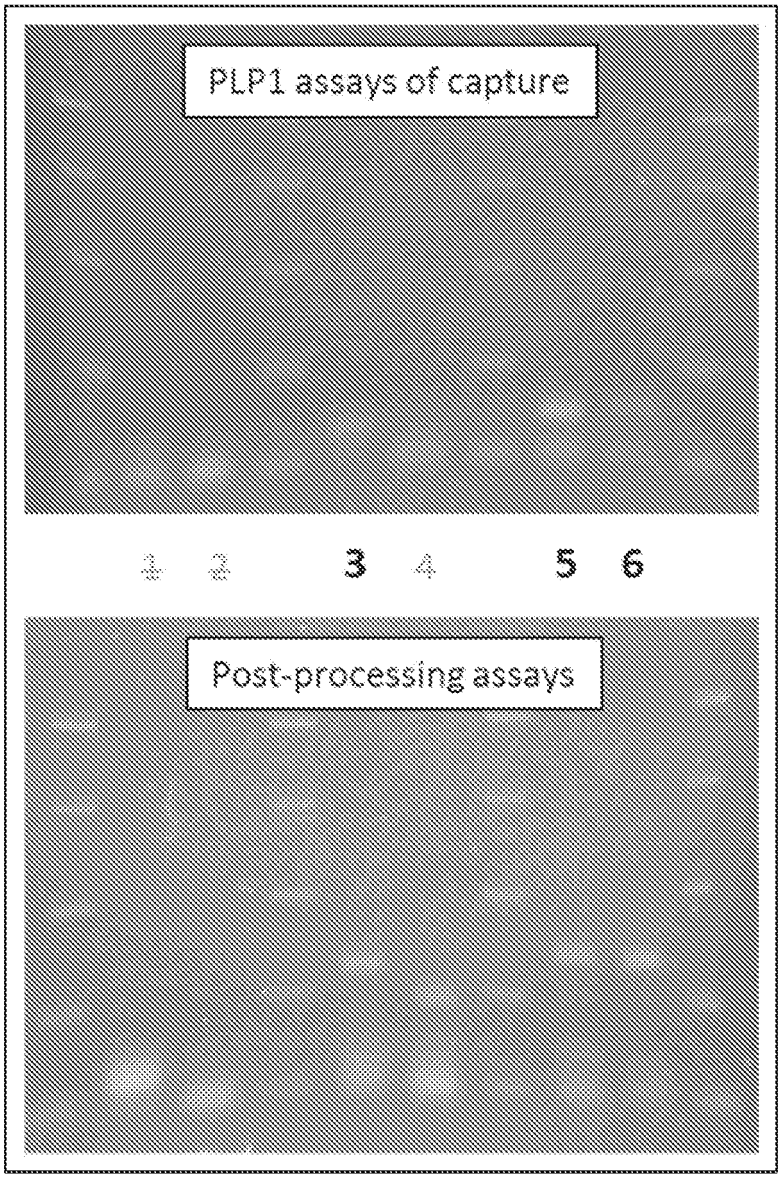


FIG. 14

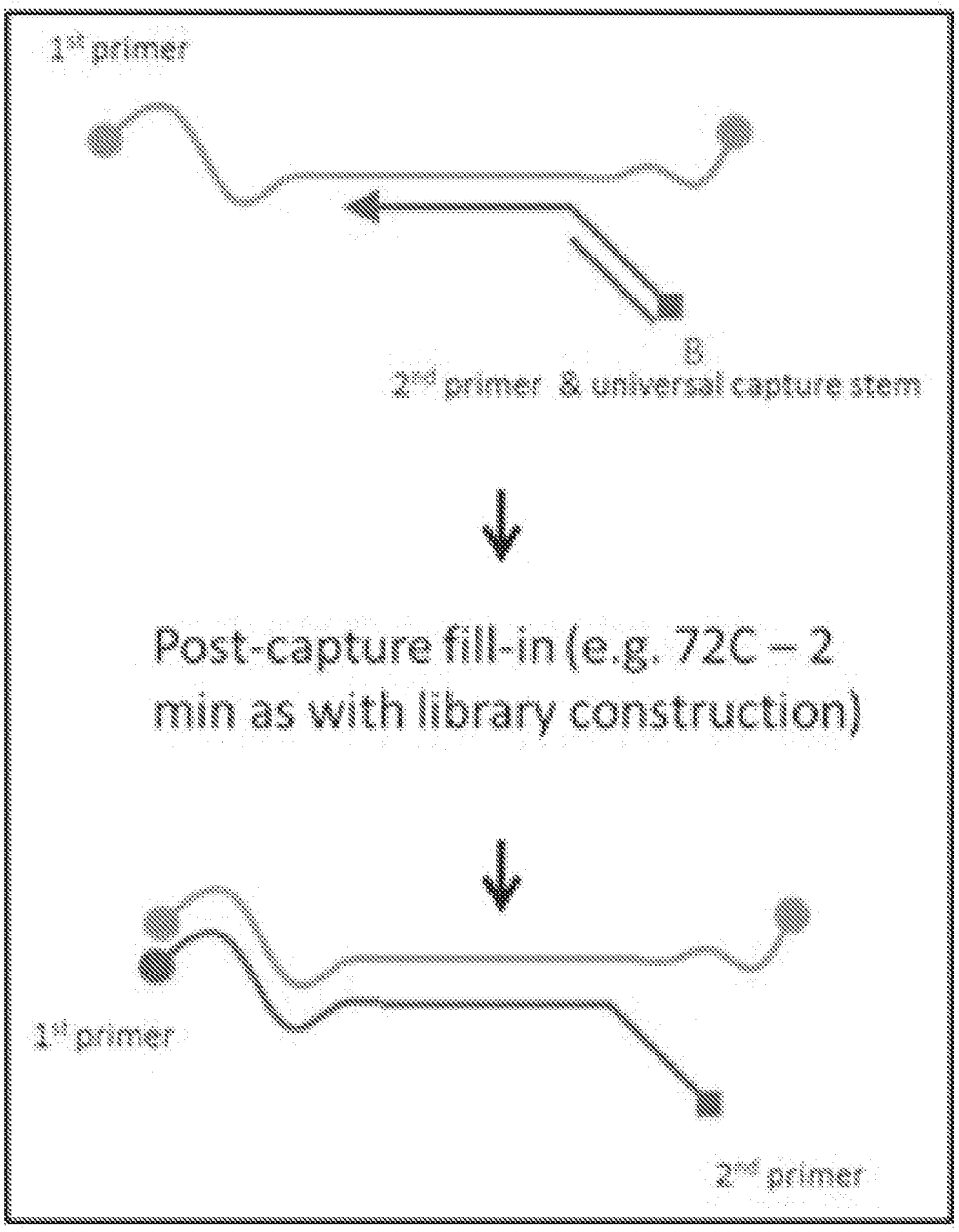

1st primer
B
2nd primer & universal capture stem
Post-capture fill-in (e.g. 72C – 2 min as with library construction)
1st primer
2nd primer

