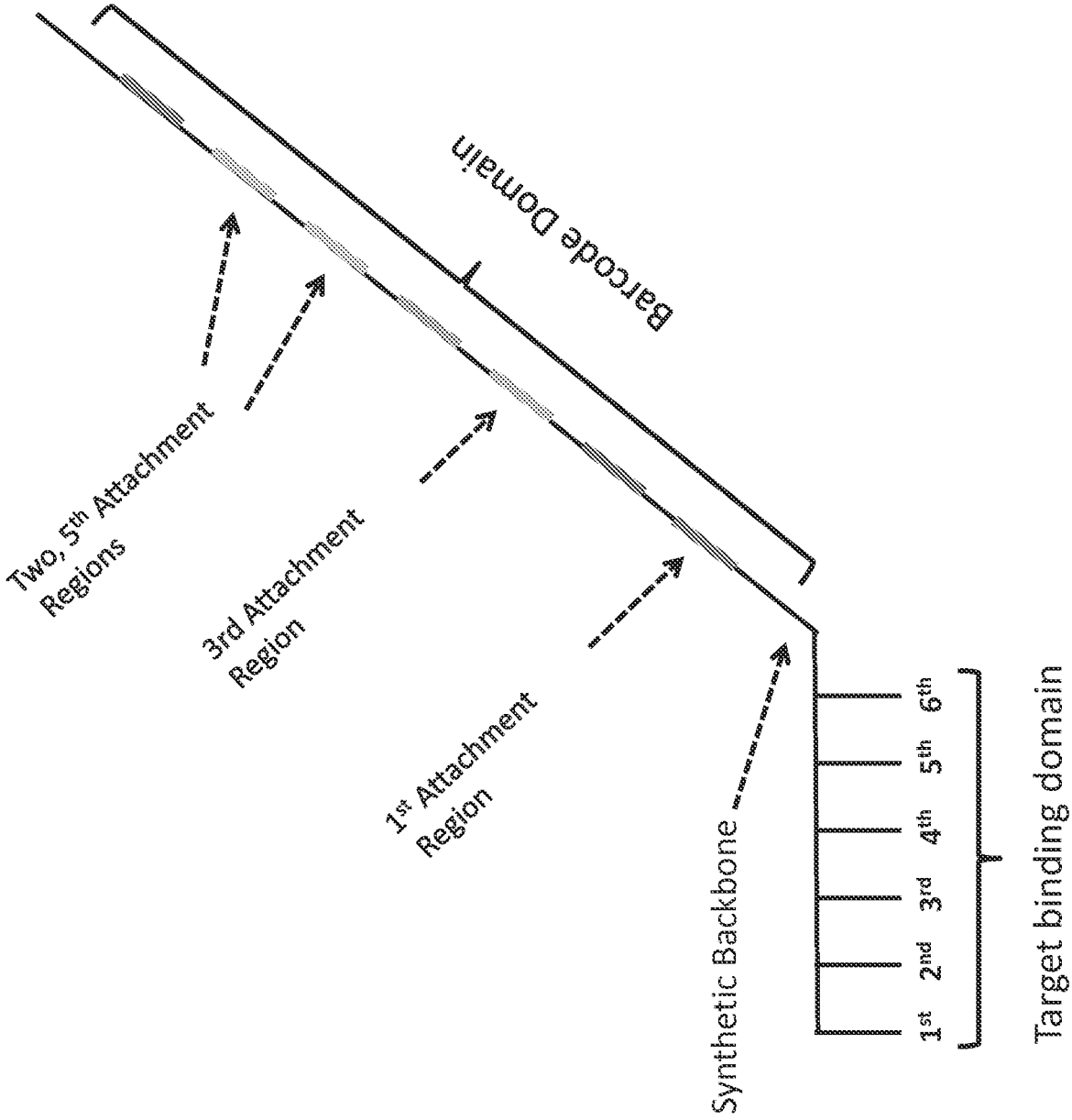

Barcode Domain
Two, 5th Attachment Regions
3rd Attachment Region
1st Attachment Region
Synthetic Backbone
1st 2nd 3rd 4th 5th 6th
Target binding domain

Figure 1

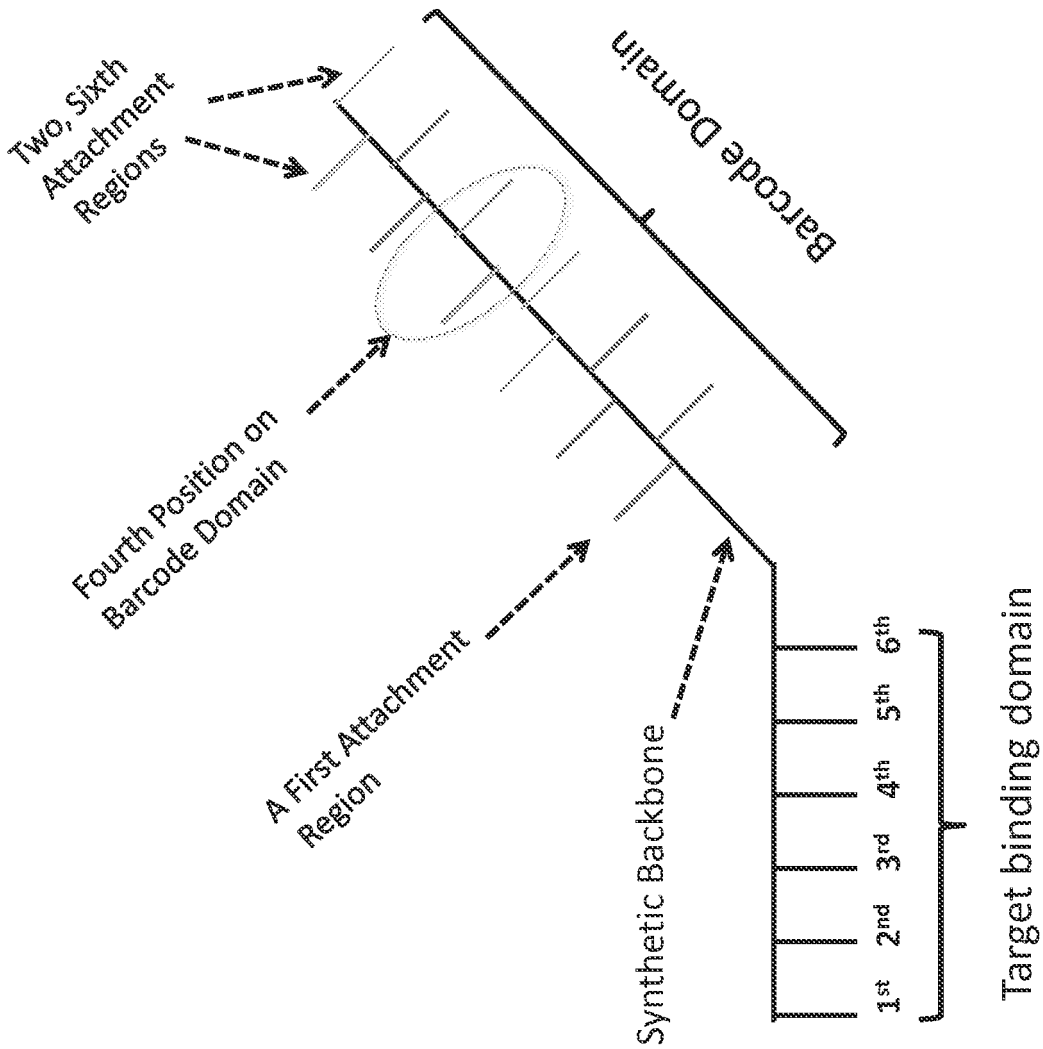

Two, Sixth Attachment Regions
Barcode Domain
Fourth Position on Barcode Domain
A First Attachment Region
Synthetic Backbone
1st 2nd 3rd 4th 5th 6th
Target binding domain

Figure 2

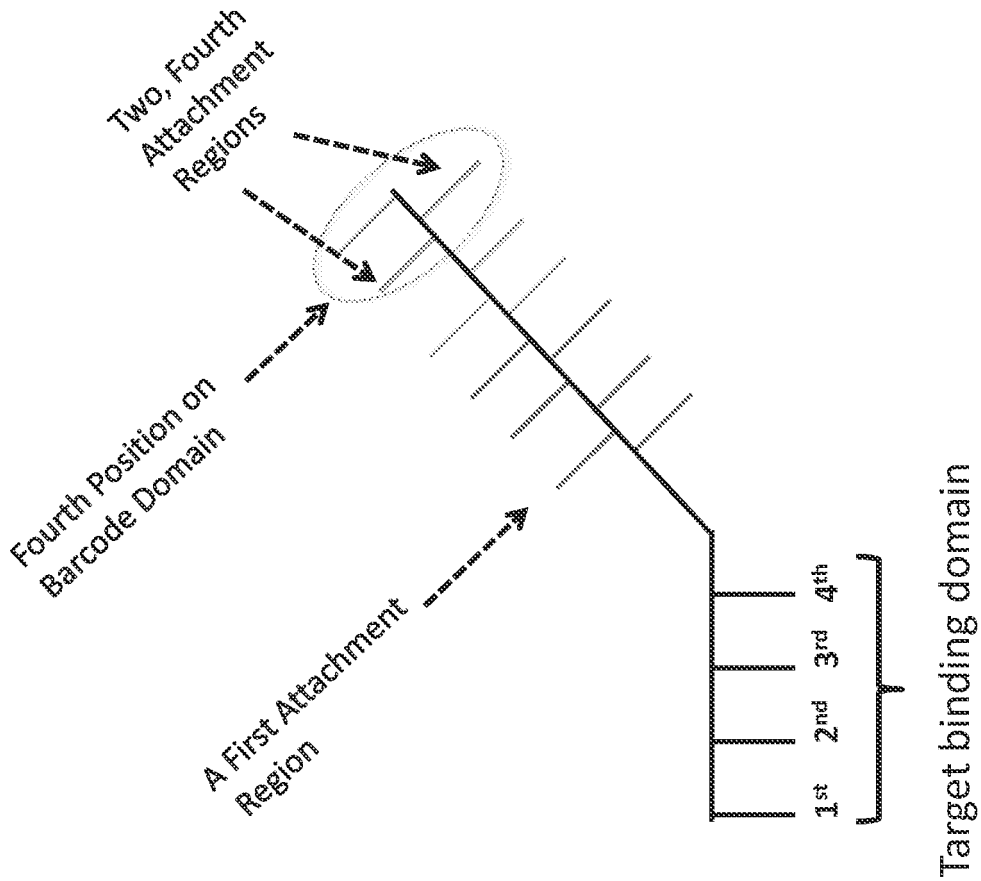

Two, Fourth Attachment Regions
Fourth Position on Barcode Domain
A First Attachment Region
1st
2nd
3rd
4th
Target binding domain

Figure 3

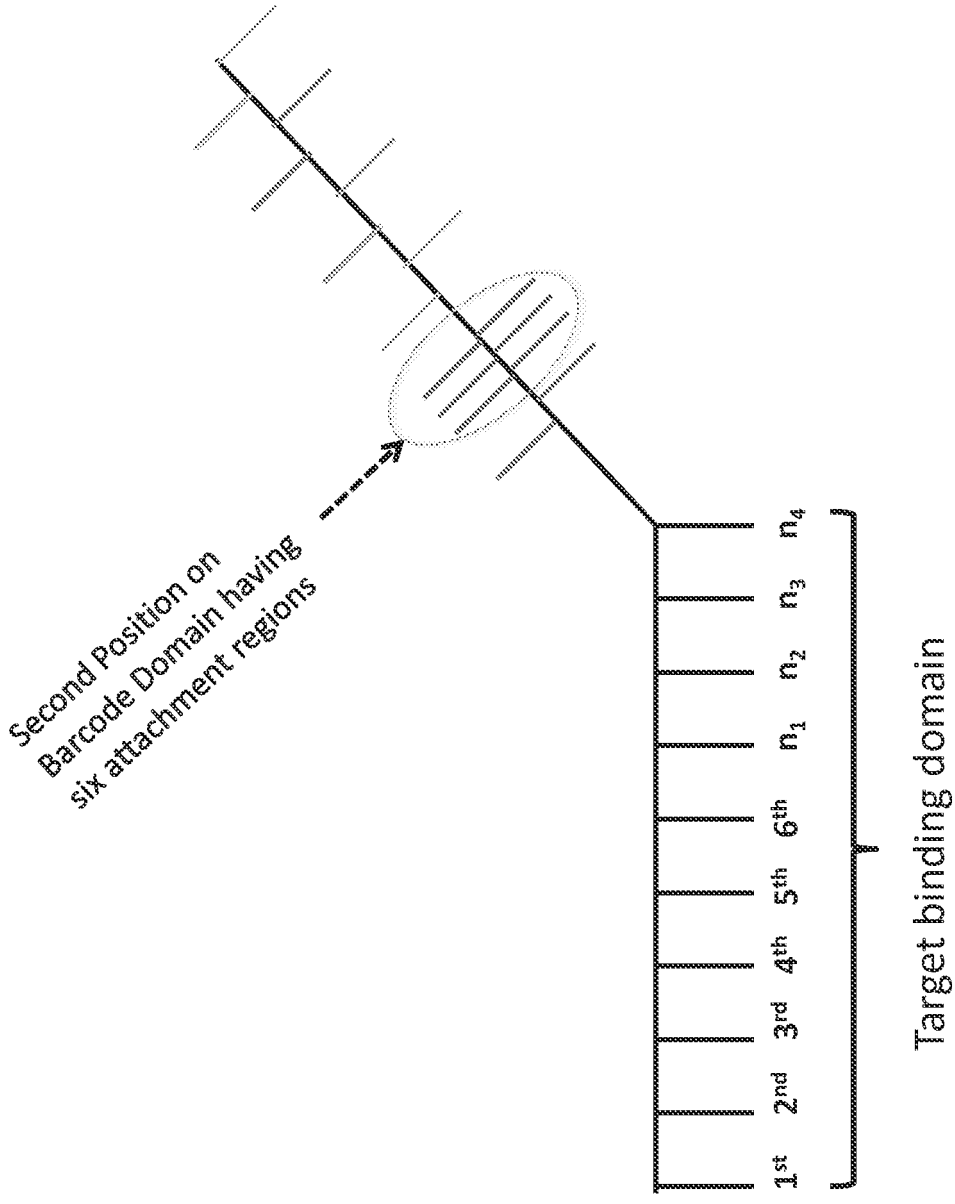

Second Position on
Barcode Domain having
six attachment regions
1st
2nd
3rd
4th
5th
6th
n1
n2
n3
n4
Target binding domain

Figure 4

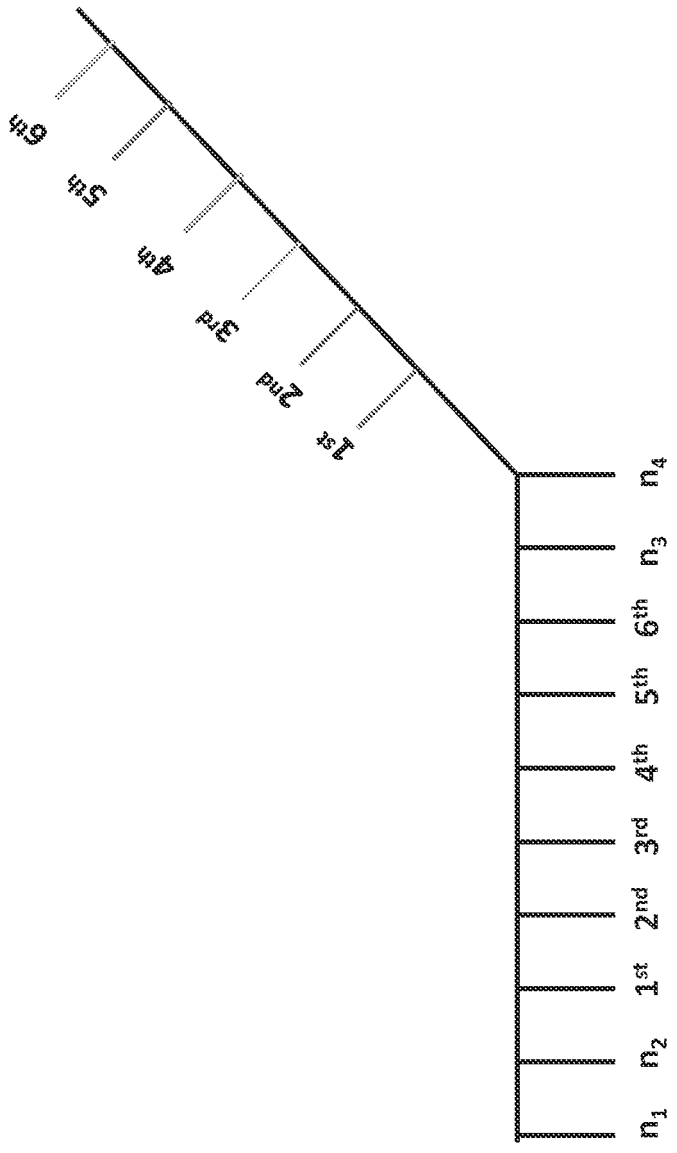


Figure 5

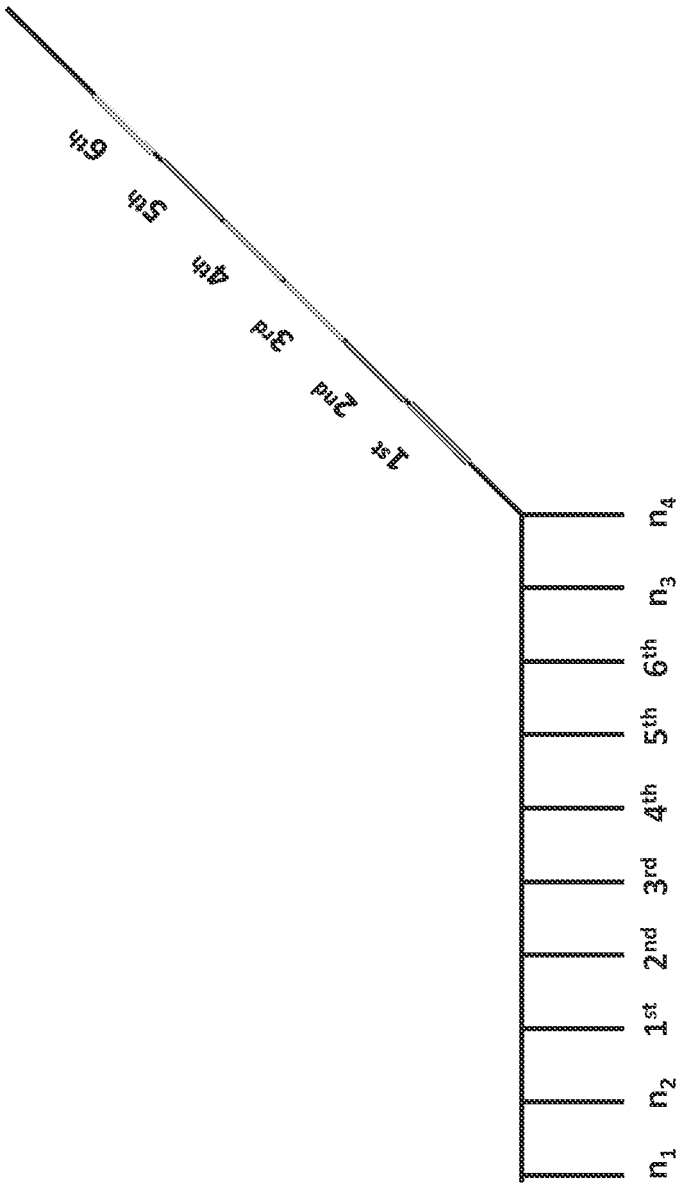


Figure 6A

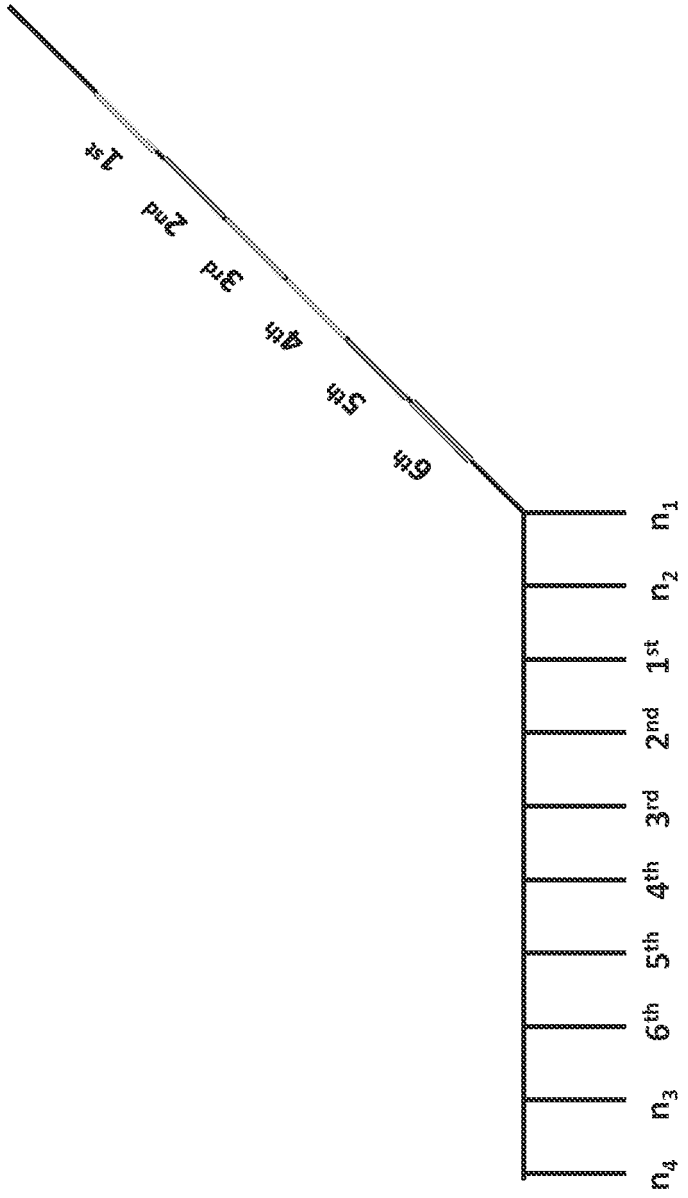


Figure 6B

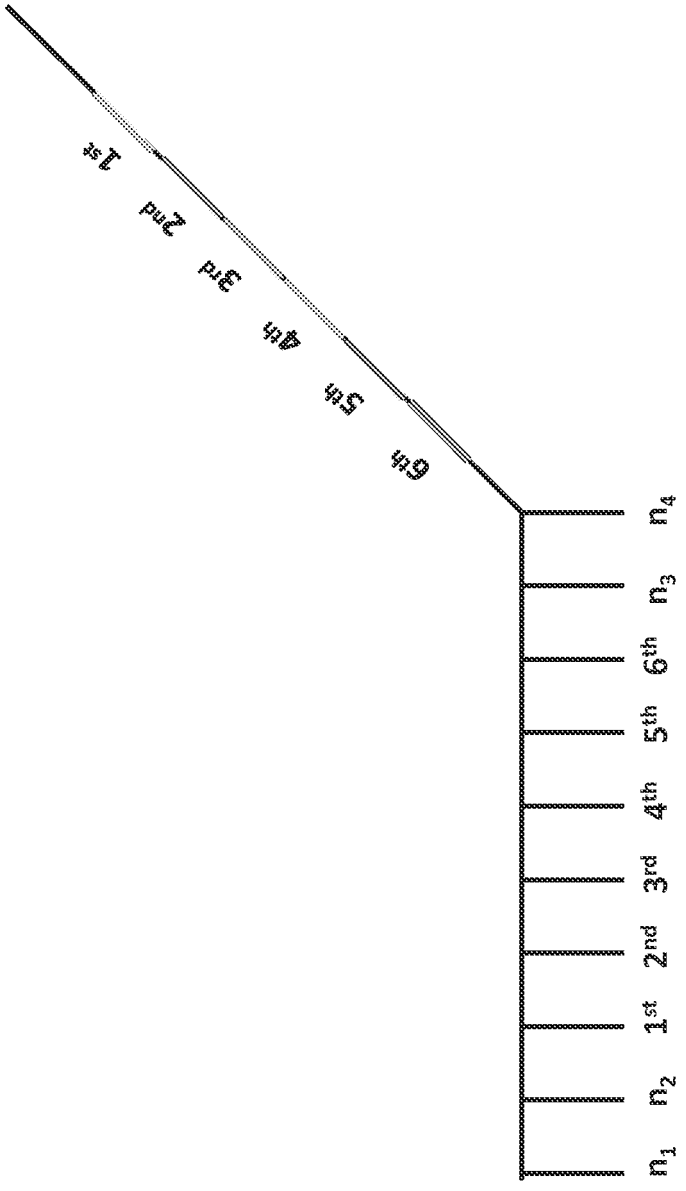


Figure 6C

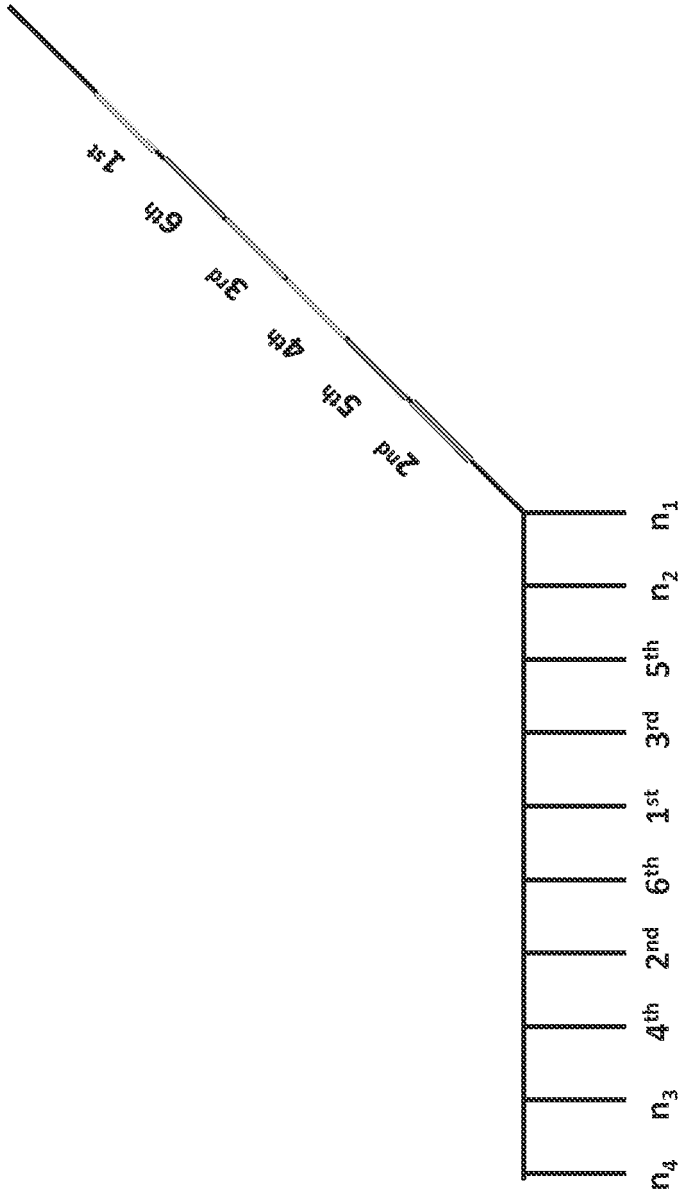
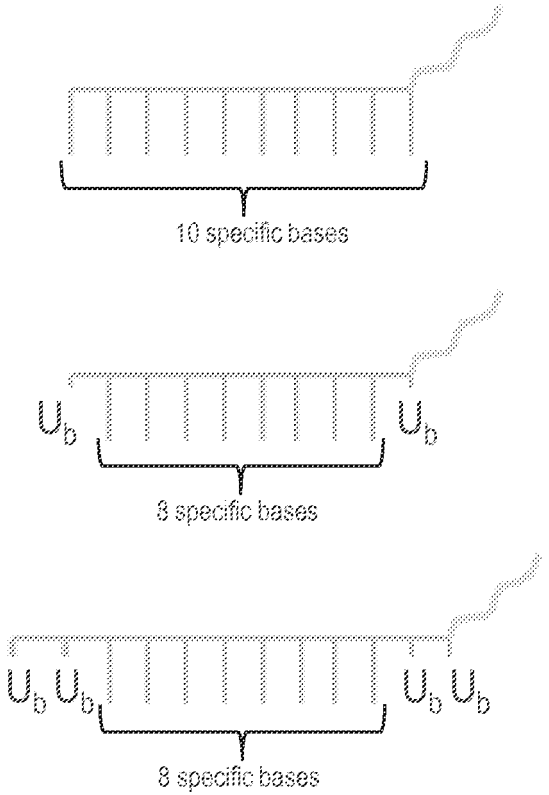


Figure 6D

Figure 7



10 specific bases: 4^{10} ~ 1 million probes

8 specific bases: 4^8 ~ 65 thousand probes

6 specific bases: 4^6 ~ 4 thousand probes

U_b = universal base (e.g., Inosine and nitro-indole)

Figure 8A

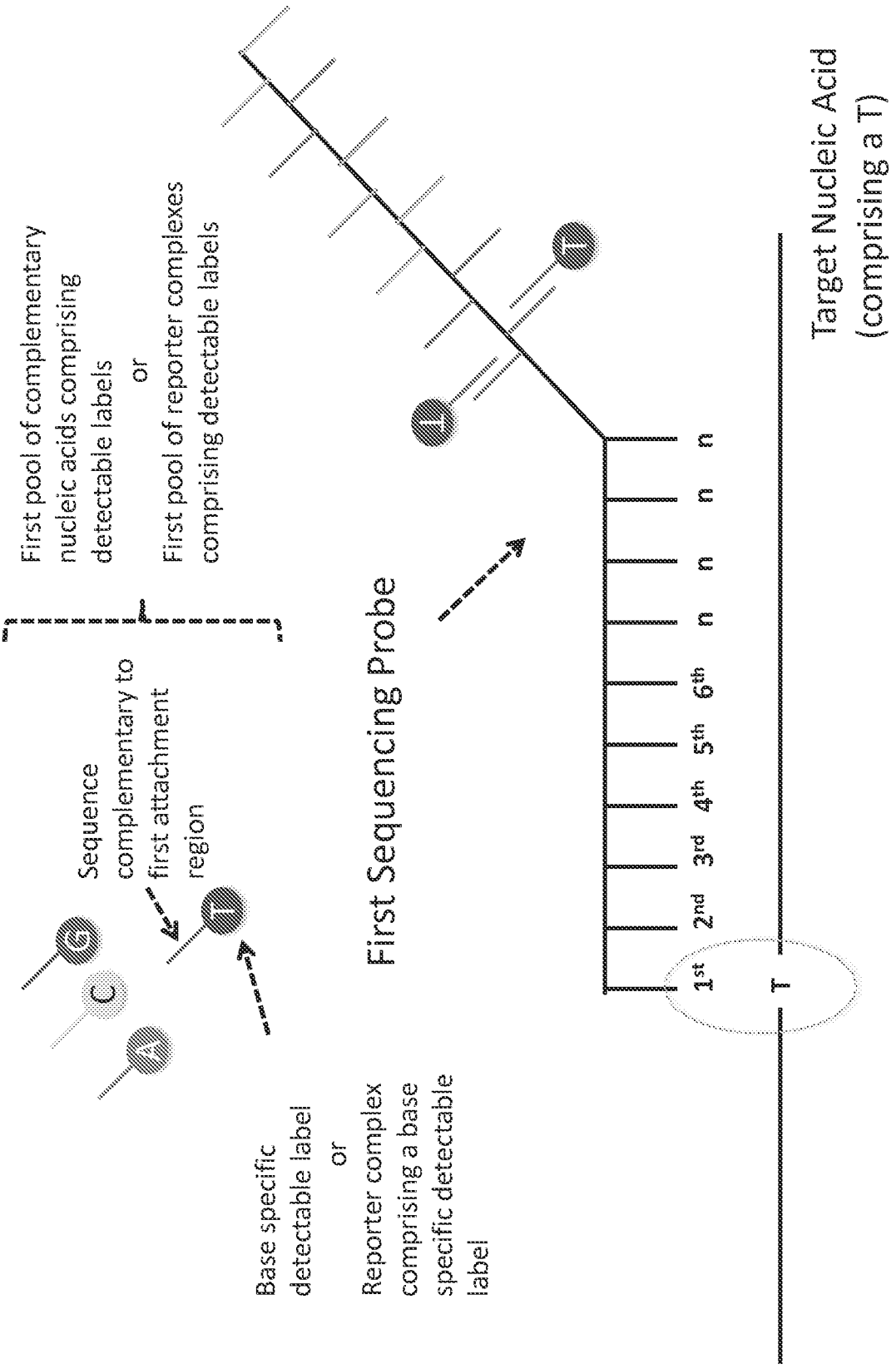

First pool of complementary nucleic acids comprising detectable labels
or
First pool of reporter complexes comprising detectable labels
Sequence complementary to first attachment region
Base specific detectable label
or
Reporter complex comprising a base specific detectable label
First Sequencing Probe
1st 2nd 3rd 4th 5th 6th n n n n n
T
Target Nucleic Acid (comprising a T)