FIG. 15

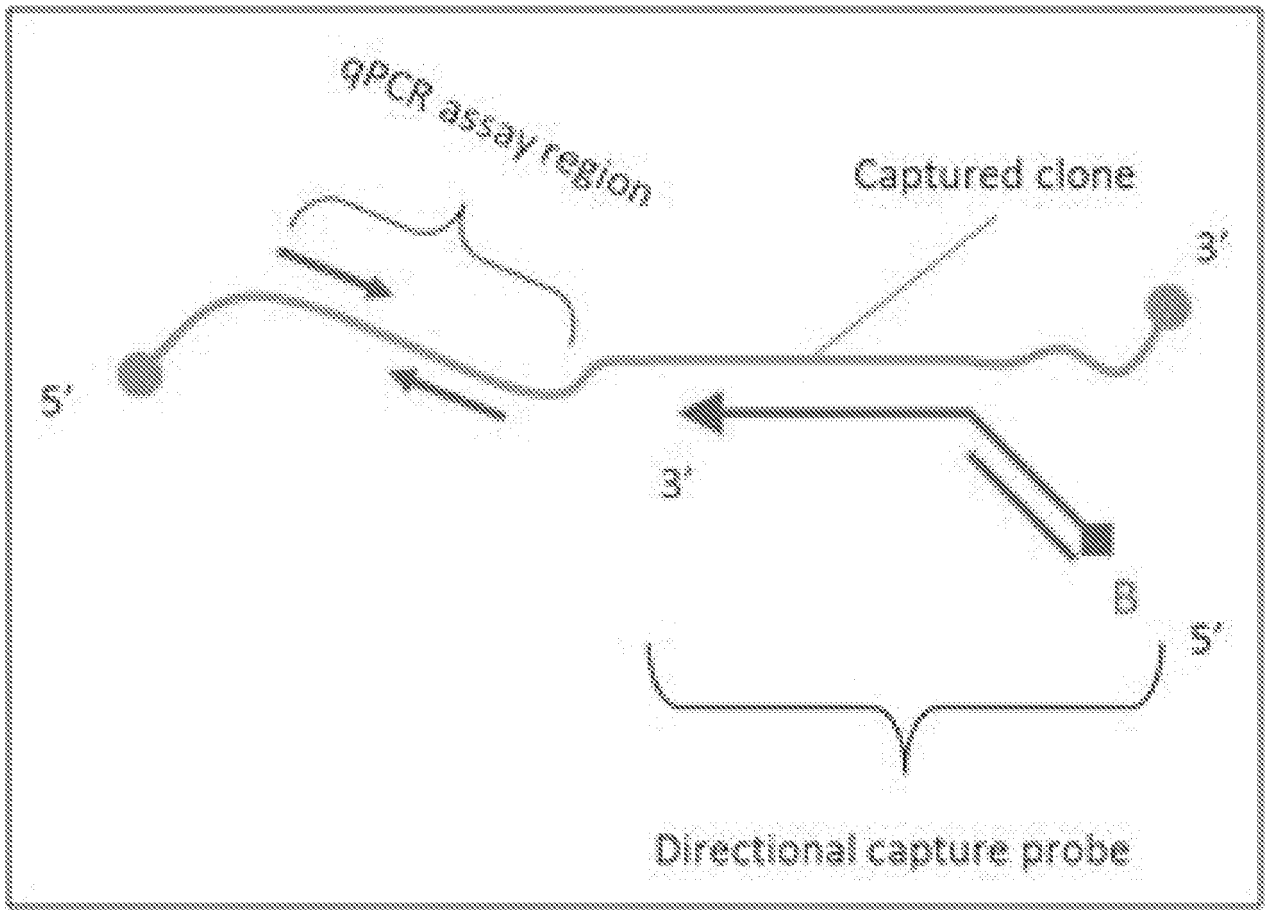

qPCR assay region
Captured clone
3'
5'
3'
B
5'
Directional capture probe

FIG. 16

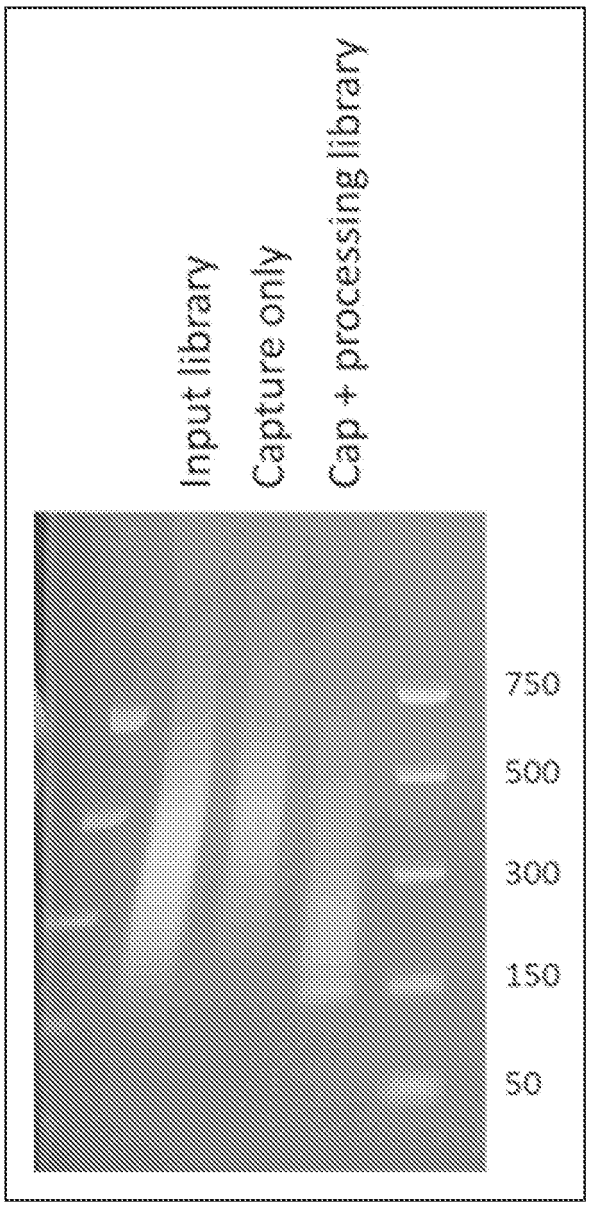

Input library
Capture only
Cap + processing library
750
500
300
150
50

FIG. 17

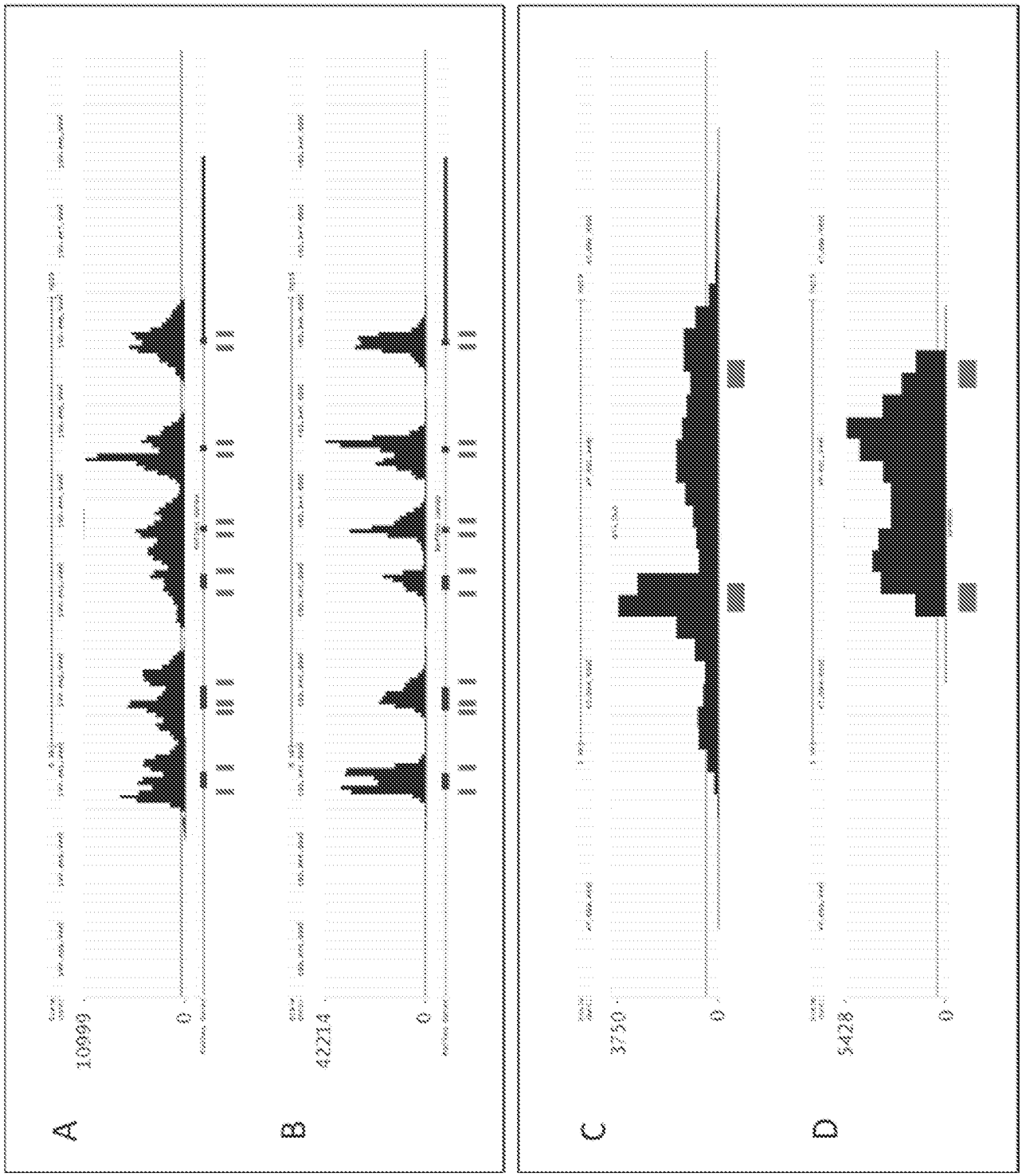


FIG. 18

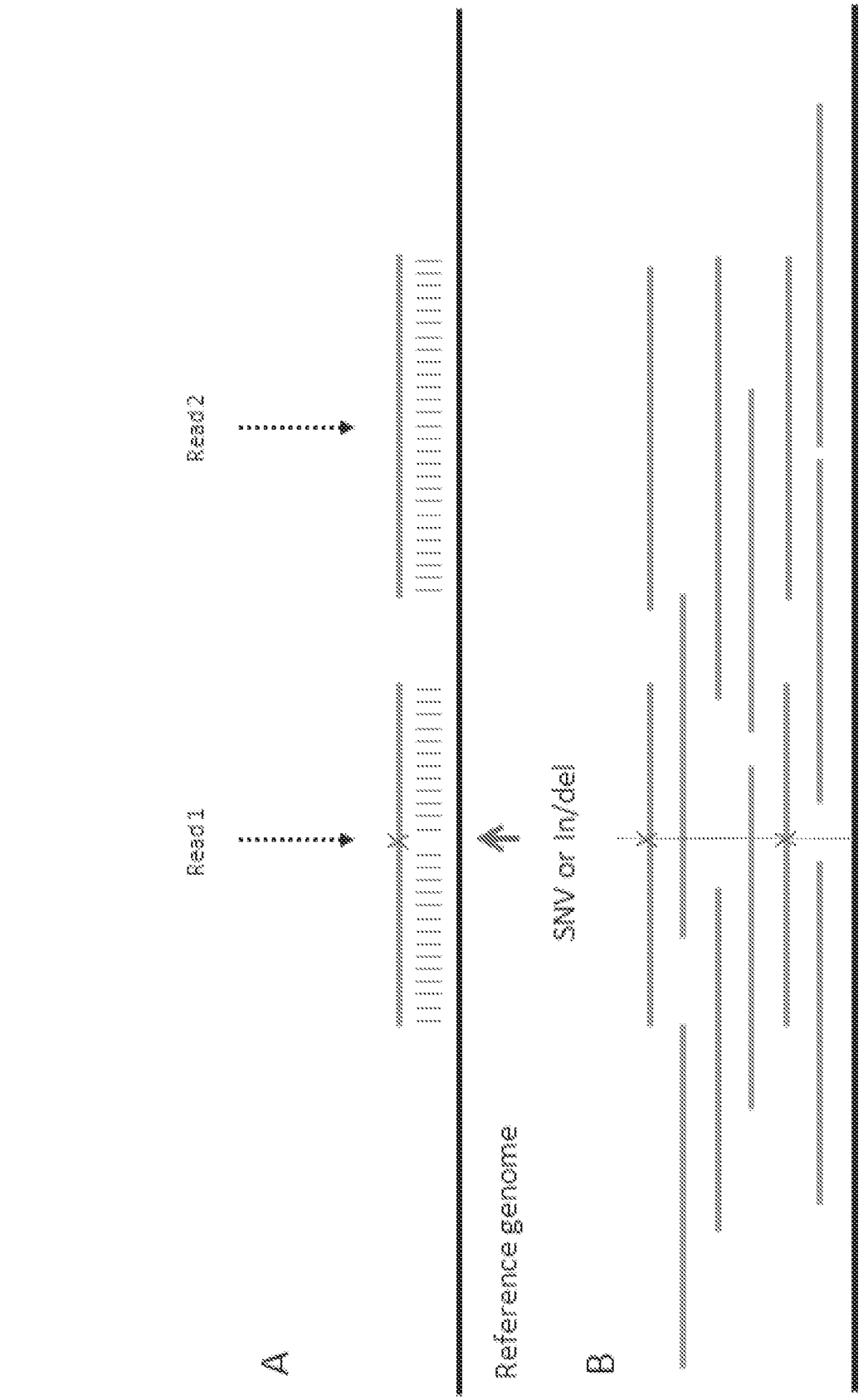

A
Read 1
Read 2
Reference genome
SNV or In/del
B
Reference genome

FIG. 19

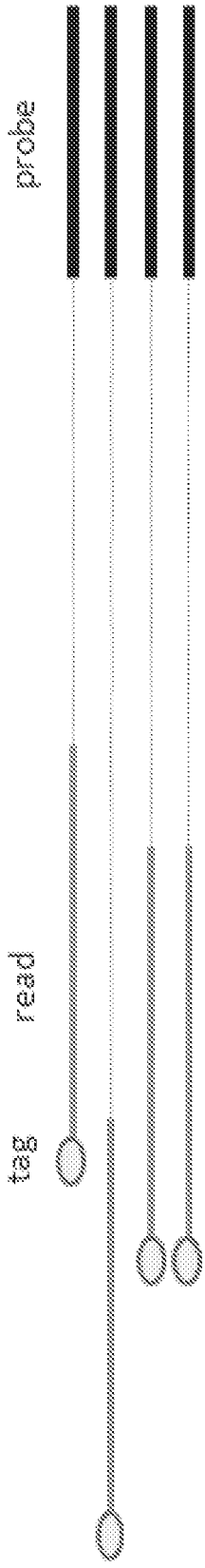


FIG. 20

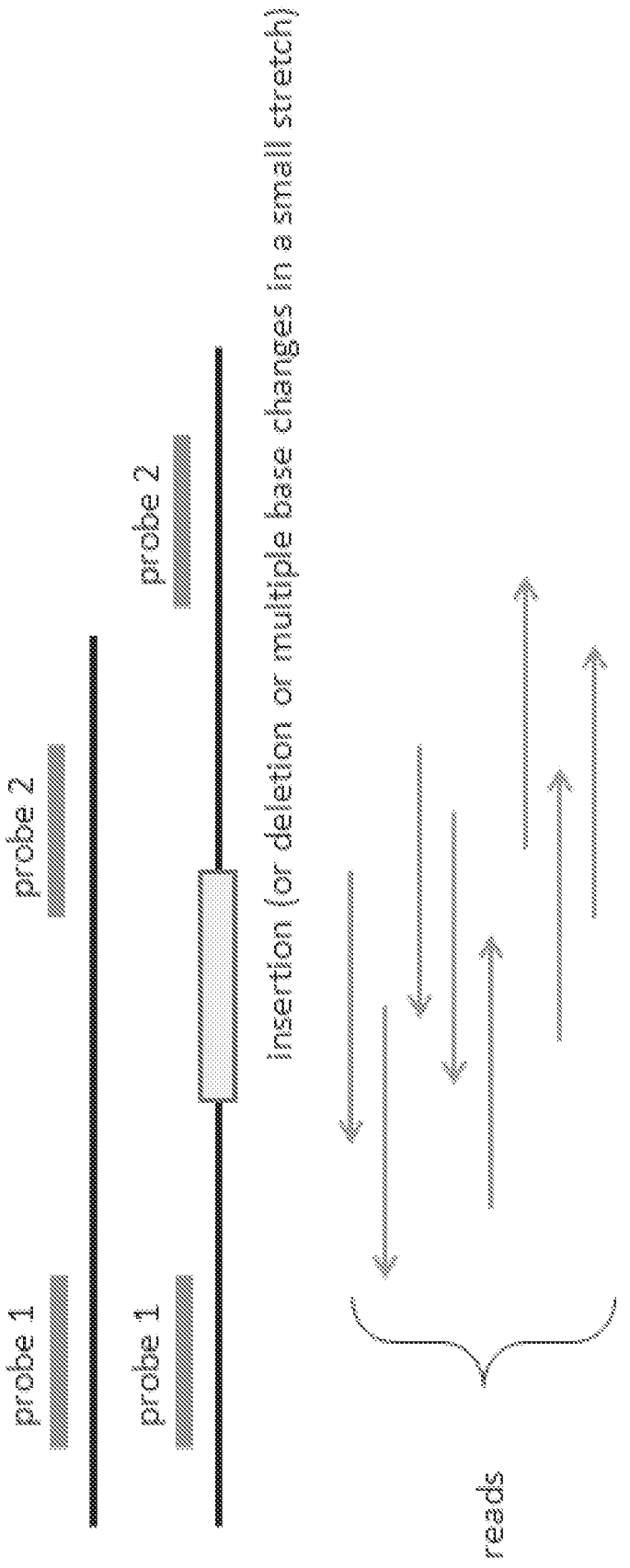

probe 1
probe 2
probe 1
probe 2
insertion (or deletion or multiple base changes in a small stretch)
reads

FIG. 21

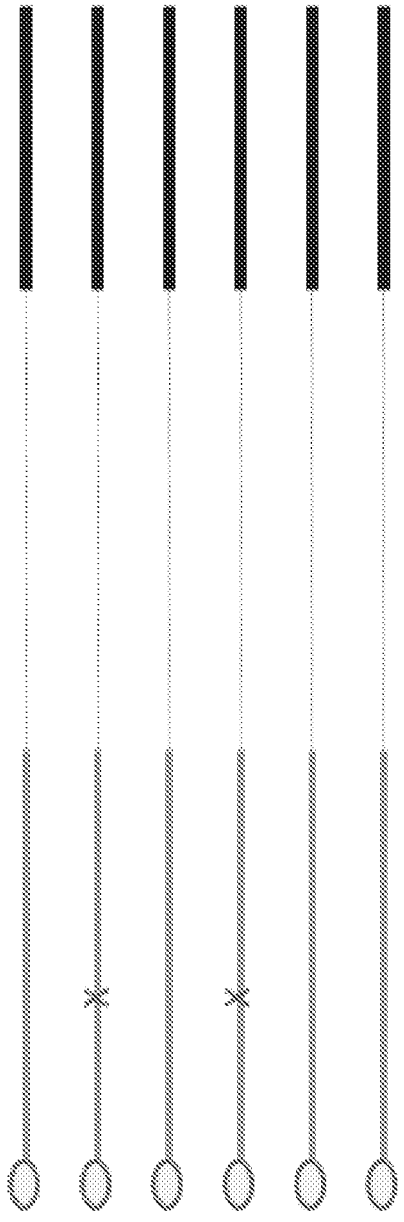

probe
same sequence start site
- same adaptor-introduced code sequence (e.g. "GTA")

FIG. 22

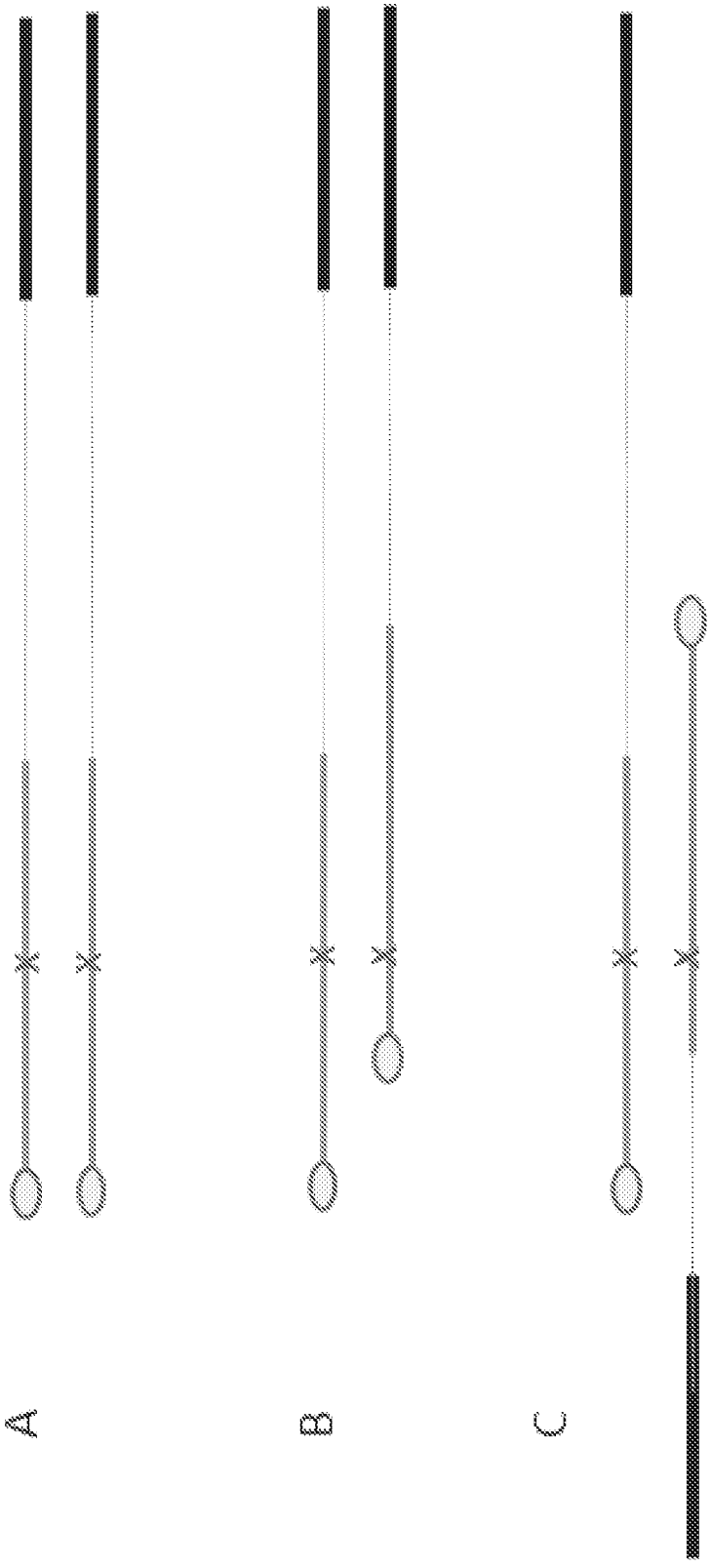


FIG. 23

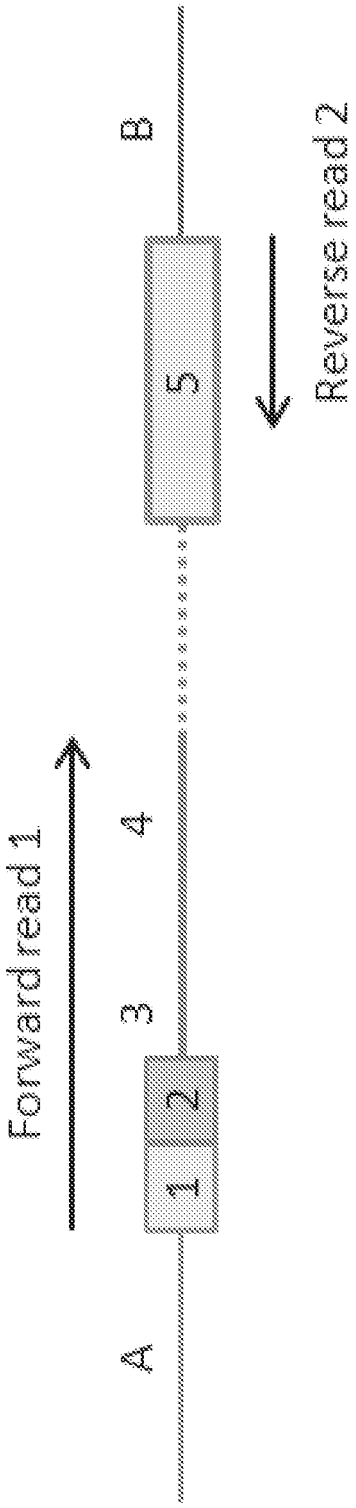


FIG. 24

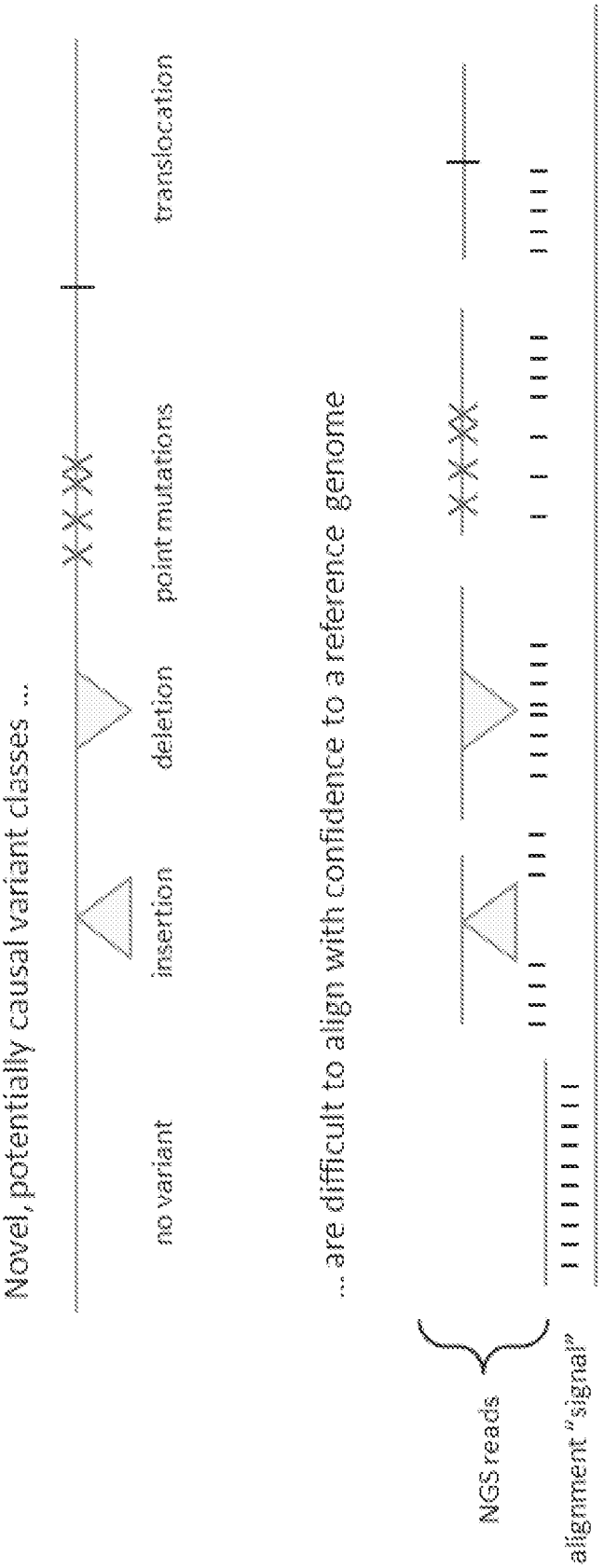

Novel, potentially causal variant classes ...
no variant
insertion
deletion
point mutations
translocation
...are difficult to align with confidence to a reference genome
NGS reads
alignment "signal"