Figure 8B

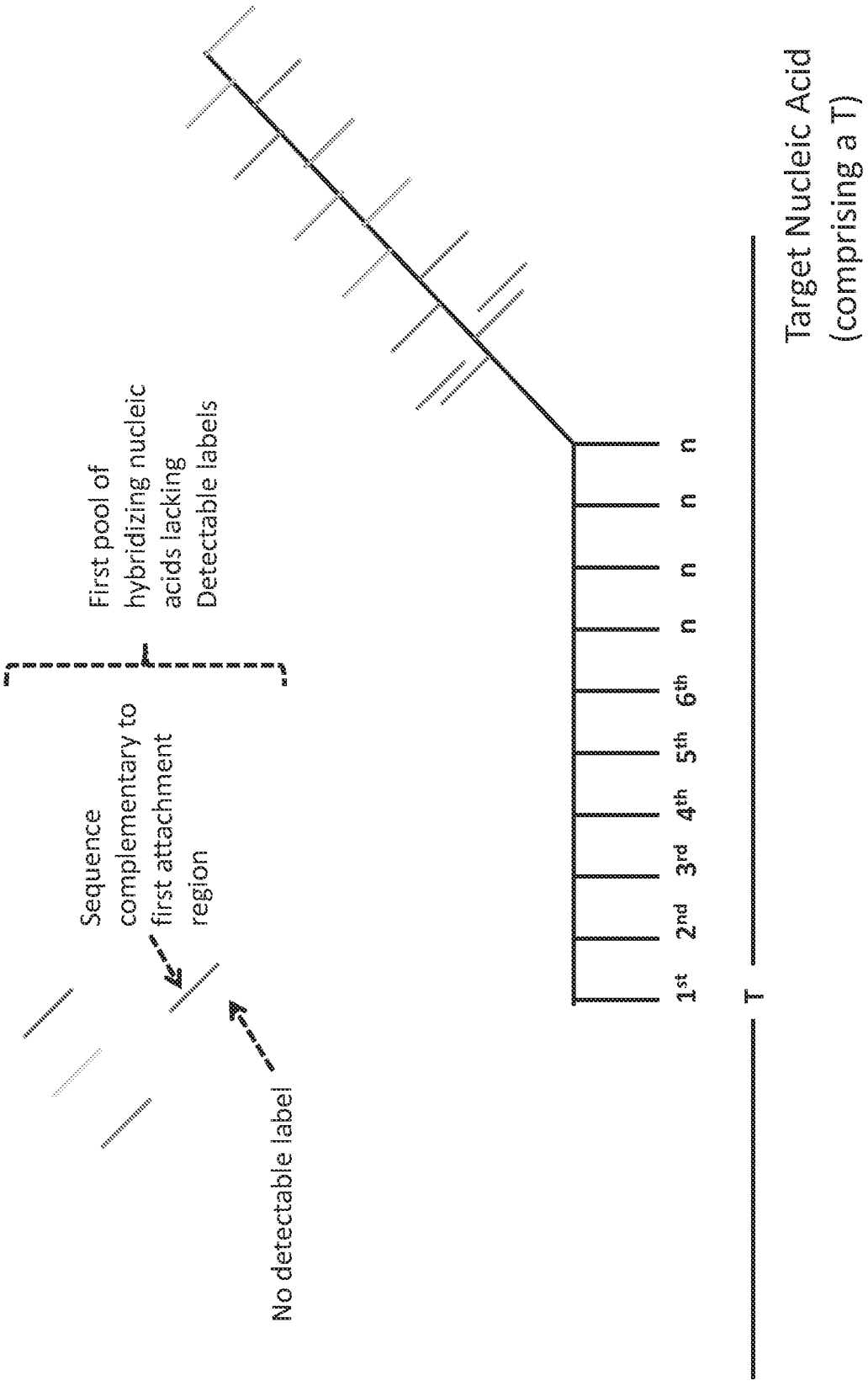

Target Nucleic Acid
(comprising a T)
T
1st 2nd 3rd 4th 5th 6th n n n n n
First pool of
hybridizing nucleic
acids lacking
Detectable labels
Sequence
complementary to
first attachment
region
No detectable label

Figure 8C

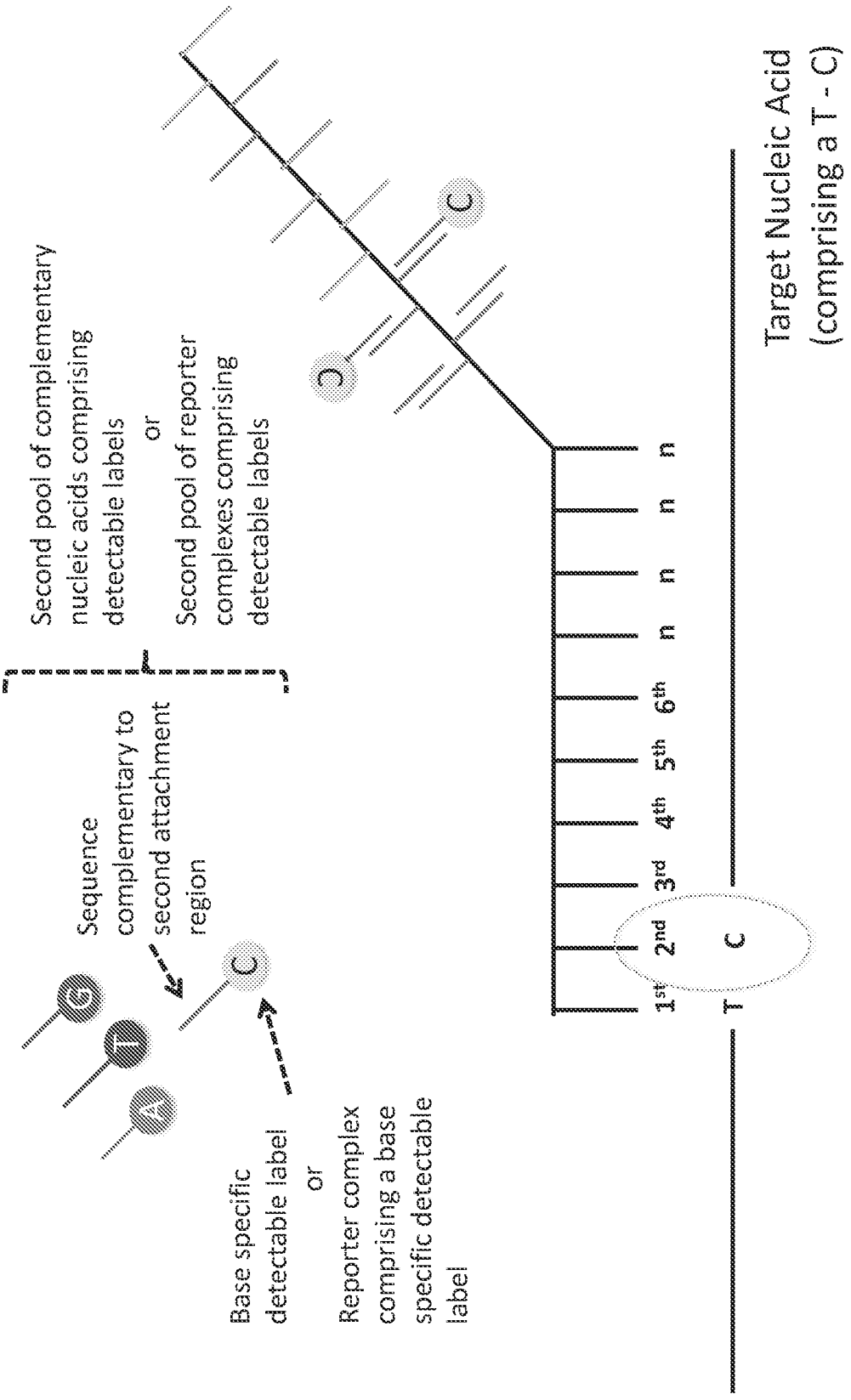
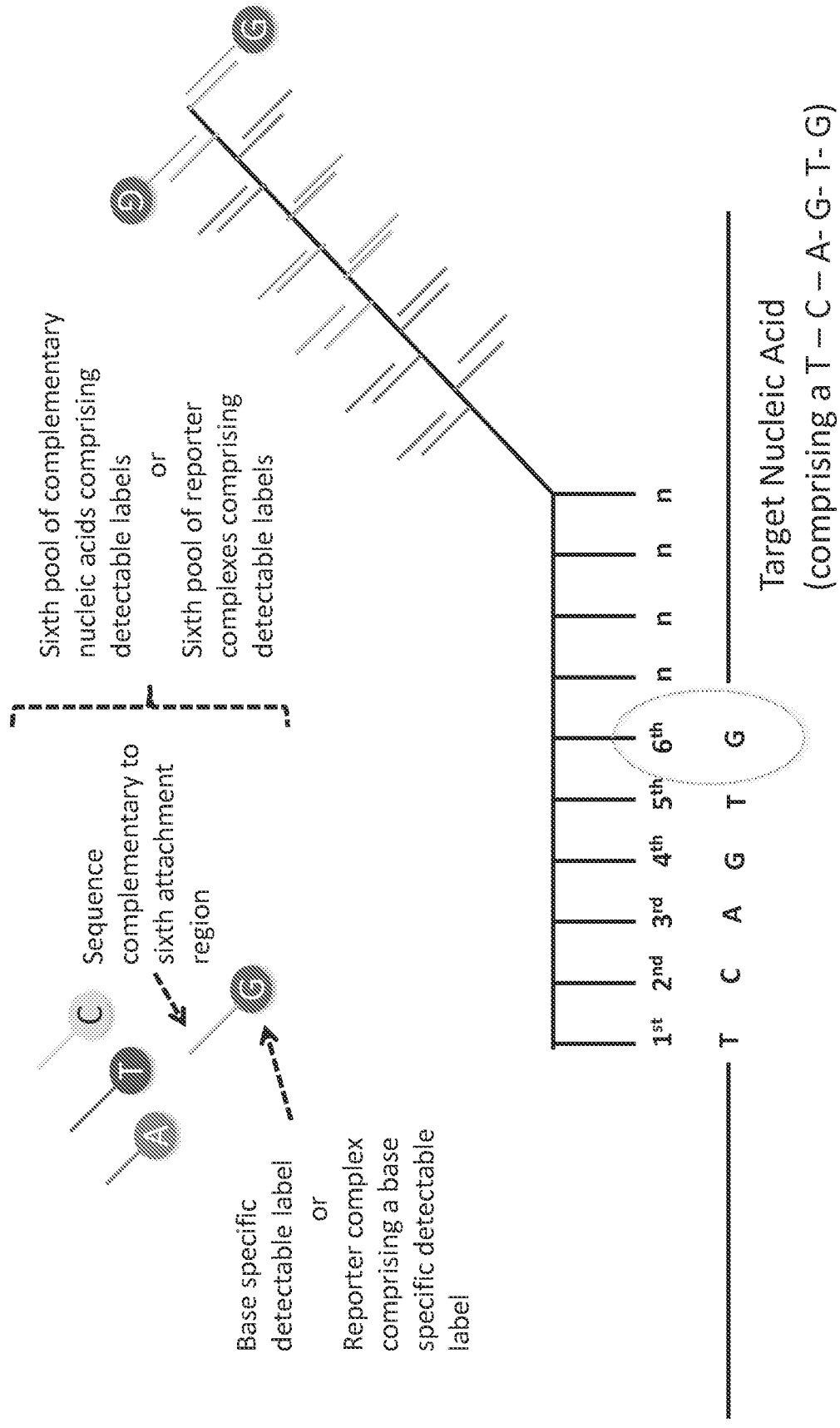

Second pool of complementary nucleic acids comprising detectable labels
or
Second pool of reporter complexes comprising detectable labels
Sequence complementary to second attachment region
Base specific detectable label
or
Reporter complex comprising a base specific detectable label
Target Nucleic Acid (comprising a T - C)
1st 2nd 3rd 4th 5th 6th
n n n n n
T C

Figure 8D


Sixth pool of complementary nucleic acids comprising detectable labels
or
Sixth pool of reporter complexes comprising detectable labels
Sequence complementary to sixth attachment region
Base specific detectable label
or
Reporter complex comprising a base specific detectable label
1st 2nd 3rd 4th 5th 6th
n n n n n n
T C A G T G
Target Nucleic Acid
(comprising a T – C – A- G- T- G)

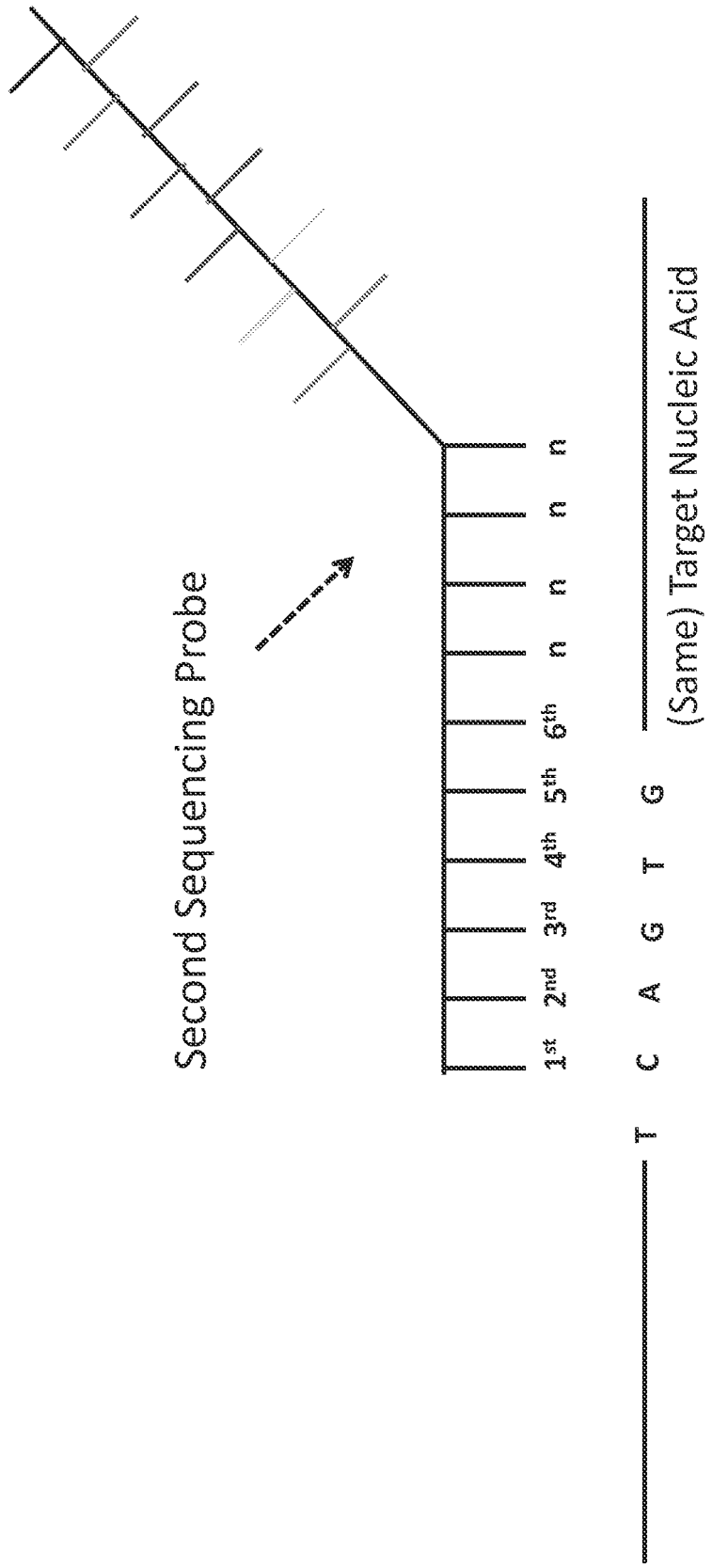

Second Sequencing Probe
1st 2nd 3rd 4th 5th 6th n n n n
T C A G T G
(Same) Target Nucleic Acid