FIG. 25

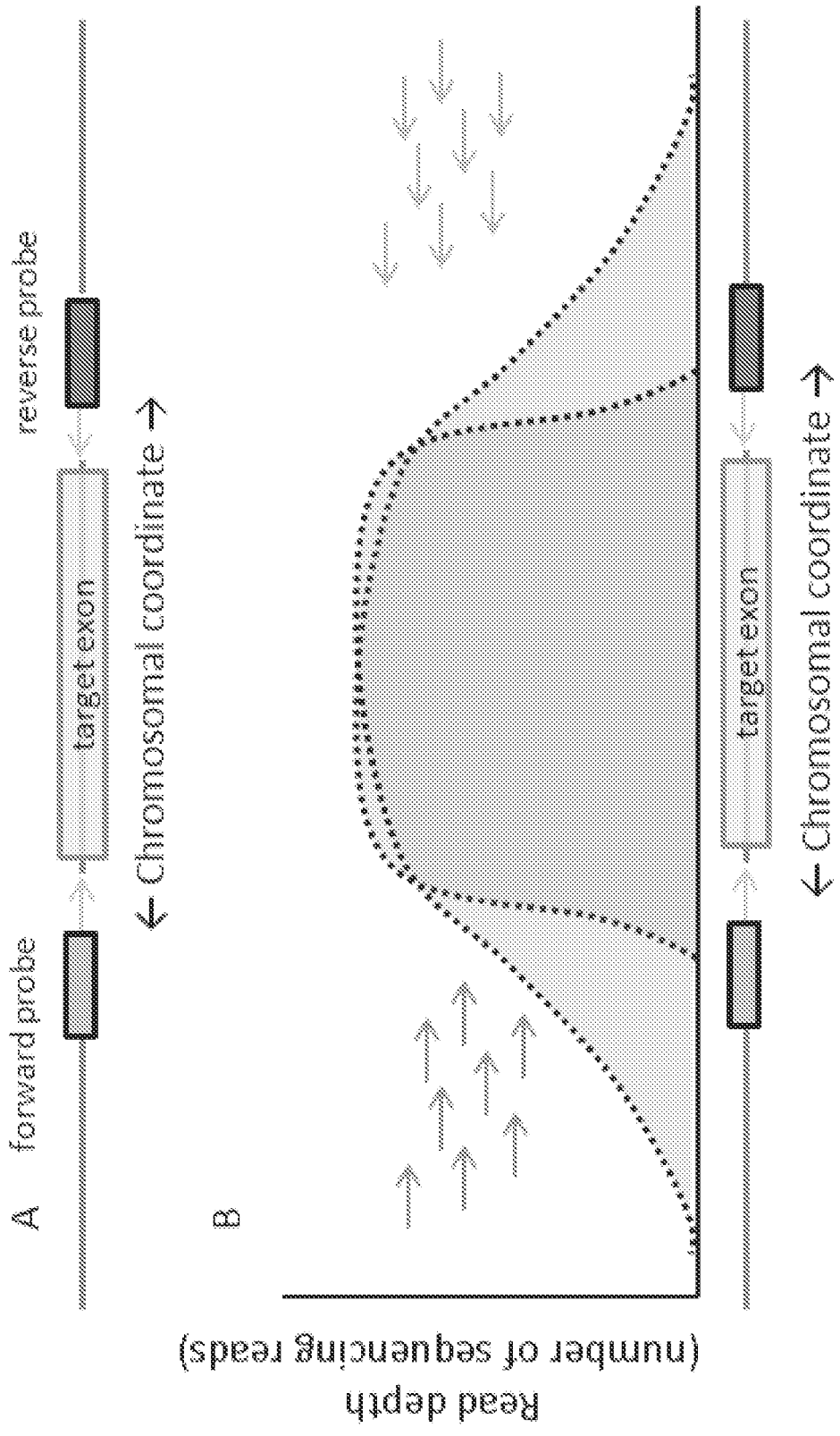

A
forward probe
reverse probe
target exon
← Chromosomal coordinate →
B
Read depth
(number of sequencing reads)
target exon
← Chromosomal coordinate →

FIG. 26

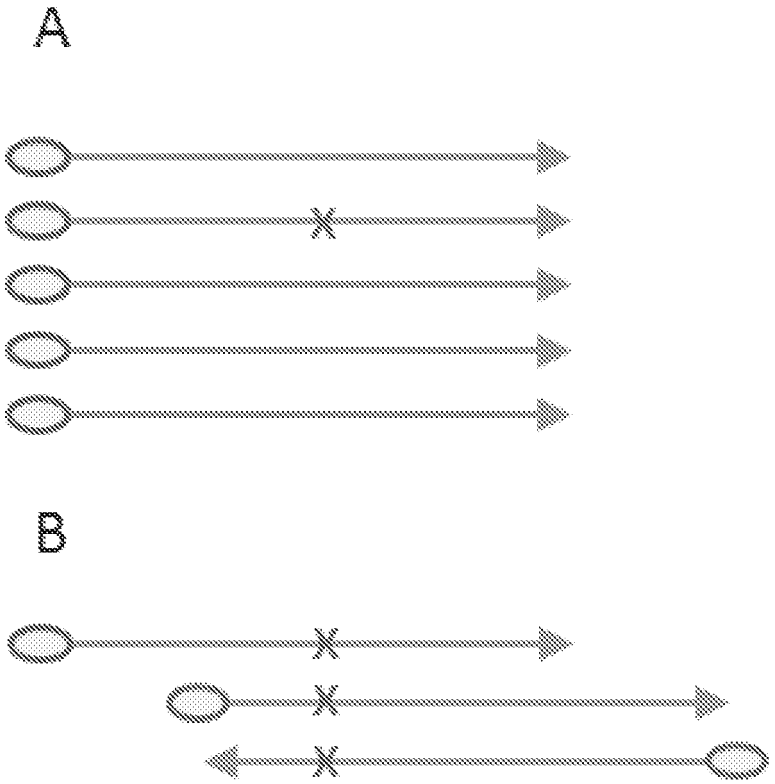
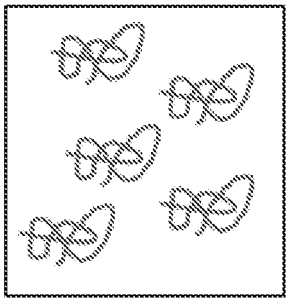
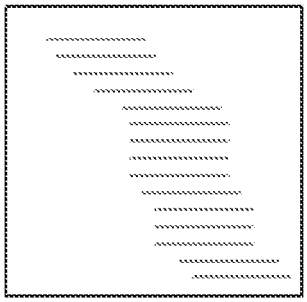


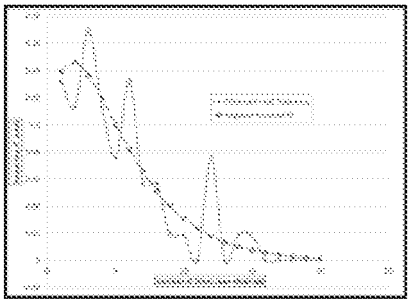
FIG. 27



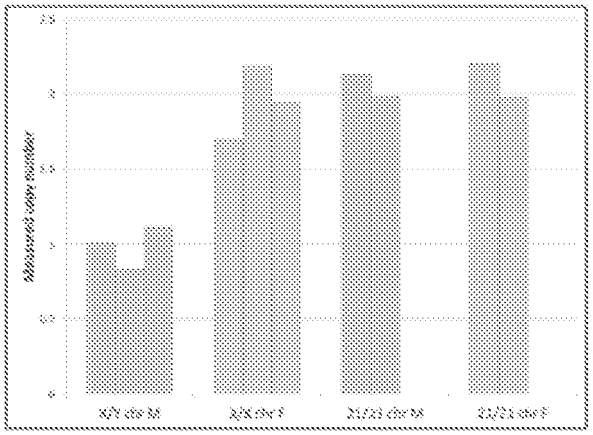
Low input genome libraries (e. g. ~50-100 genome equivalents) created with molecular annotation and subsequently captured with targeting probes.



Variant filtering for true positive read set on a probe-by-probe basis. This allows the number and frequency of observed unique reads to be determined.



Observed read frequencies are fitted to a negative binomial distribution, which allows the "inferred class" of unobserved reads to be estimated. The number of observed reads + inferred reads = copy number.



Representative copy determination at three probes for the X-chromosomal gene PLP1 in a male and a female sample. For comparison, two probes for the autosomal chromosome 21 gene TNFRSF14 gene are also shown.

FIG. 28

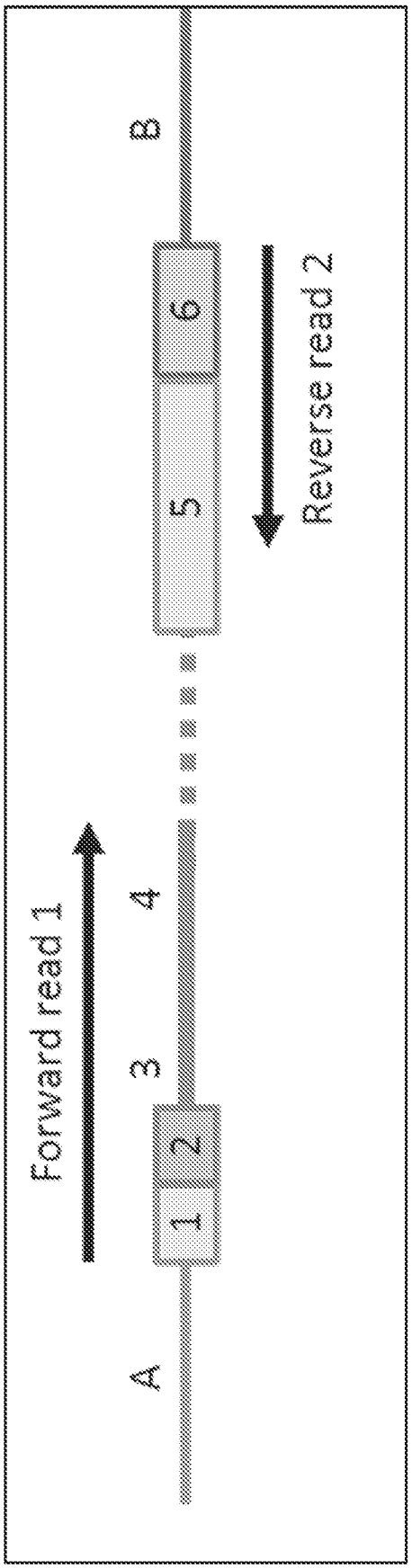


FIG. 29

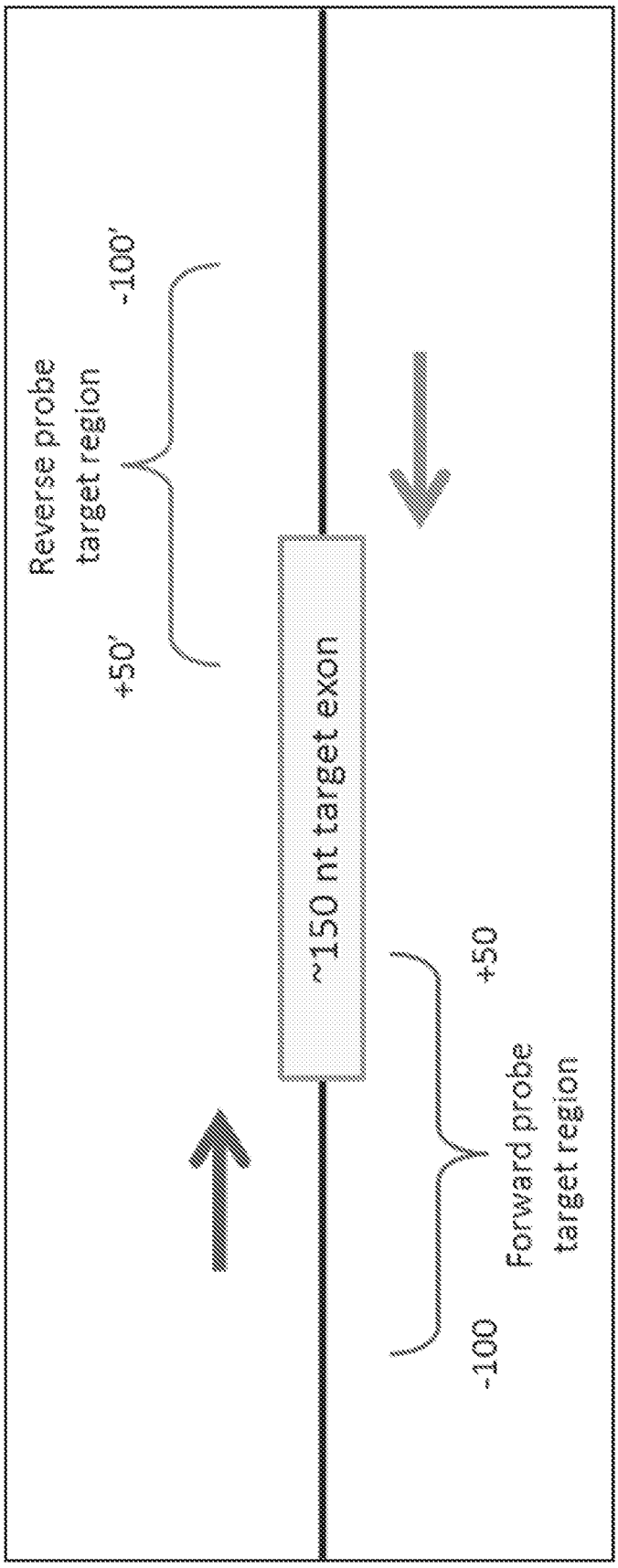

Reverse probe target region
-100'
+50'
~150 nt target exon
+50
Forward probe target region
-100