Figure 8E

Figure 9A

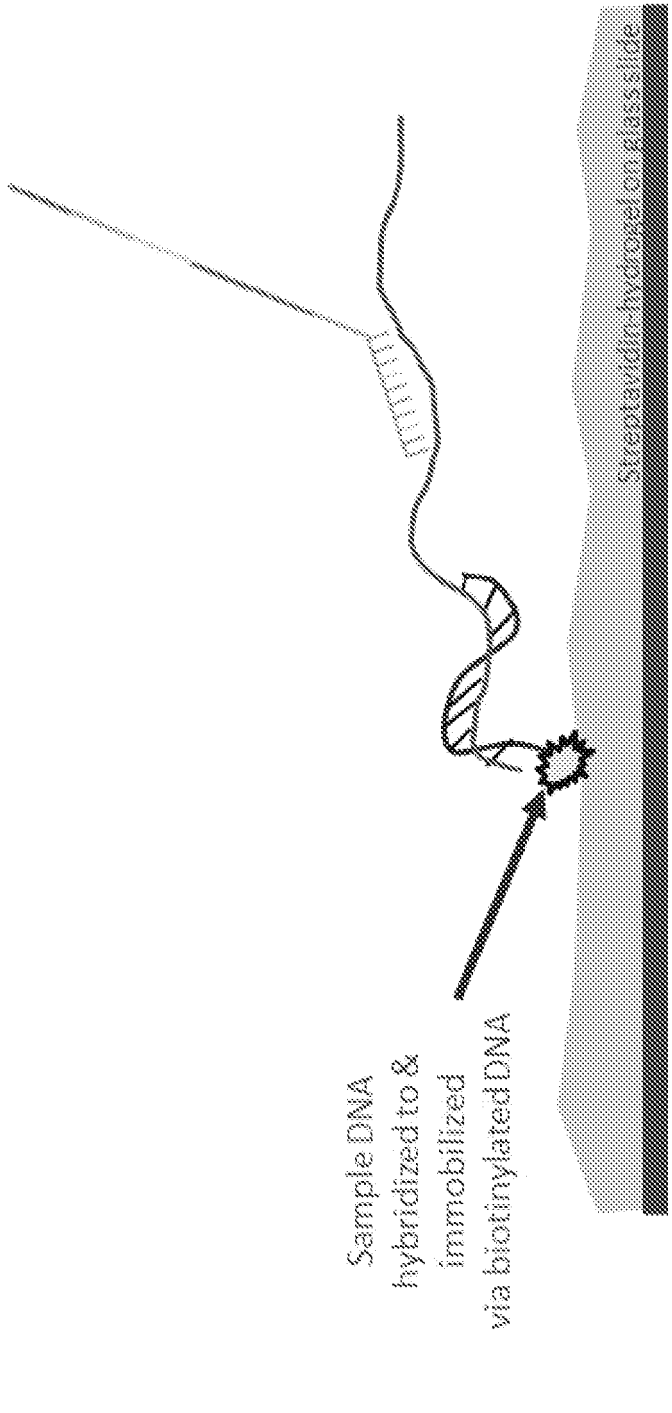

Sample DNA
hybridized to &
immobilized
via biotinylated DNA
Streptavidin hydrogel on glass slide
Cycle 1: → probe Hyb

Figure 9B

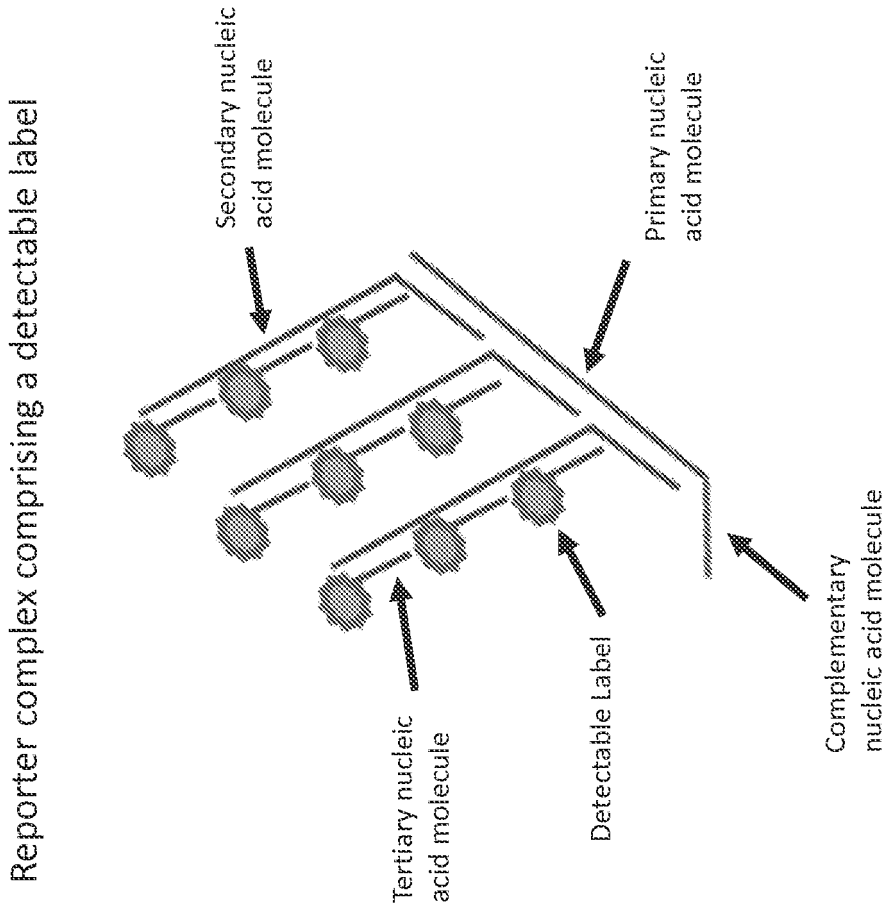

Reporter complex comprising a detectable label
Secondary nucleic acid molecule
Primary nucleic acid molecule
Tertiary nucleic acid molecule
Detectable Label
Complementary nucleic acid molecule

Figure 9C

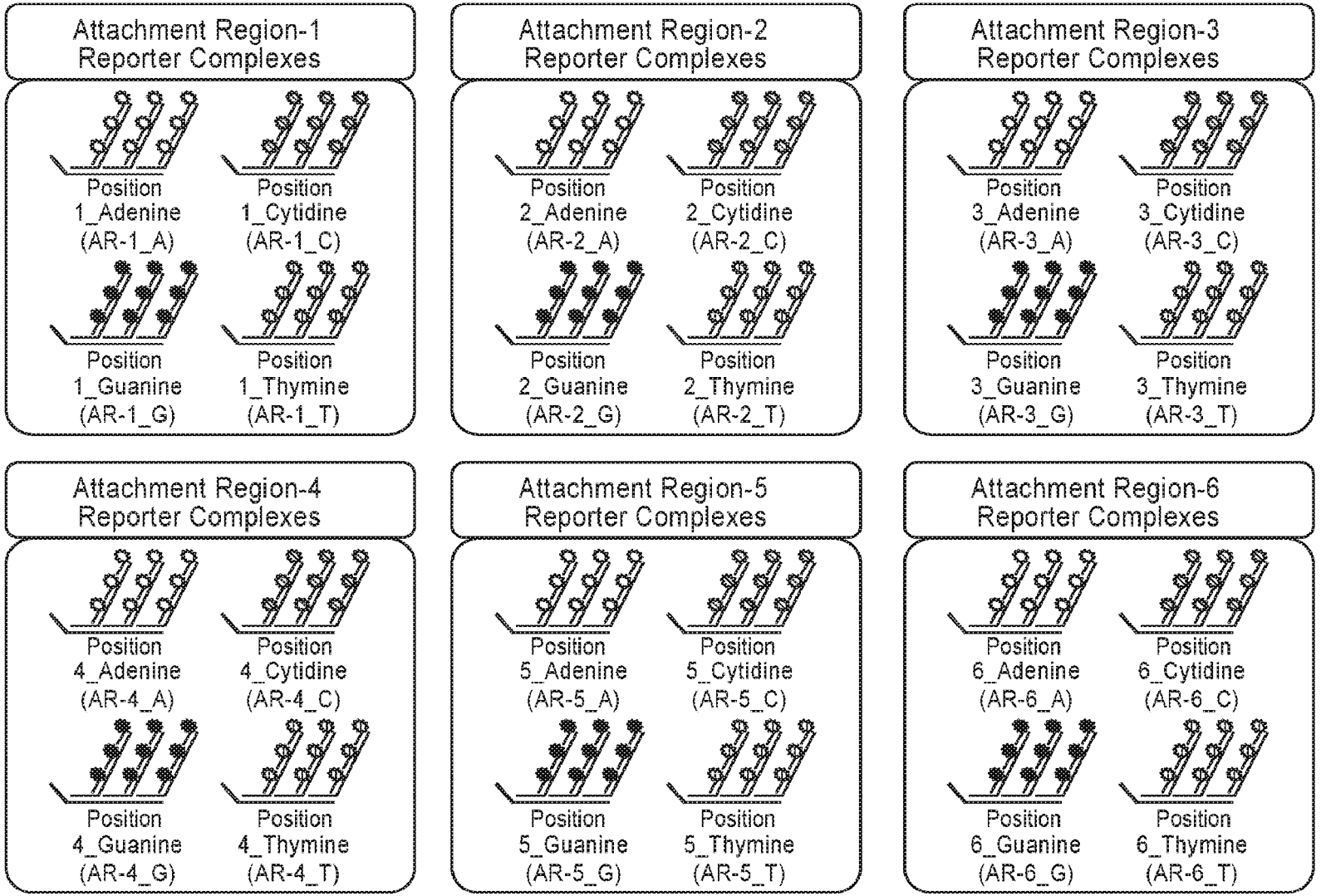

Attachment Region-1 Reporter Complexes
Position 1_Adenine (AR-1_A)
Position 1_Cytidine (AR-1_C)
Position 1_Guanine (AR-1_G)
Position 1_Thymine (AR-1_T)
Attachment Region-2 Reporter Complexes
Position 2_Adenine (AR-2_A)
Position 2_Cytidine (AR-2_C)
Position 2_Guanine (AR-2_G)
Position 2_Thymine (AR-2_T)
Attachment Region-3 Reporter Complexes
Position 3_Adenine (AR-3_A)
Position 3_Cytidine (AR-3_C)
Position 3_Guanine (AR-3_G)
Position 3_Thymine (AR-3_T)
Attachment Region-4 Reporter Complexes
Position 4_Adenine (AR-4_A)
Position 4_Cytidine (AR-4_C)
Position 4_Guanine (AR-4_G)
Position 4_Thymine (AR-4_T)
Attachment Region-5 Reporter Complexes
Position 5_Adenine (AR-5_A)
Position 5_Cytidine (AR-5_C)
Position 5_Guanine (AR-5_G)
Position 5_Thymine (AR-5_T)
Attachment Region-6 Reporter Complexes
Position 6_Adenine (AR-6_A)
Position 6_Cytidine (AR-6_C)
Position 6_Guanine (AR-6_G)
Position 6_Thymine (AR-6_T)

Figure 9D

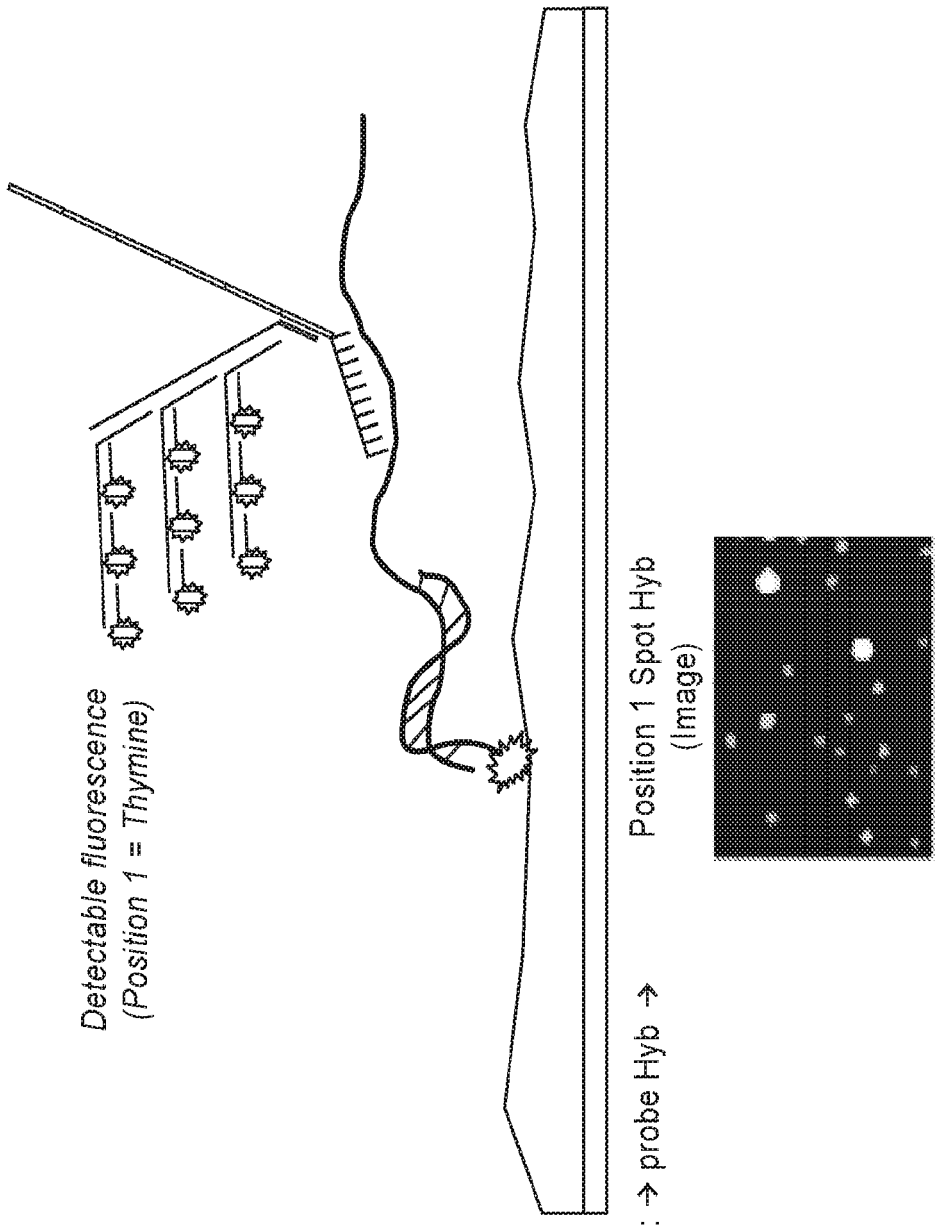
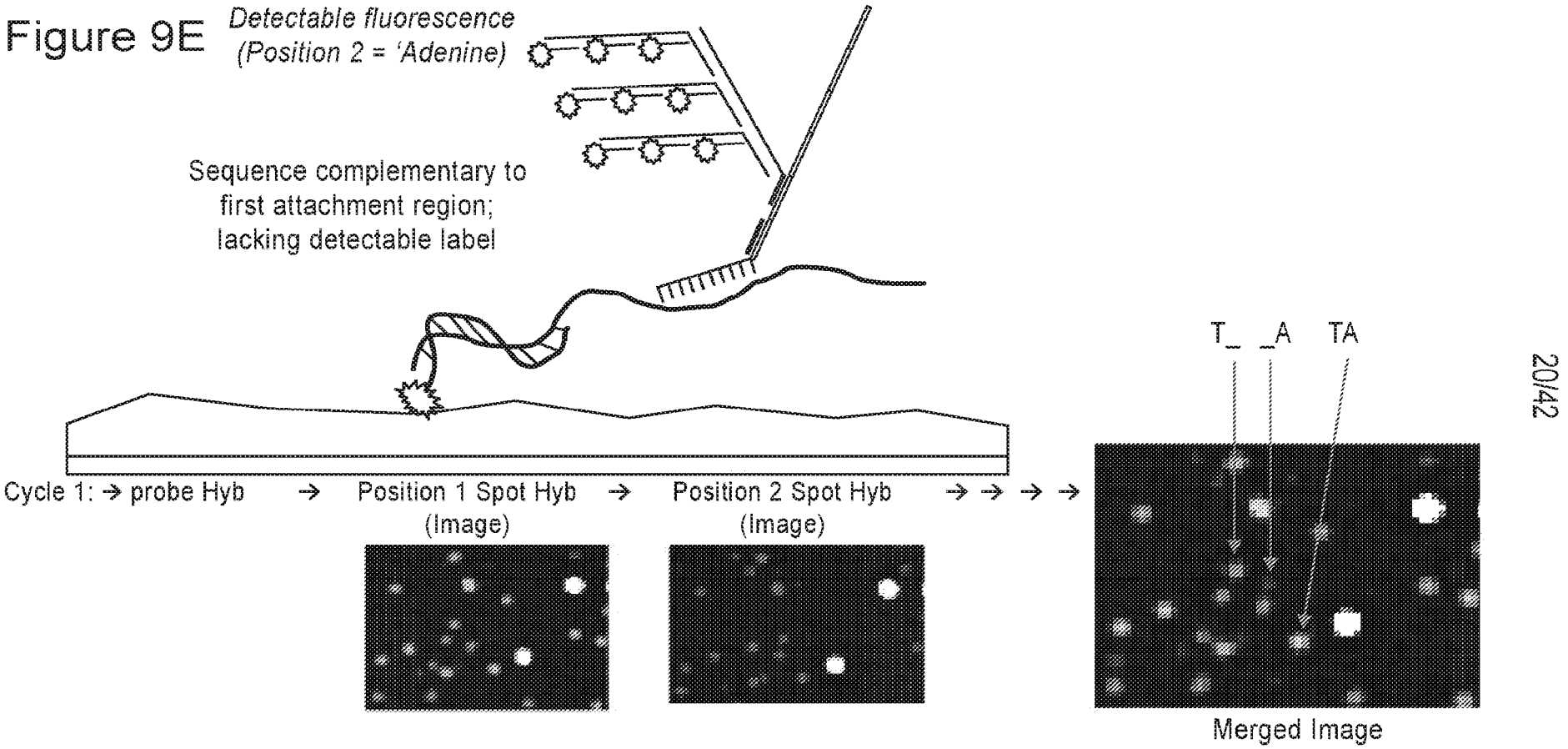

Detectable fluorescence
(Position 1 = Thymine)
Cycle 1: → probe Hyb →
Position 1 Spot Hyb
(Image)

Figure 9E


Detectable fluorescence
(Position 2 = 'Adenine)
Sequence complementary to
first attachment region;
lacking detectable label
Cycle 1: → probe Hyb → Position 1 Spot Hyb → Position 2 Spot Hyb → → → →
(Image)
(Image)
T_
_A
TA
Merged Image

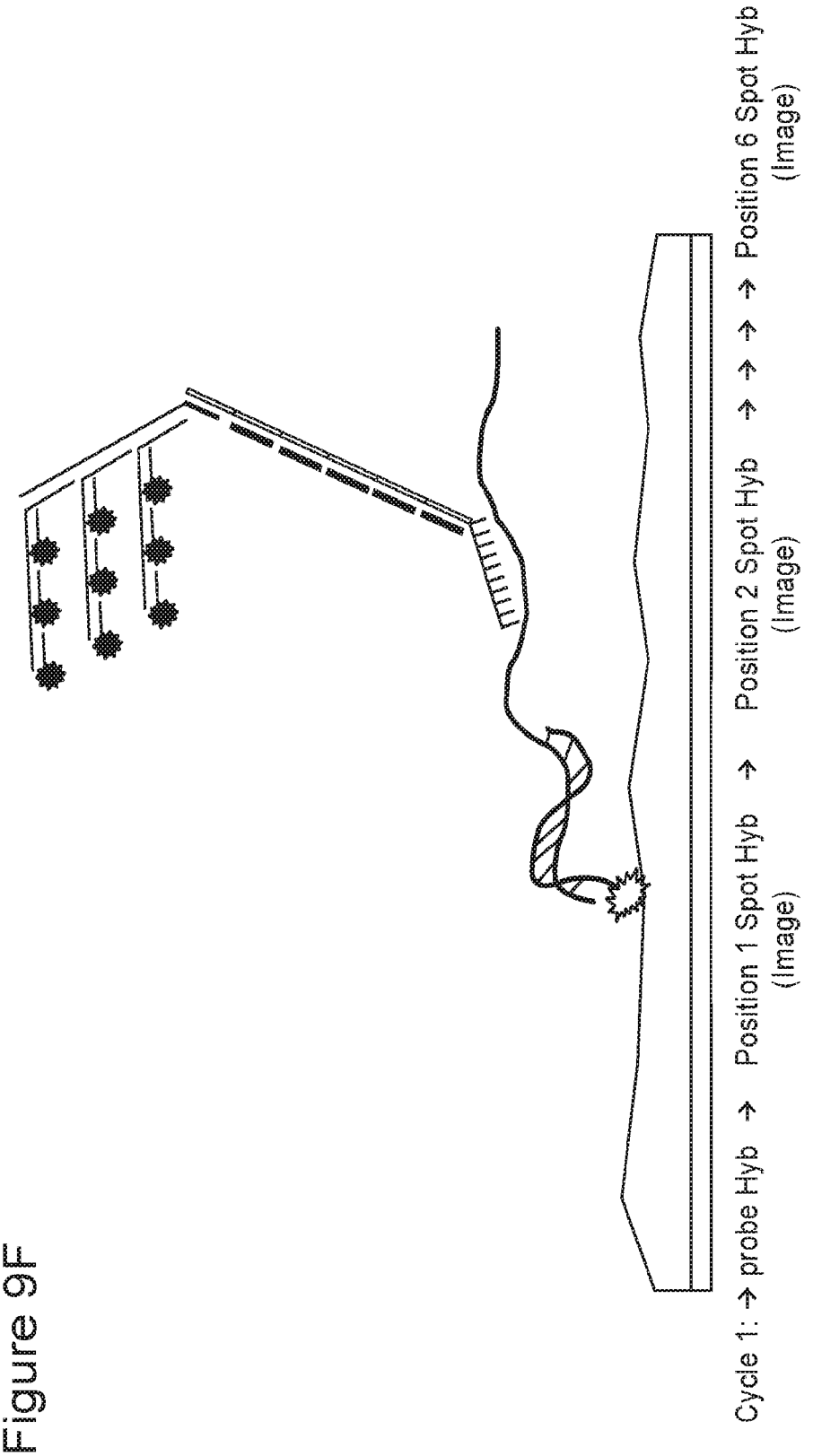

Cycle 1: → probe Hyb → Position 1 Spot Hyb (Image) → Position 2 Spot Hyb (Image) → → → → Position 6 Spot Hyb (Image)

Figure 9F

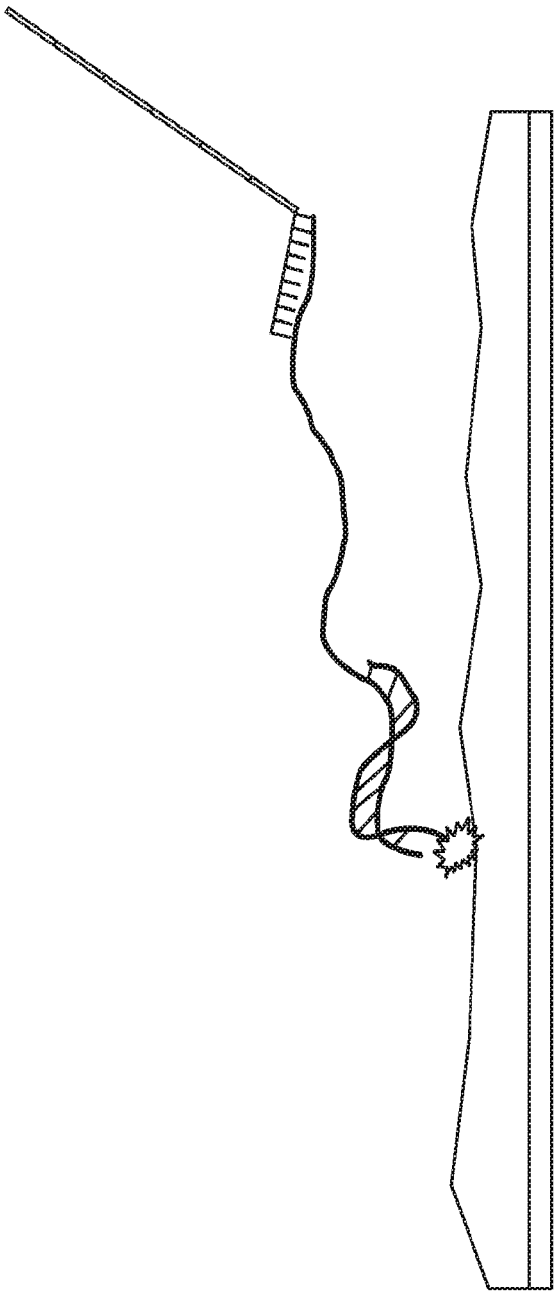


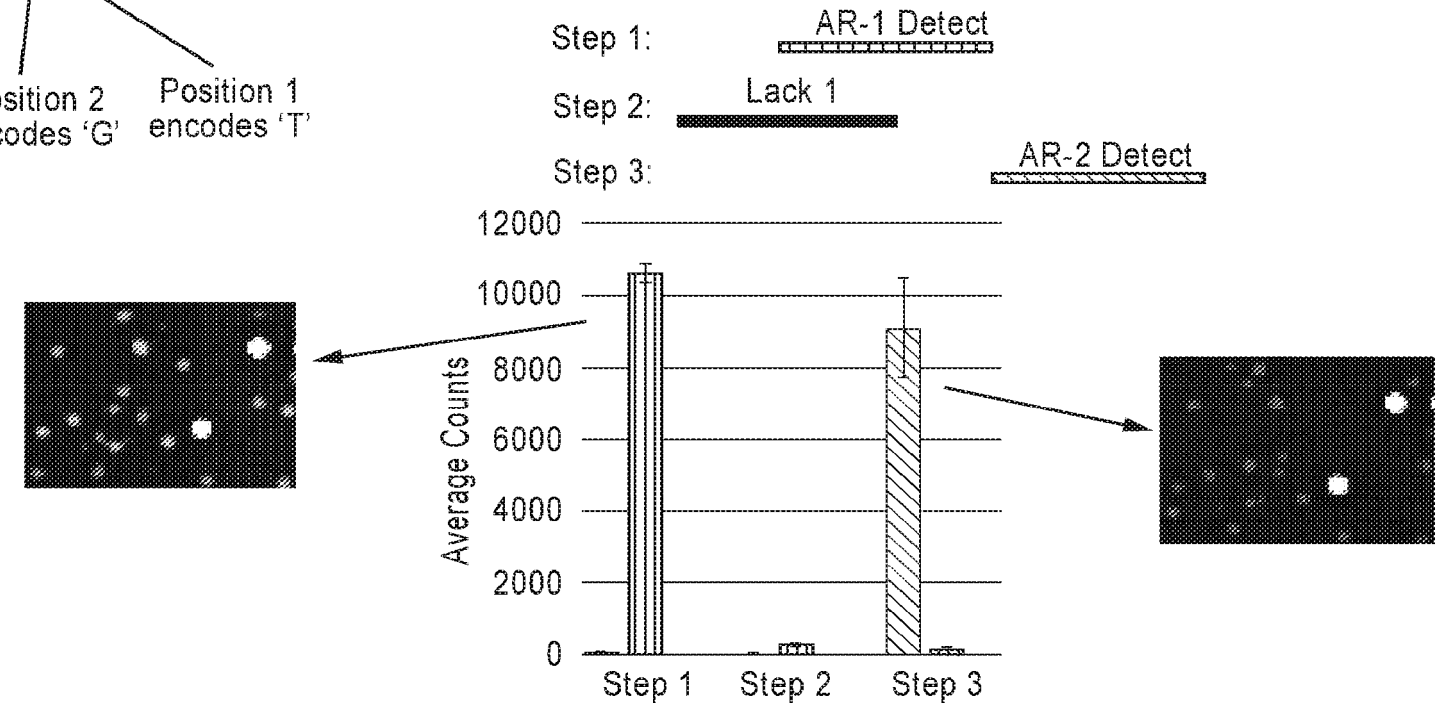
Figure 9G

Figure 10

	Target Binding Domain	(Optional) dsDNA spacer	Flank 1	AR-1	AR-1/ Flank 2	AR- 2	AR-2/ Flank 3	AR-3	AR-3/ Flank 4	AR-4	AR-4/ Flank 5	
V2 Design	TTCACTGTAG	CTGTCTCATTTTGCTGCATCCTGTCCGTTCACGTTG	GAGCTT	GTCATC	CGTCCT	CTTTTC	ACTCCT	AGGCAT	TTGCCT	ATTCGG	CGTCCT	...

Position 2 encodes 'G'

Position 1 encodes 'T'



Flank = Flanking single-stranded polynucleotide
AR = Attachment Region
Detect = Complementary Nucleic Acid Comprising Detectable Label
Lack = Hybridizing Nucleic Acid Lacking Detectable Label

Figure 11

	Domain	(Optional) dsDNA spacer	Flank 1	AR-1	AR-1/ Flank 2	AR- 2	AR-2/ Flank 3	AR-3	AR-3/ Flank 4	AR-4	AR-4/ Flank 5	
V2 Design	TTCACTGTAG	CTGTCTCATTTTGCTGCATCCTGTCCGTTCACGTTG	GAGCTT	GTCATC	CGTCCT	CTTTTC	ACTCCT	AGGCAT	TTGCCT	ATTCGG	CGTCCT	...