FIG. 30

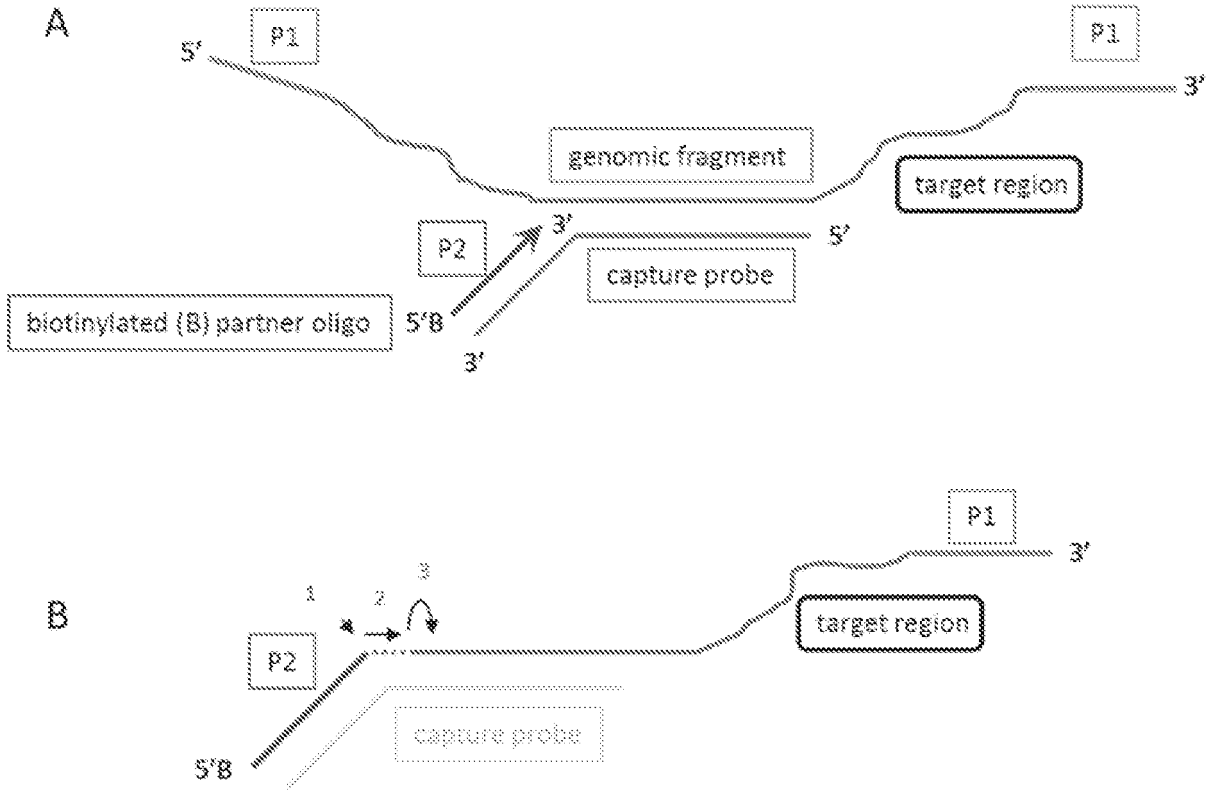

A
P1
5'
P1
3'
genomic fragment
target region
P2
3'
5'
capture probe
biotinylated (B) partner oligo
5'B
3'
B
P1
3'
1
2
3
target region
P2
capture probe
5'B

FIG. 31

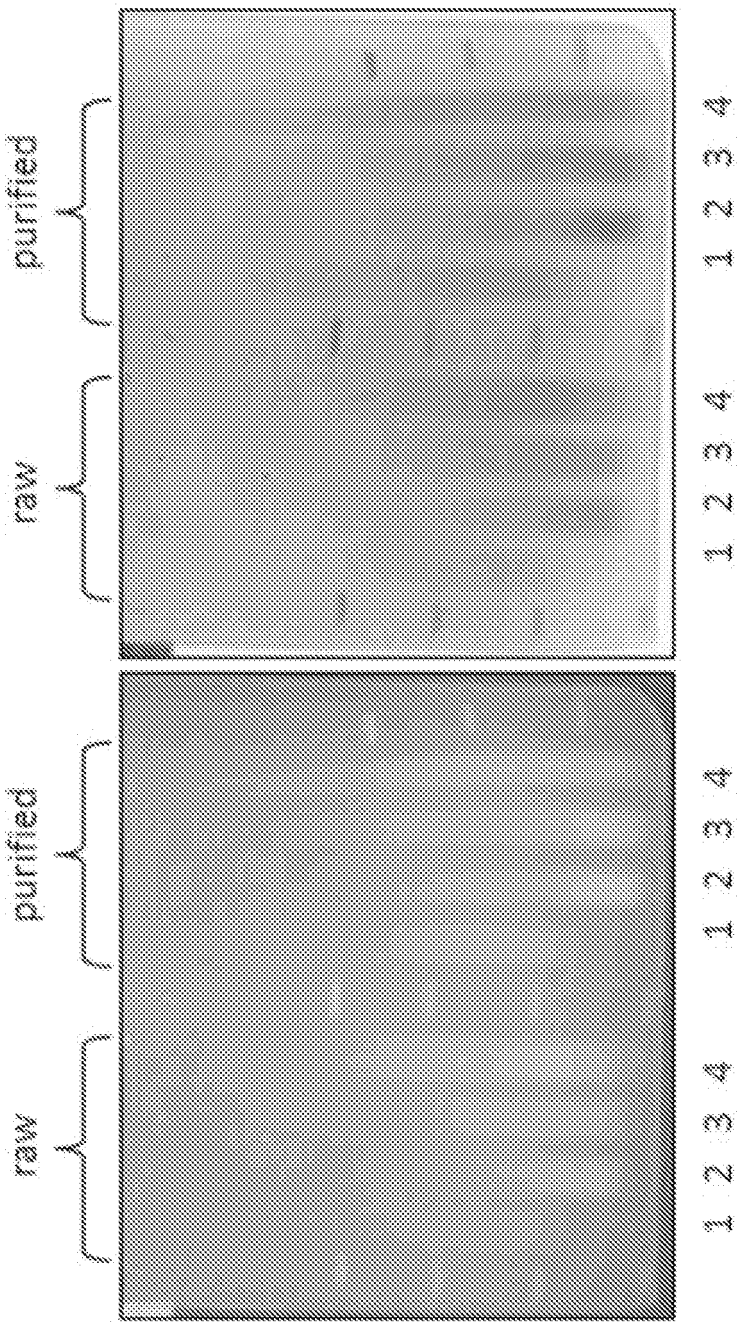


FIG. 32

Step 1 —
fragment gDNA to ~400 bp
or
perform 1st strand cDNA synthesis with (N)15 in the presence of ddNTPs
please see slide 4 for special comments about cDNA profiling
Step 2 —
melt gDNA
anneal gDNA or cDNA with labeled capture probes, purify
gDNA or 1st strand cDNA
5'
3'
3'
probe
P2
5'
Biotin

FIG. 33

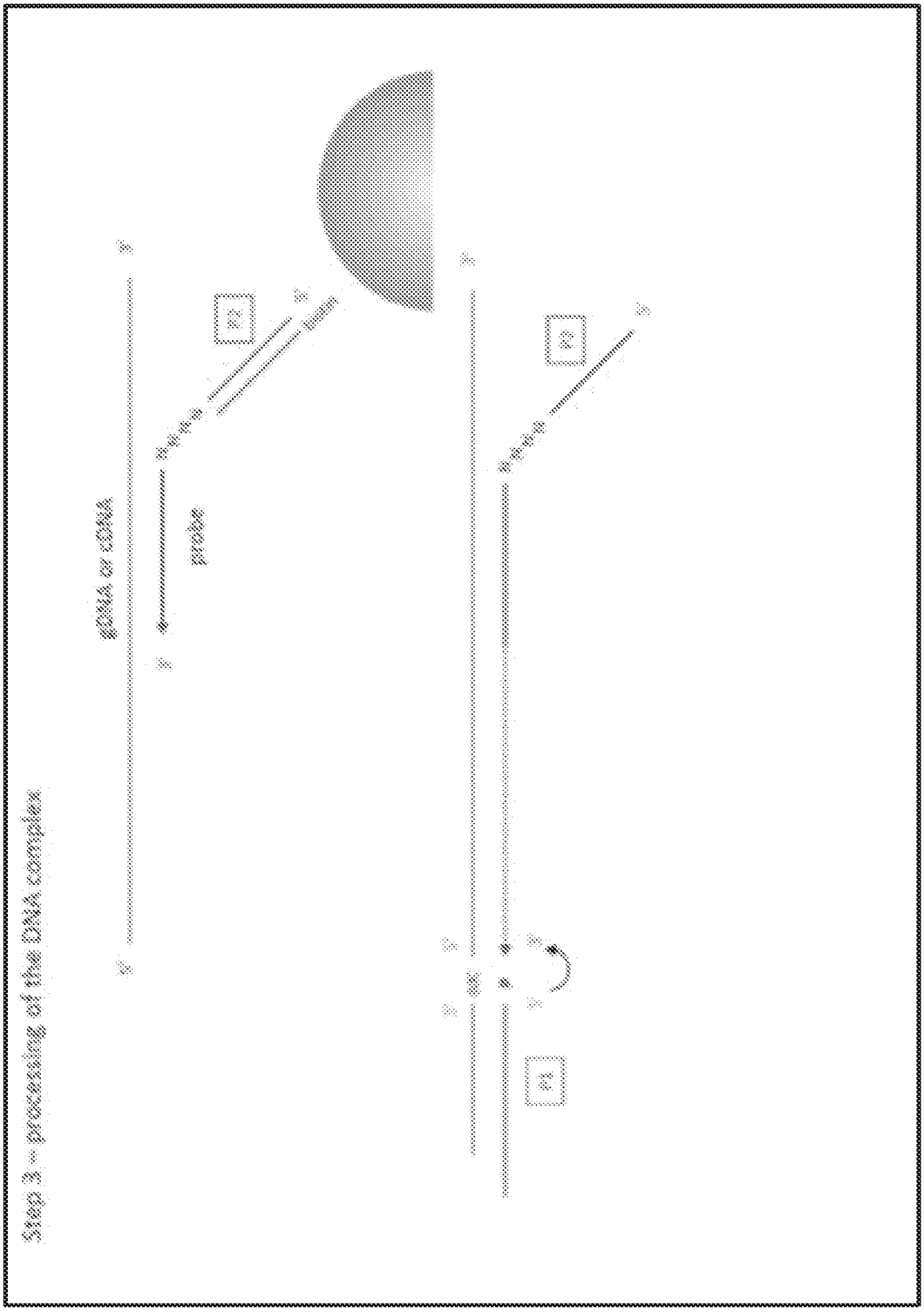


FIG. 34

Step 4 – PCR amplification
P1
Target region (ragged ends)
P2
Index
PCR is used to add flow cell compatible sequences and to introduce sample-specific indexes to each reaction.
Step 5 – DNA sequencing
read_1
read_1_index
read_2
read_1: target sequence 51 nt
read_1_index: sample label 8 nt
read_2: sequence label, probe tag 16 nt

FIG. 35

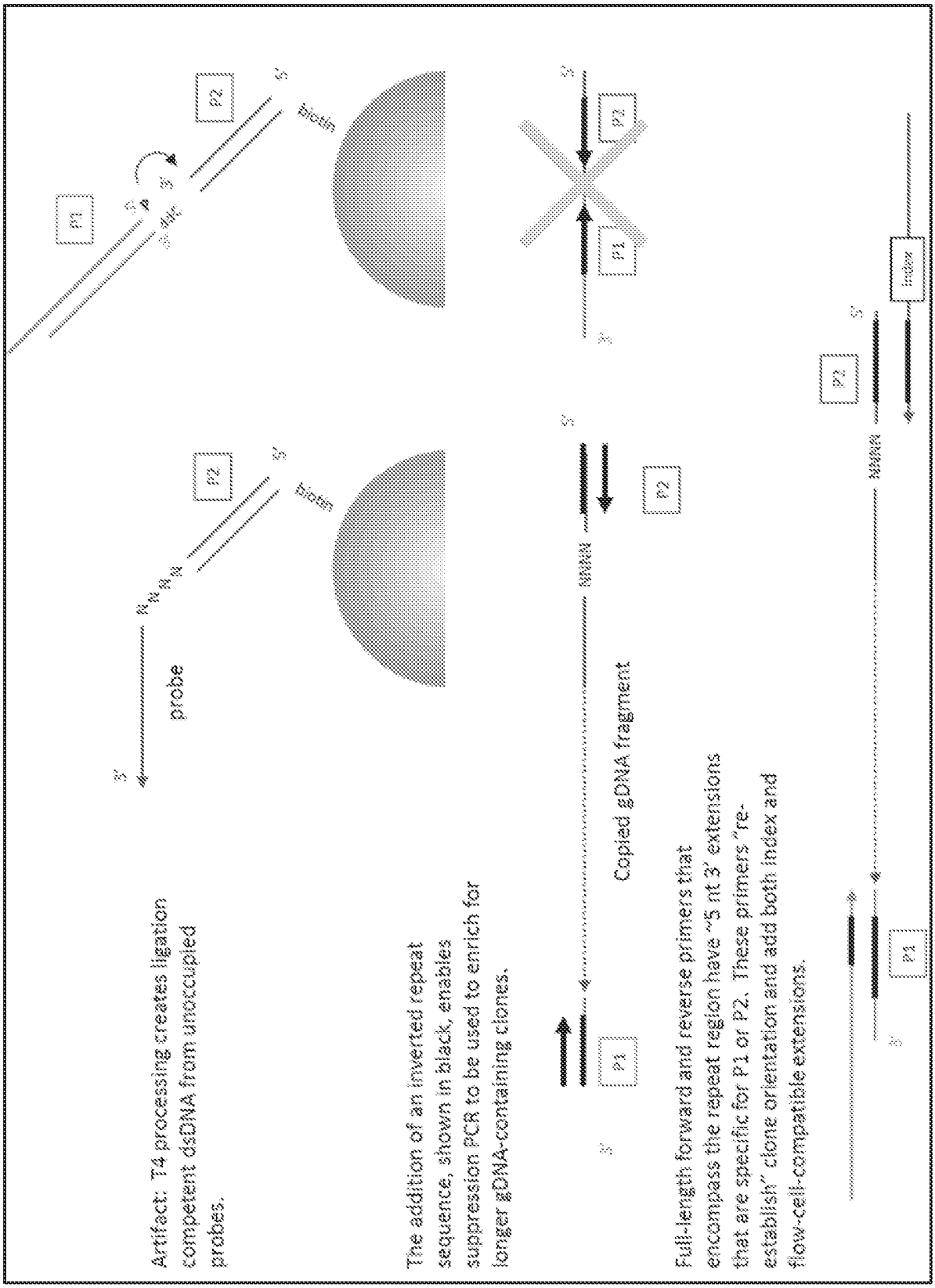

Artifact: T4 processing creates ligation competent dsDNA from unoccupied probes.
probe
P1
P2
biotin
The addition of an inverted repeat sequence, shown in black, enables suppression PCR to be used to enrich for longer gDNA-containing clones.
Copied gDNA fragment
Full-length forward and reverse primers that encompass the repeat region have ~5 nt 3' extensions that are specific for P1 or P2. These primers "re-establish" clone orientation and add both index and flow-cell-compatible extensions.
Index