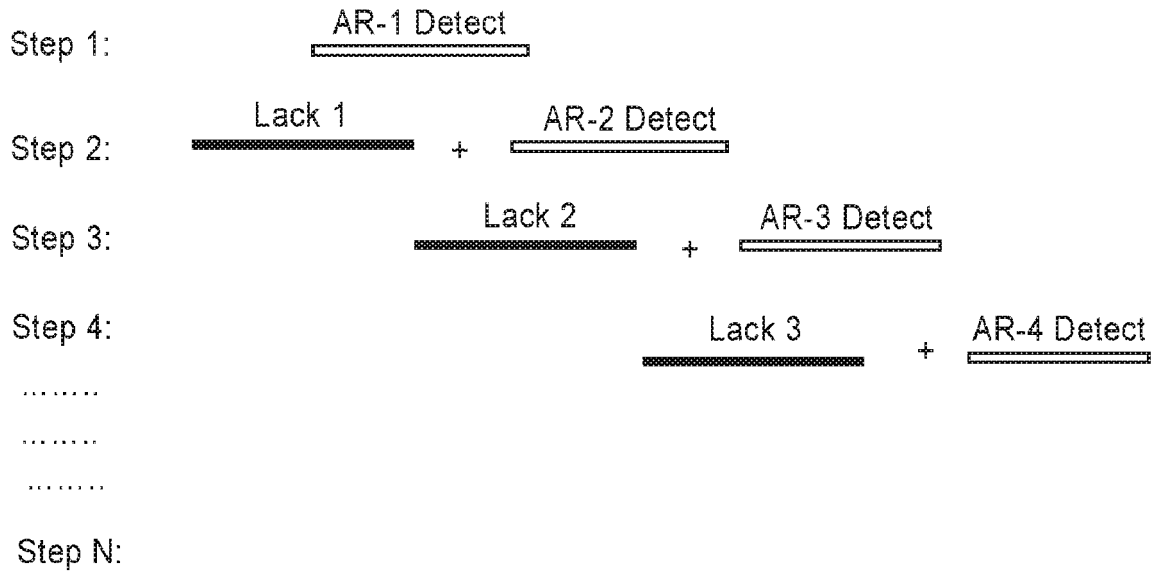
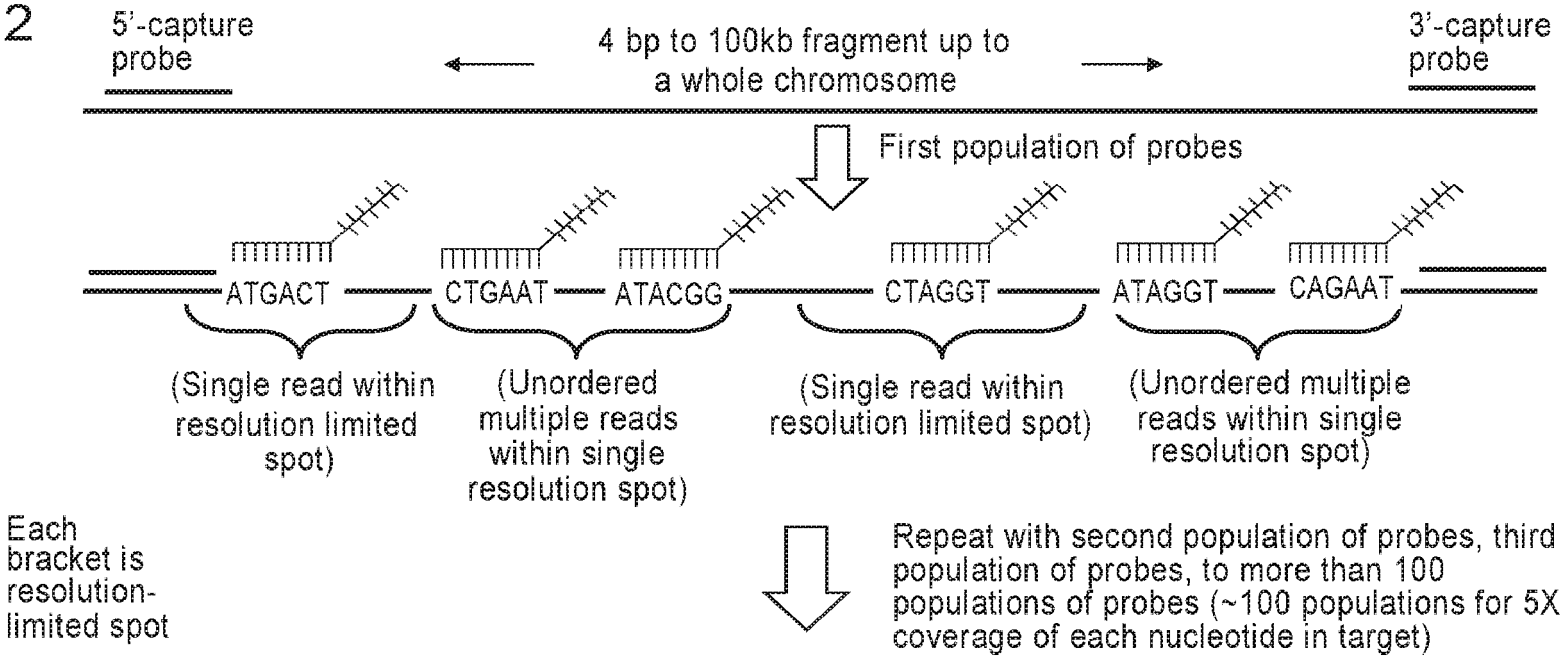


Figure 12



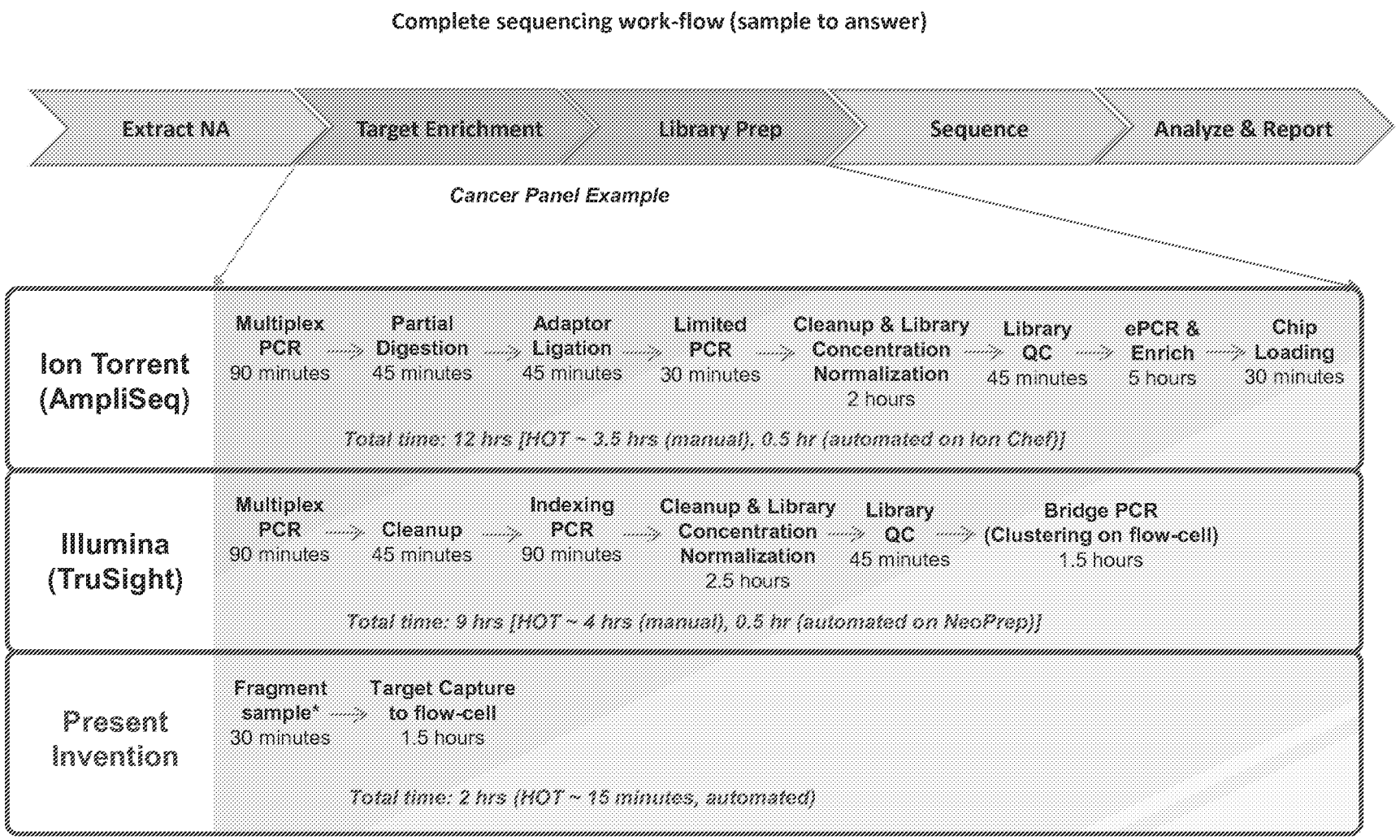
~ 20 FOV (for 100X coverage of a gene) X 8 mins / FOV = 2.7 hour run = 3 Gb/24hrs
(~ 100% MiSeq)

~ 20 FOV X 5000 genes/FOV = 100,000 gene reads (yields 100X of 1000 genes)

~ 20 FOV (for 100X coverage of a gene) X 8 mins / FOV = 2.7 hour run = 12 Gb/24hrs*
(~ 400% MiSeq)

*(ordered-array NSTG throughput estimate)

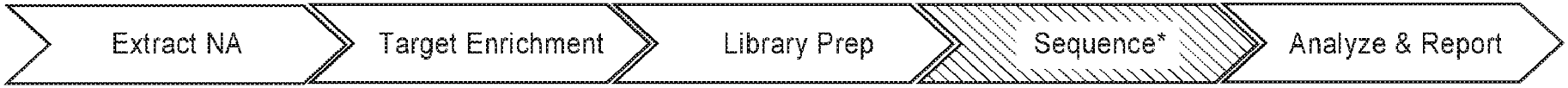
Figure 13



*Need for fragmentation will be dependent on sequencing application (i.e. long-read apps such as HLA sequencing will not require this step)

Figure 14

Complete sequencing work-flow (sample to answer)



Run-time calculator based on sequencing application

Technical Attributes	Raw Accuracy (%)	98	
	Read length (bp)	-	
Application specific seq requirements	# of Target regions	100	
	Target region size (bp)	100	
	Variant detection threshold	0.01	
	# of Target copies required (X)	3000	
	Total # of Reads required (M)	0.3	
Probes detecting 6 nucleotide positions	Total reads per FOV	250000	
	# of FOV required	1.2	
	Total number of hyb&seq cycle	430	
	Hardware motion time (s)	0.5	
	Exposure time (s)	0.2	
	Focus time (s)	0.3	
	Total imaging time per FOV (s)	1	
	Total # of spot per hyb&seq cycle	6	
	k-mer probe hyb time (s)	5	
	Spot exchange time (s)	3	
	Wash time (s)	5	
	Total chemistry time per hyb&seq cycle (s)	28	
	Total seq run time* (hr)	4.62	

Run-times:

100-plex = 4.6 hrs

1000-plex = 16 hrs

* Sequencing run-time includes the time for primary data processing (i.e. basecalling, QV assignment, probe label readout counting & binning)

Figure 15

Cycling calculator based on probe occupancy

Total # of single molecule targets on flow-cell	30000	
Average length of 'gapped' target (bp)	100	
Barcode occupancy (%) per seq cycle	0.75	
# of barcode cycle required to 'hit' >95% of targets on flow-cell at least once	3	
Dye occupancy (%) per position	0.8	
Total number of Dye positions	6	
Total # of 1 bp encoded read per seq cycle	6	35
Total # of 2 bp encoded read per seq cycle	23	346
Total # of 3 bp encoded read per seq cycle	92	1843
Total # of 4 bp encoded read per seq cycle	369	5530
Total # of 5 bp encoded read per seq cycle	1475	8847
Total # of 6 bp encoded read per seq cycle	5898	5898
Total number of 'usable' read per seq cycle		22499
# of barcode cycle required to cover given single molecule target with ~3X coverage	143	
# of barcode cycle required to cover given single molecule target with ~5X coverage	238	
# of barcode cycle required to cover all single molecule target on flow-cell with ~3X coverage	430	
# of barcode cycle required to cover all single molecule target on flow-cell with ~5X coverage	713	

Figure 16

	Gigabases	time hrs.	# Reads	Clinical utility
Hiseq 2500	900Gb	11 days	3 Billion !	Genome sequence cancer sequence
Next Seq	120GB	27hrs	400M	Genome sequence exome Transcriptome
Proton PI	10Gb	2-4	70M	exome Transcriptome
Present Invention	3 to 12 Gb	3	0.5M	Targeted (Dx, cDx) plus multi-analyte plus sample-to-answer, long-reads
MISeq	3.5Gb	24hrs	25M	transcriptome targeted (Dx)
PGM 318	up to 2Gb	7.3/4.4	4M	Targeted
PacBio RS II	2.2Gb/day	3 hrs	50k	Niche genomes, long-reads, high error rate
PGM 314	0.1Gb	3.7/2.3	400k	Targeted

Figure 17

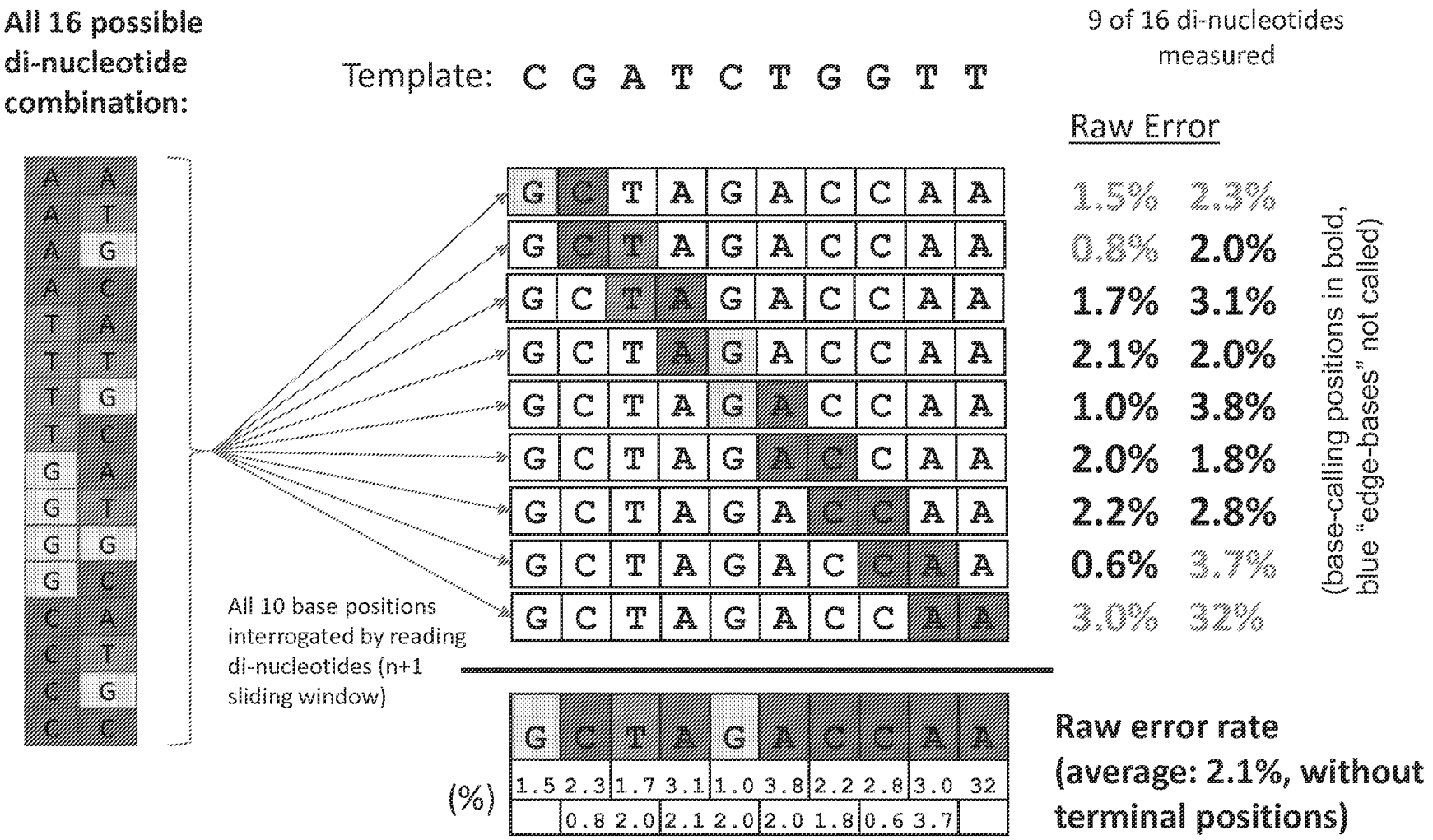

All 16 possible di-nucleotide combination:
Template: C G A T C T G G T T
9 of 16 di-nucleotides measured
Raw Error
1.5% 2.3%
0.8% 2.0%
1.7% 3.1%
2.1% 2.0%
1.0% 3.8%
2.0% 1.8%
2.2% 2.8%
0.6% 3.7%
3.0% 32%
(base-calling positions in bold, blue "edge-bases" not called)
All 10 base positions interrogated by reading di-nucleotides (n+1 sliding window)
(%)
1.5 2.3 1.7 3.1 1.0 3.8 2.2 2.8 3.0 32
0.8 2.0 2.1 2.0 2.0 1.8 0.6 3.7
Raw error rate (average: 2.1%, without terminal positions)