FIG. 36

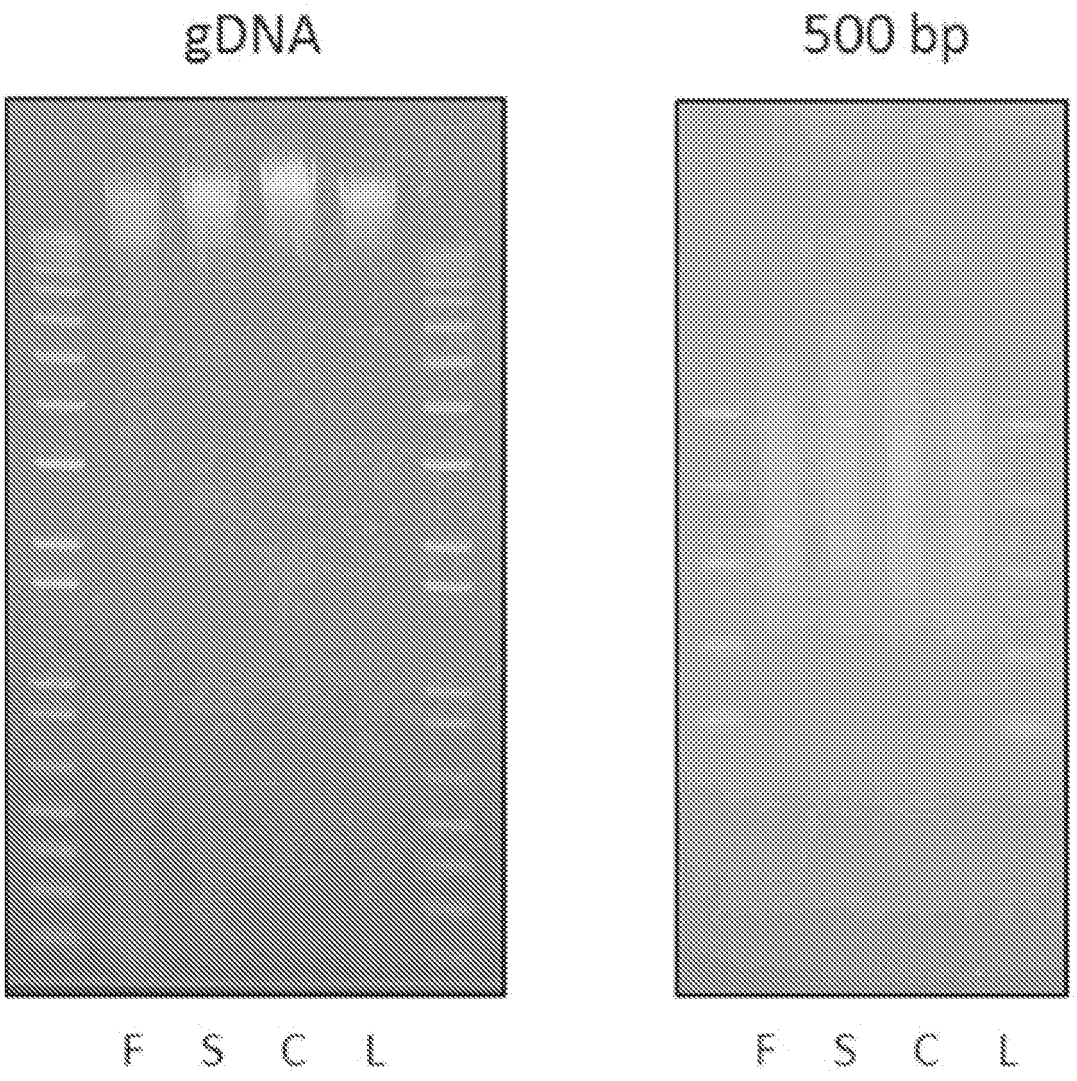


FIG. 37

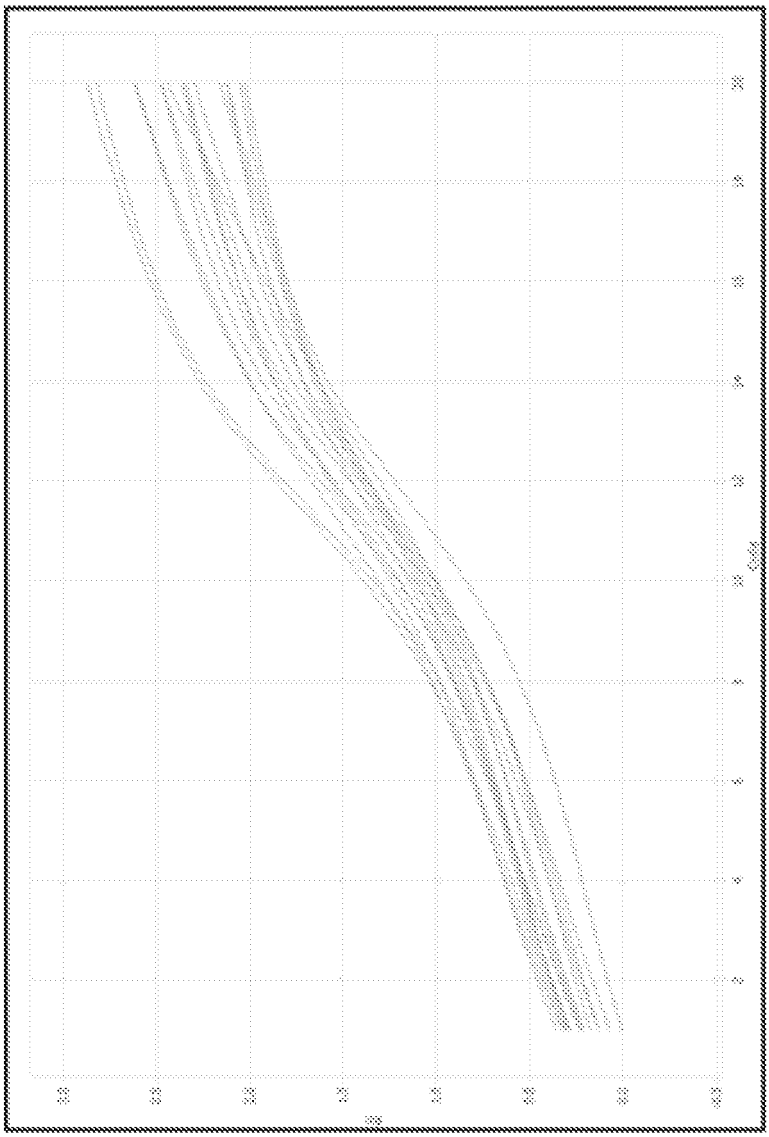


FIG. 38

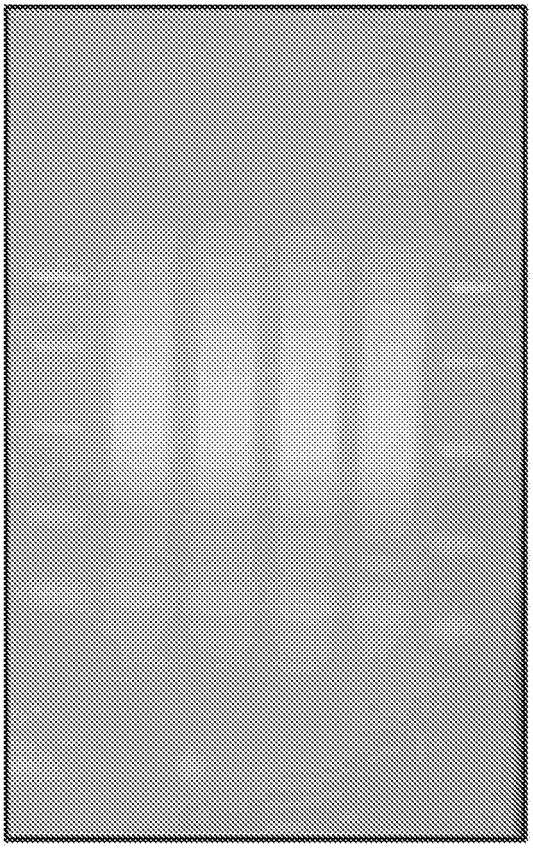


FIG. 39

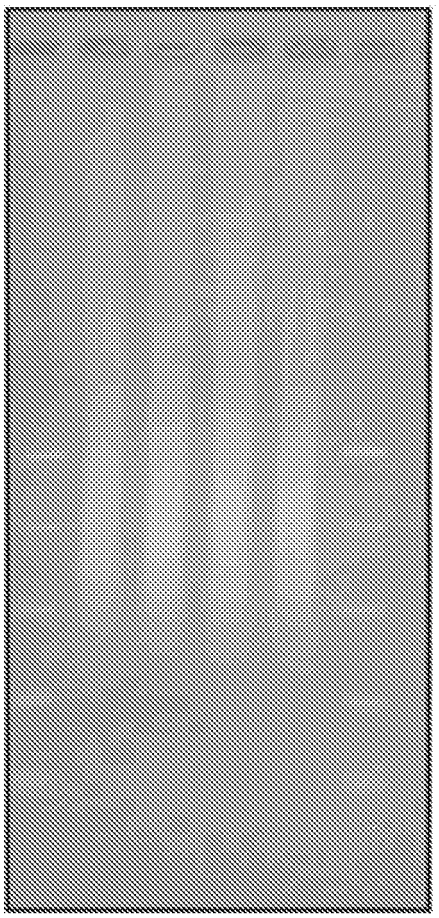


FIG. 40

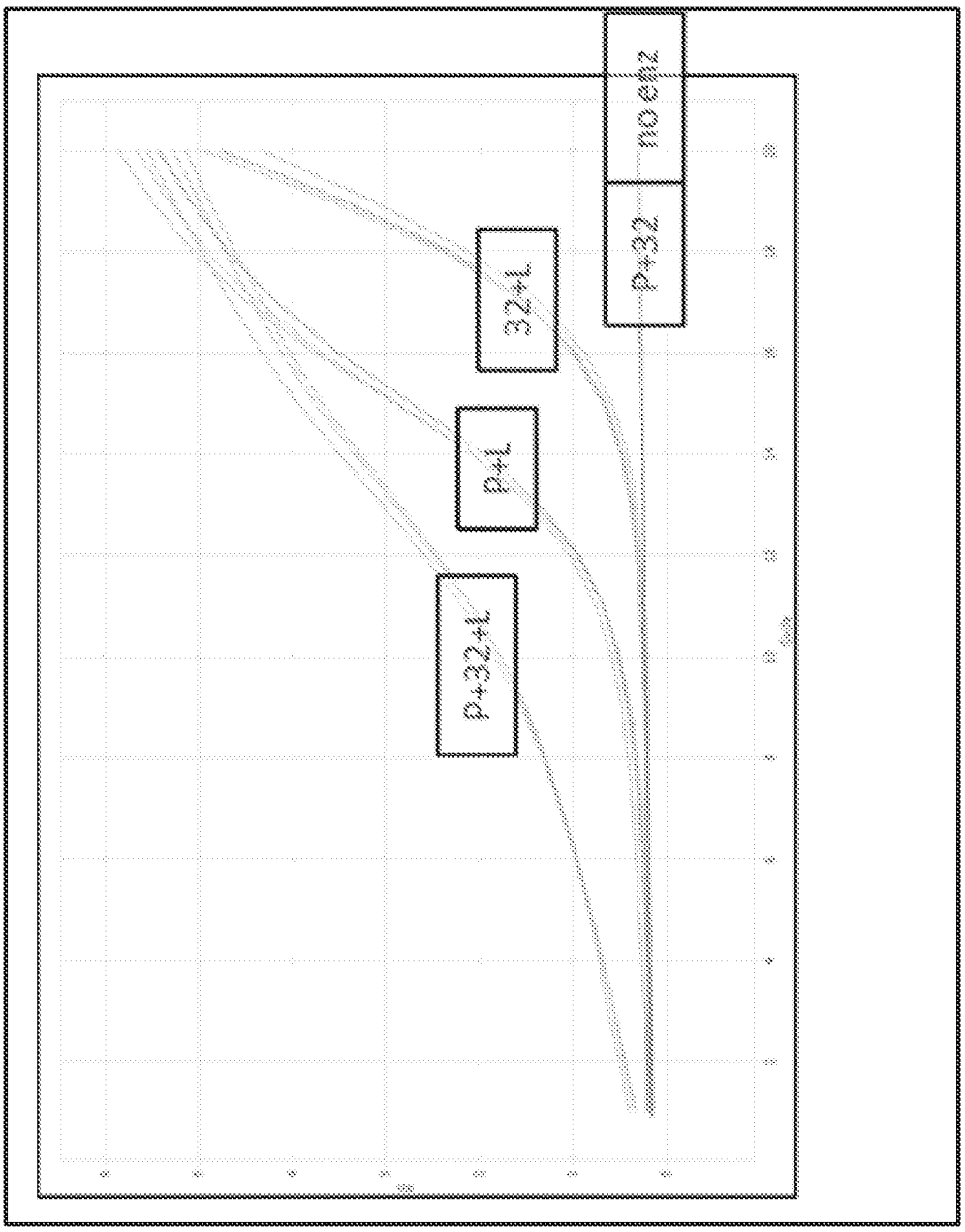


FIG. 41

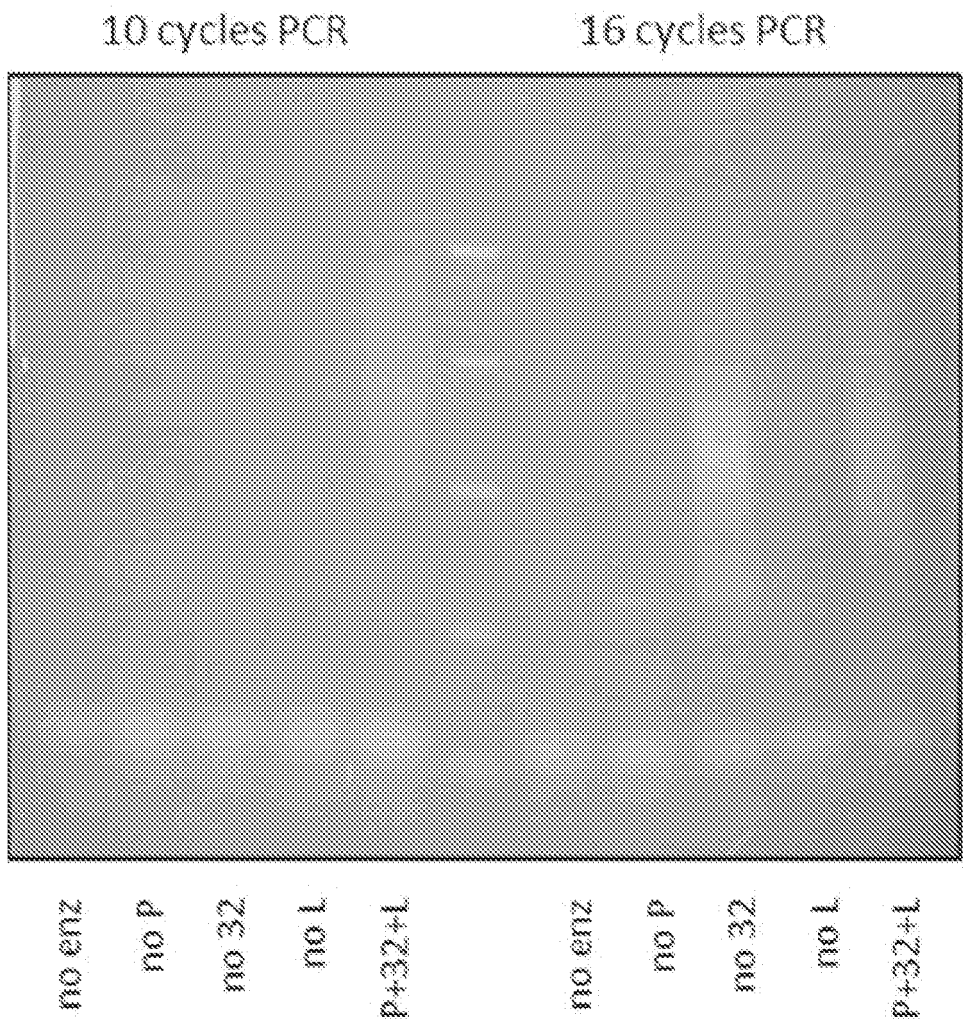

10 cycles PCR
16 cycles PCR
no enz
no P
no 32
no L
P+32+L
no enz
no P
no 32
no L
P+32+L