Figure 18

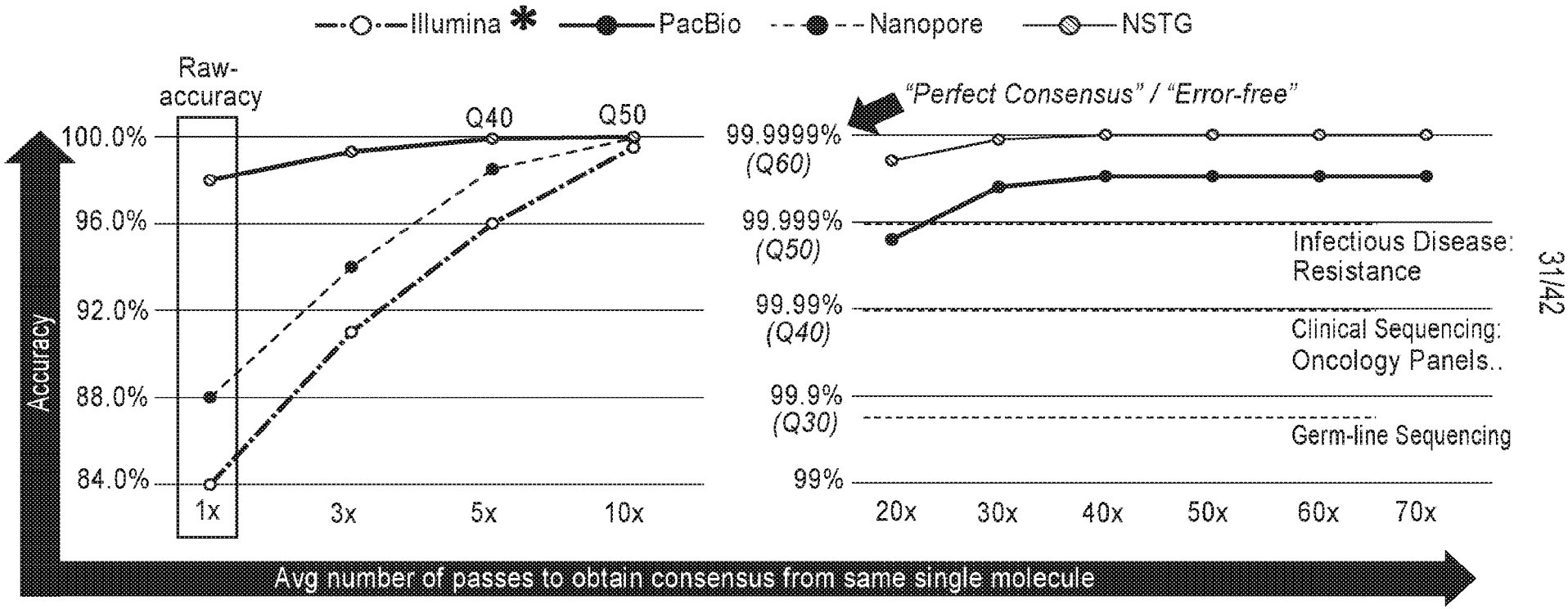
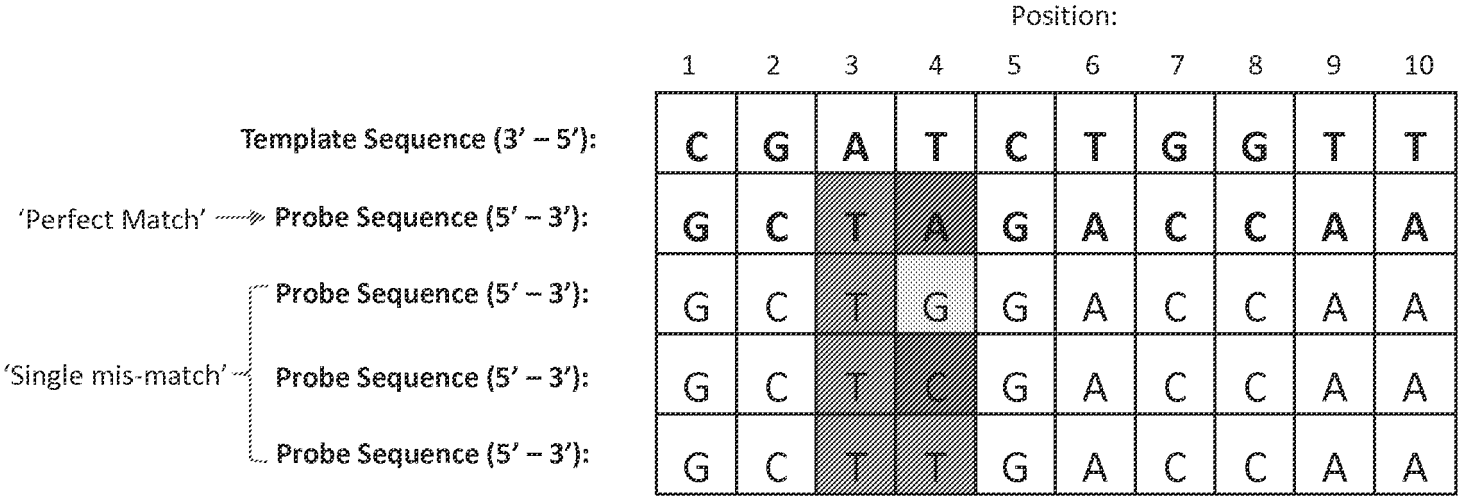

Illumina
PacBio
Nanopore
NSTG
Raw-accuracy
Q40
Q50
"Perfect Consensus" / "Error-free"
100.0%
96.0%
92.0%
88.0%
84.0%
99.9999% (Q60)
99.999% (Q50)
99.99% (Q40)
99.9% (Q30)
99%
Infectious Disease: Resistance
Clinical Sequencing: Oncology Panels..
Germ-line Sequencing
1x
3x
5x
10x
20x
30x
40x
50x
60x
70x
Accuracy
Avg number of passes to obtain consensus from same single molecule

Figure 19



Counts observed*	% of Total valid counts
2543	98%
40	1.5%
13	0.5%
1	0%
2636	100%

Total:

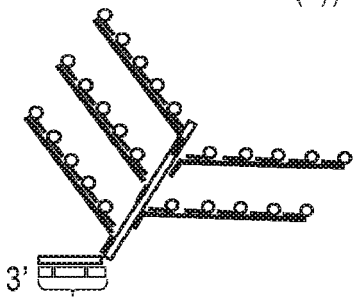
* Counts observed value represent counts after background count subtraction

**All hybridization done under buffer condition: 1xSSPE/0.1% Tween20

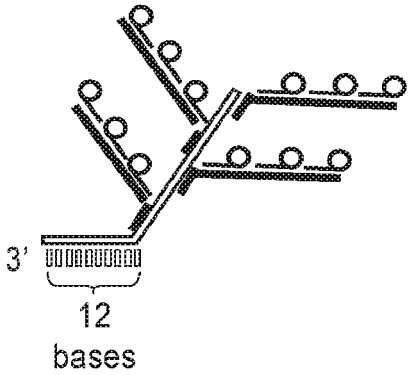
Figure 20A

5x5

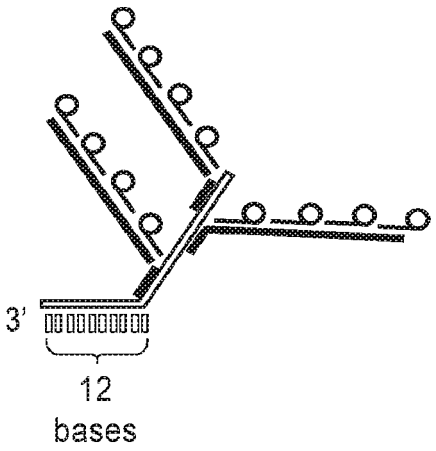
(1 primary nucleic acid molecule, 5 secondary nucleic acid molecules, 5 tertiary nucleic acid molecule with detectable label(s))



4x3



3x4



Spacer 3x4

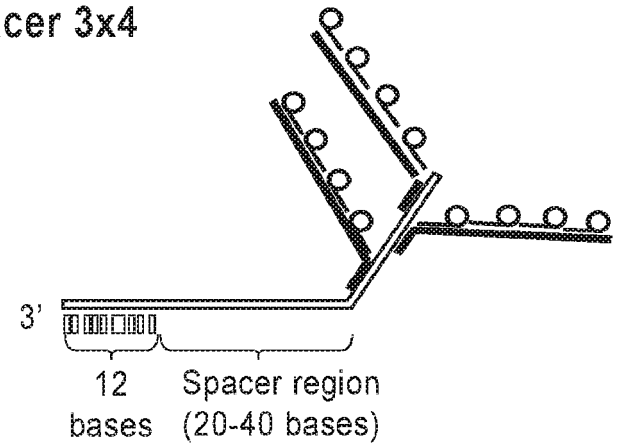


Figure 20B

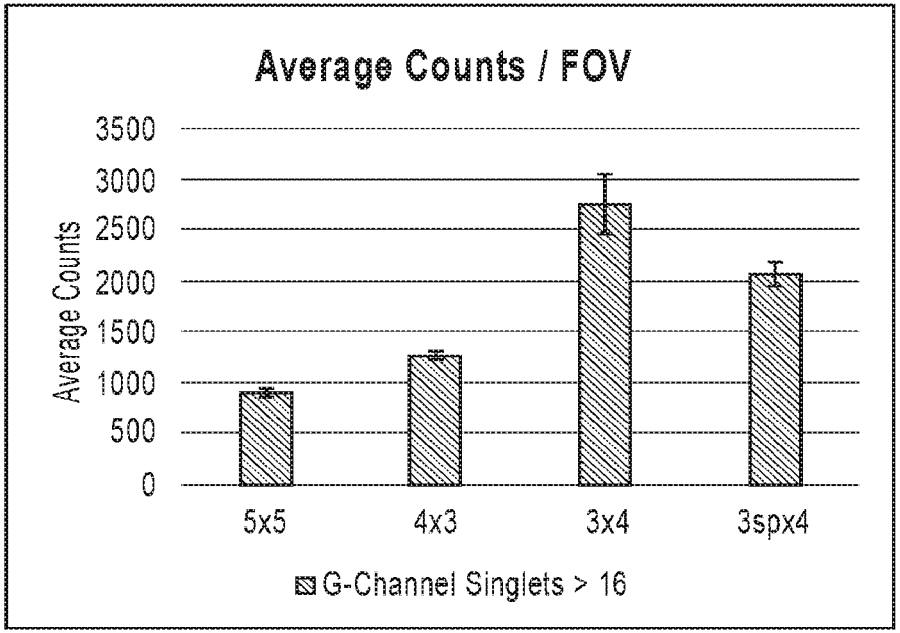


Figure 20C

Exemplary recipe for constructing reporter complexes comprising detectable labels

Volume (uL)	Primary n. a. molecule (10uM)	Secondary n. a. molecule (10uM)	Tertiary n. a. molecule (100uM)	H20
5x4	1	4.5	2.25	92.25
5x3	1	4.5	1.8	92.7
4x4	1.28	4.5	2.25	91.97
4x3	1.28	4.5	1.8	92.42
3x4	1.8	4.5	2.25	91.45