FIG. 42

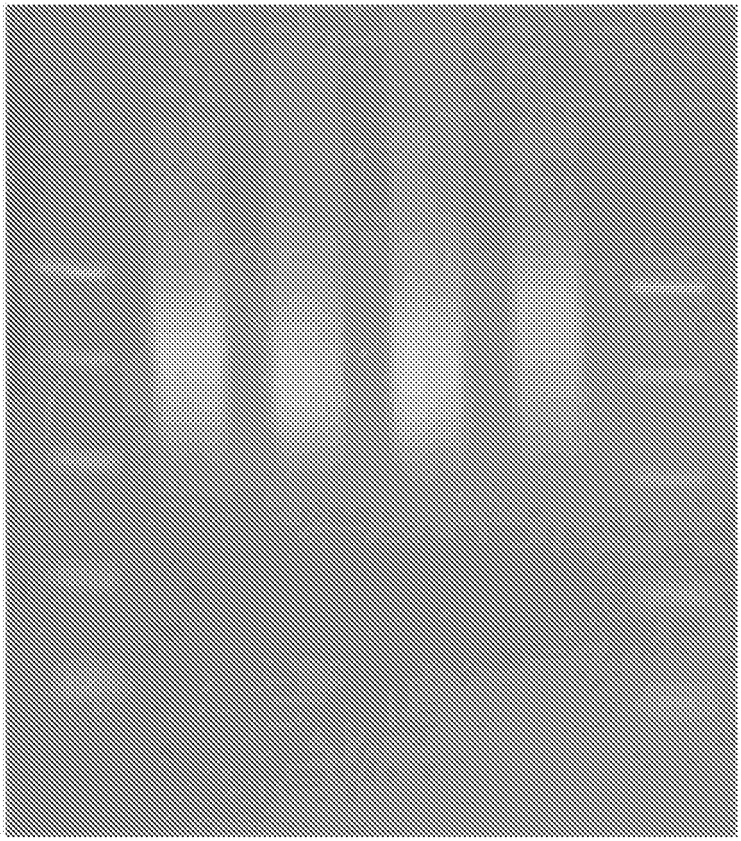


FIG. 43

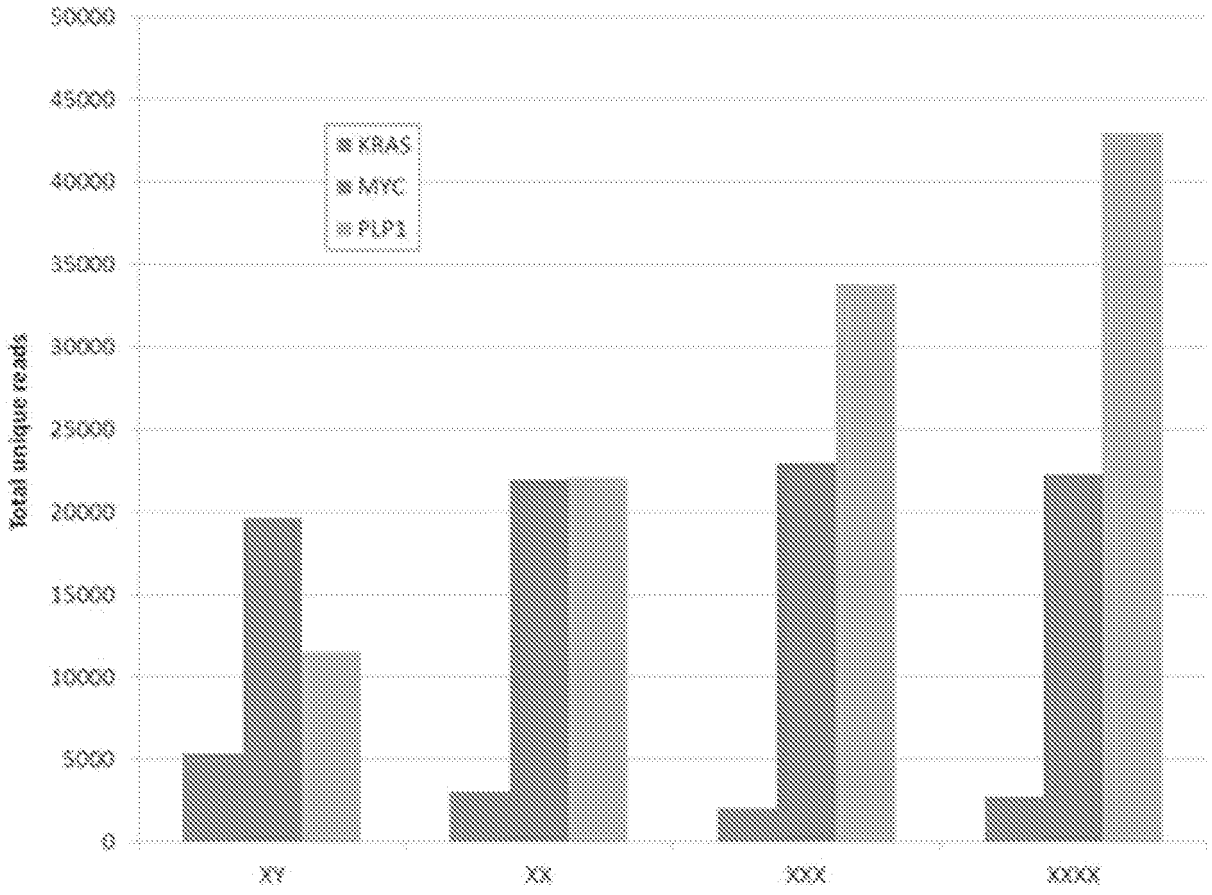

50000
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
Total unique reads
KRAS
MYC
PLP1
XY
XX
XXX
XXXX

FIG. 44

FIG. 45

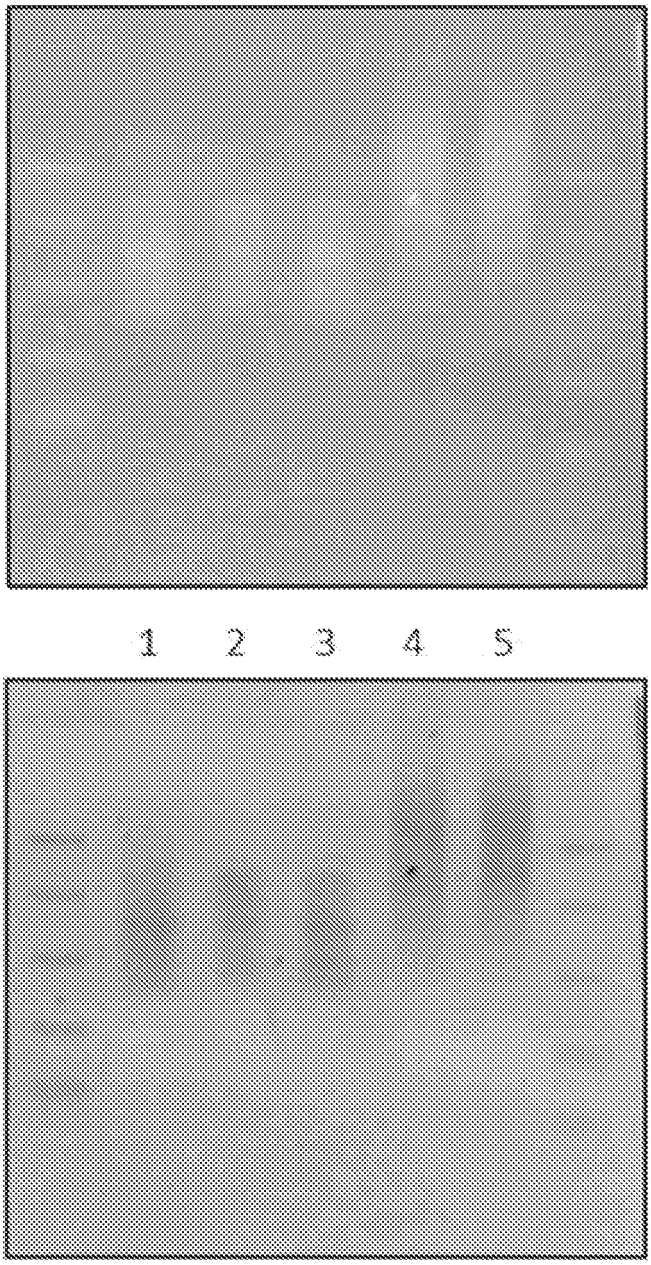


Figure 1. DNA gel of RNA-seq samples.

1. Post capture RNA-seq C+P
2. Pippin prep lane 4
3. Pippin prep lane 5
4. ACA2-20 dT library
5. ACA2_FLFP + exome CAC3_FLRP dT library

FIG. 46

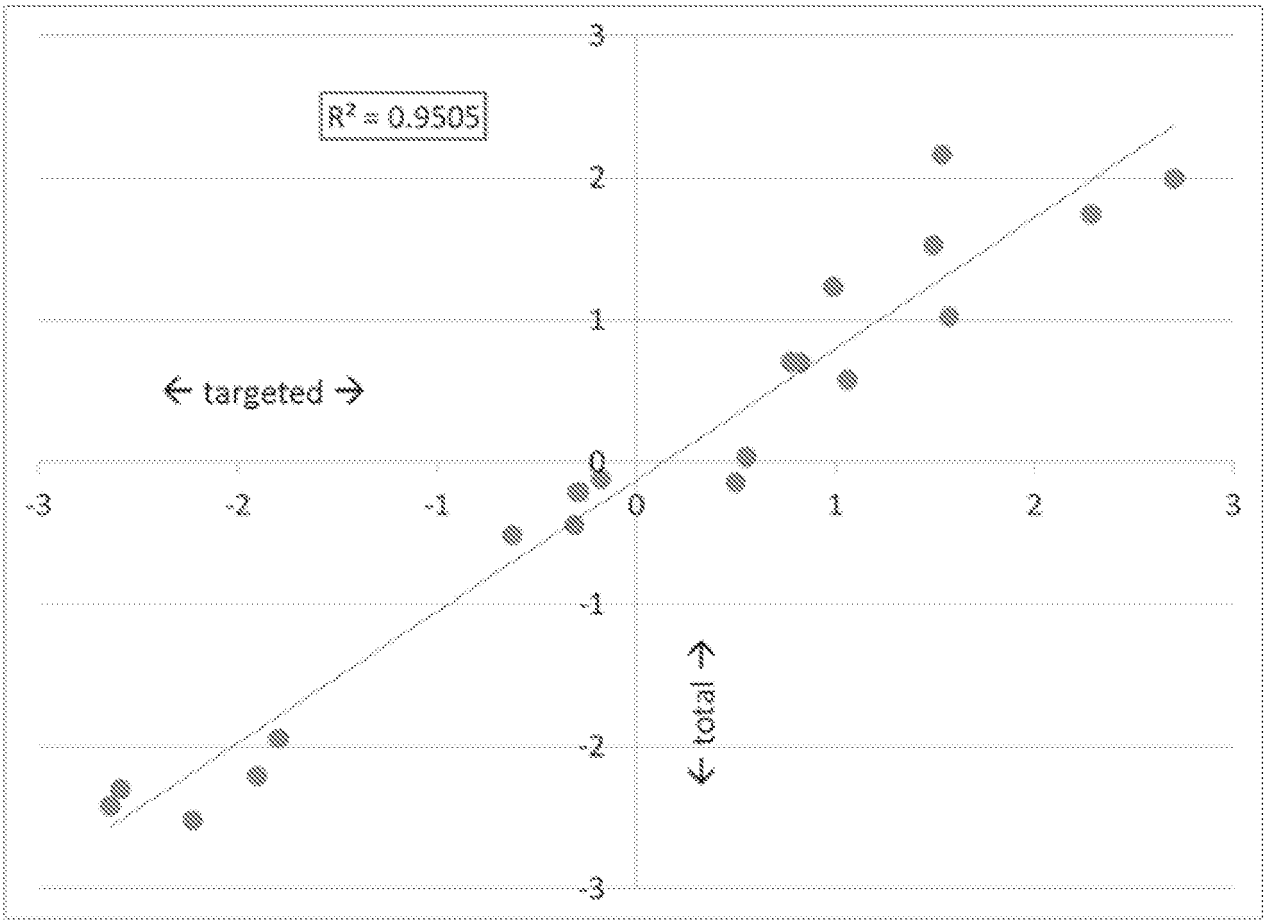


FIG. 47

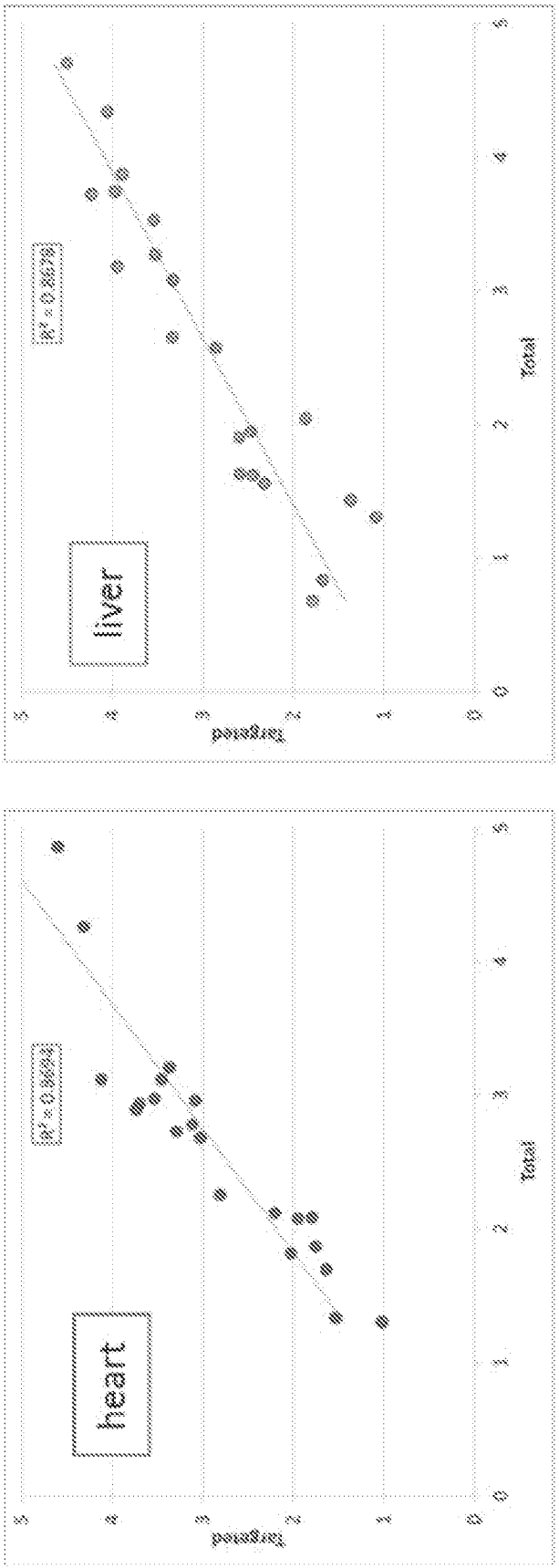


FIG. 48