Colors in table correspond to colors in diagrams of complexes shown at left

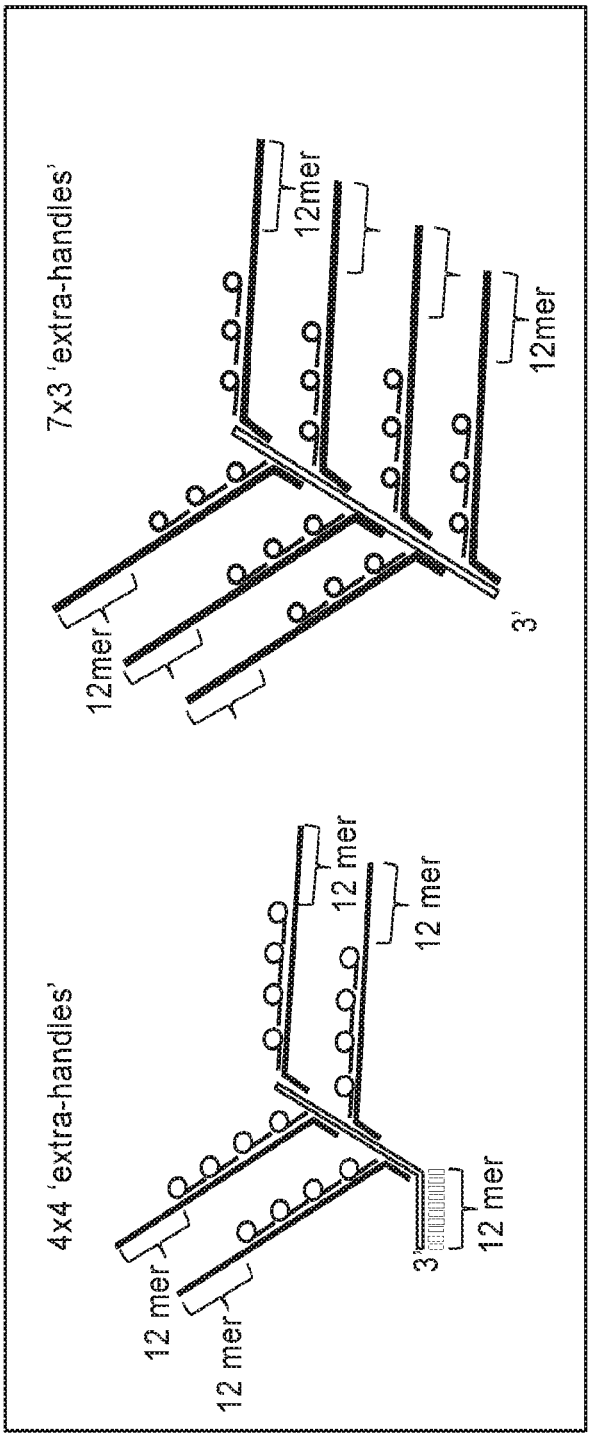


Figure 21A

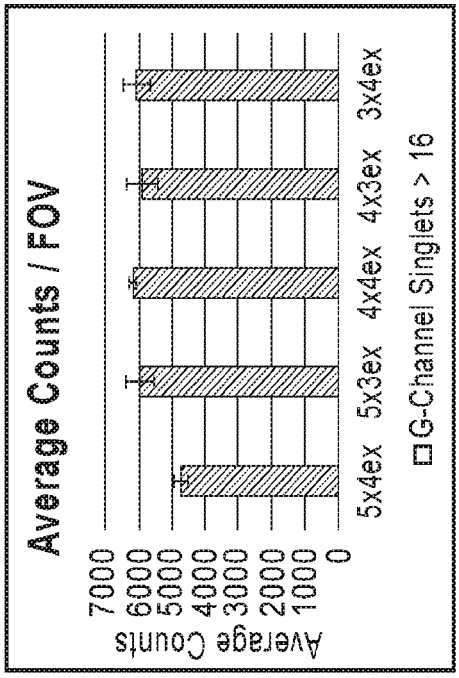


Figure 21B

Figure 22A

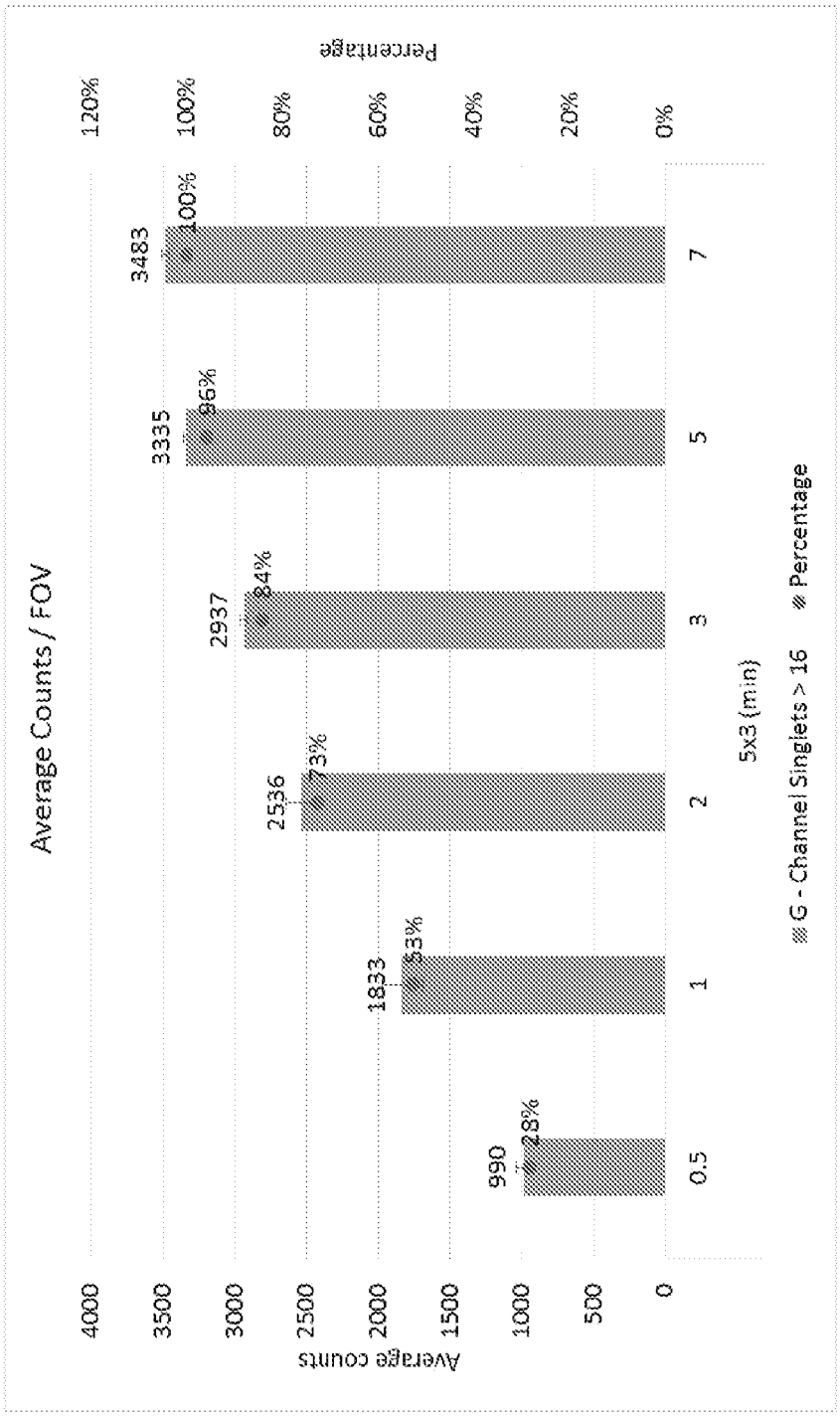
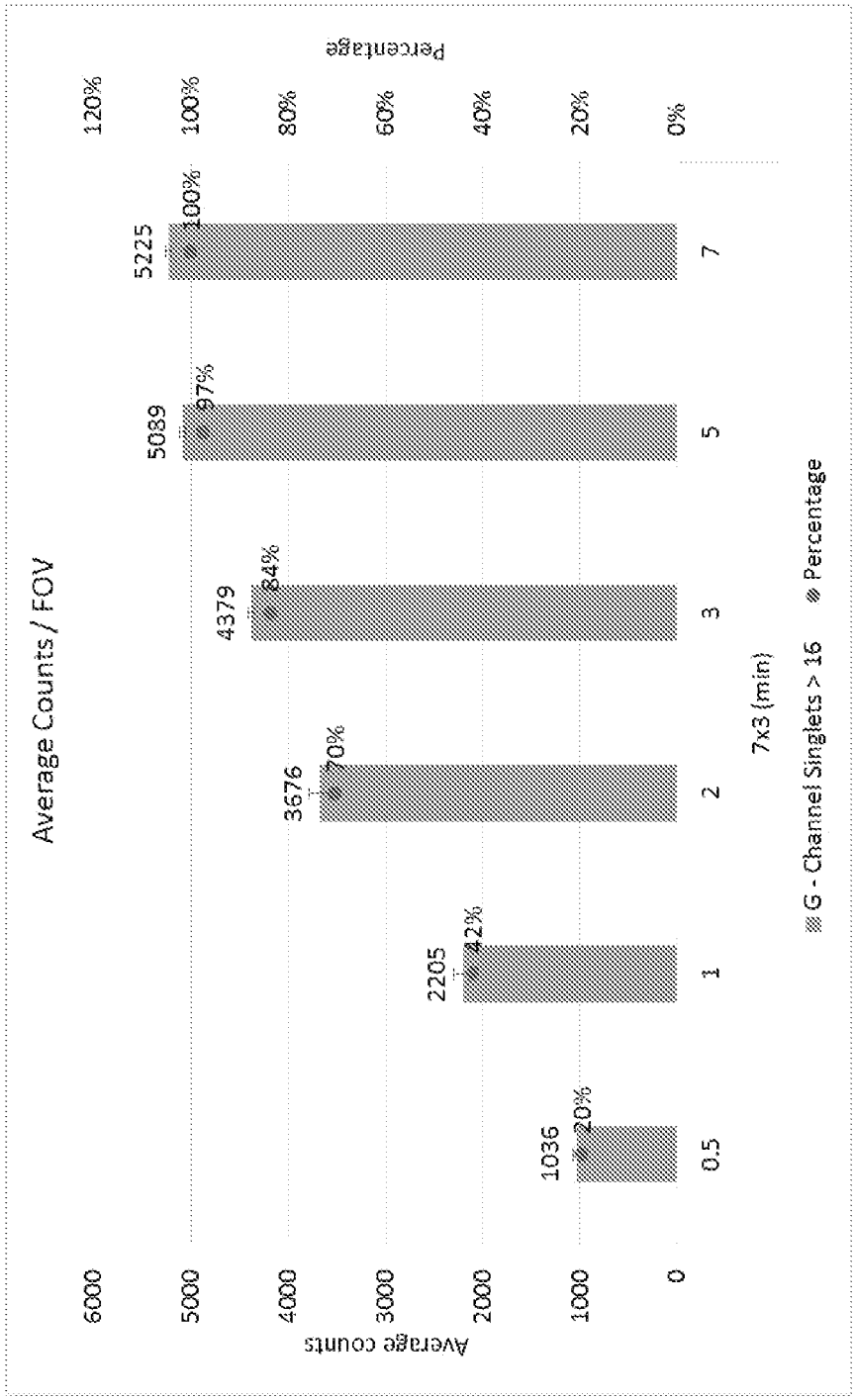

'5x3' Kinetics
Average Counts / FOV
Average counts
Percentage
4000
3500
3000
2500
2000
1500
1000
500
0
120%
100%
80%
60%
40%
20%
0%
990
28%
1833
53%
2536
73%
2937
84%
3335
96%
3483
100%
0.5
1
2
3
5
7
5x3 (min)
G - Channel Singlets > 16
Percentage

Figure 22B


'7x3 extra handles' Kinetics
Average Counts / FOV
Average counts
Percentage
6000
5000
4000
3000
2000
1000
0
120%
100%
80%
60%
40%
20%
0%
1036
20%
2205
42%
3676
70%
4379
84%
5089
97%
5225
100%
0.5
1
2
3
5
7
7x3 (min)
G - Channel Singlets > 16
Percentage

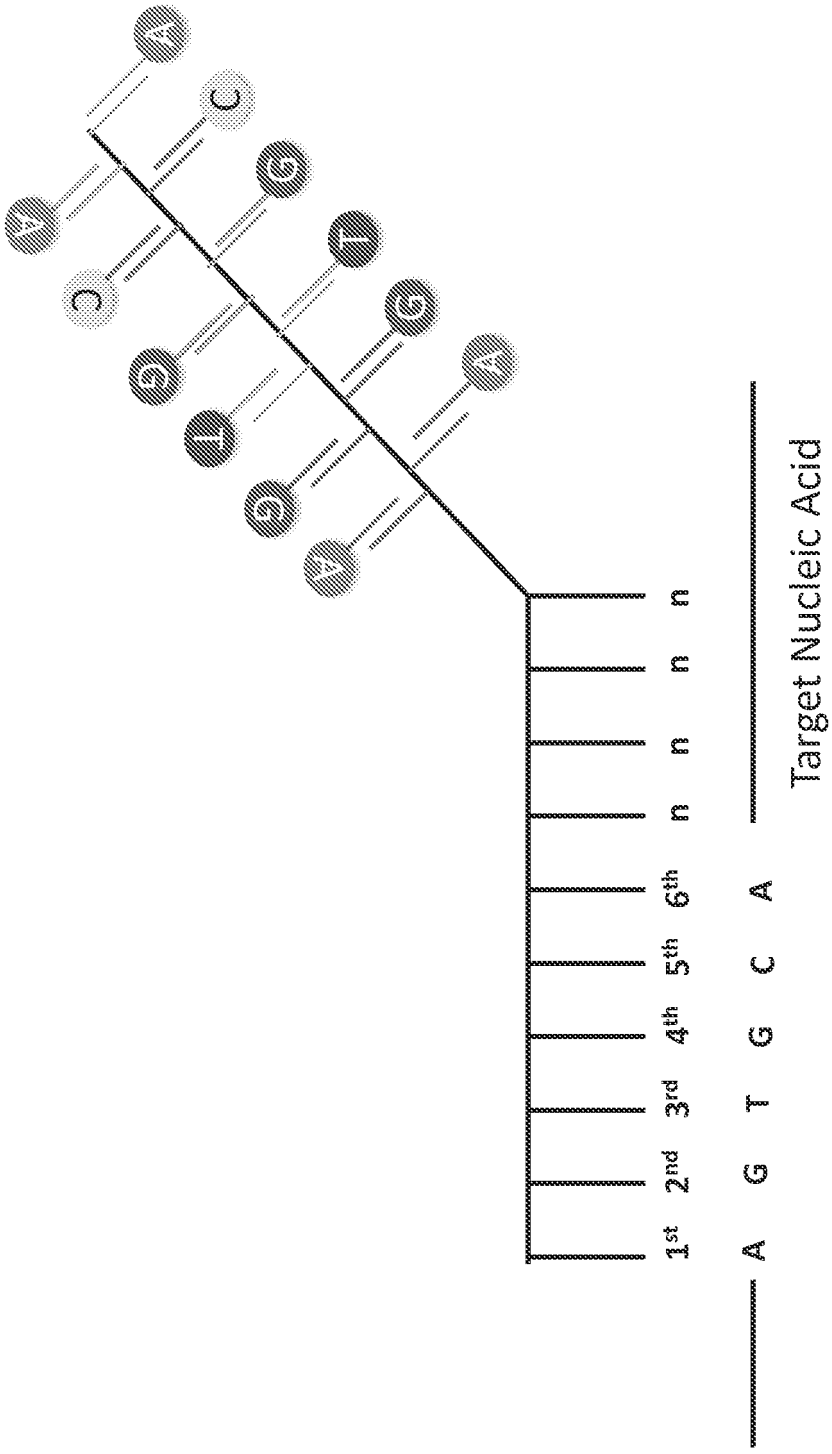

A
C
G
T
G
A
A
C
G
T
G
A
1st 2nd 3rd 4th 5th 6th
n n n n n n
A G T G C A
Target Nucleic Acid

Figure 23

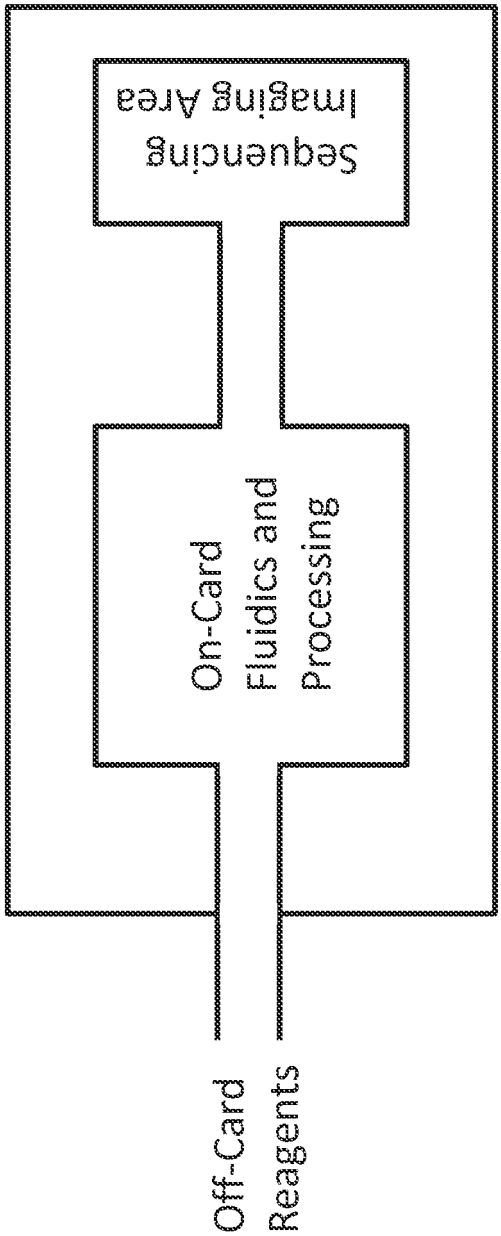

Sequencing Imaging Area
On-Card Fluidics and Processing
Off-Card Reagents

Figure 24

Figure 25

code	sequences
YGBYGR	GAGACTGTAC - original
BGYRGB	GAGACAGTAC
YBRYGB	GAGACCGTAC
RGBYGR	GAGACGGTAC
YRBGYR	GAGAGTGTAC
GRYGRB	GAGAATGTAC
RGYBRG	GAGATTGTAC

code	counts	sequences
YGBYGR	16074	GAGACTGTAC
RGYBRG	1132	GAGATTGTAC
YBRYGB	998	GAGACCGTAC
YRBGYR	784	GAGAGTGTAC
YBRYBG	712	
YBGRBY	699	
YGRYBG	692	
YBRGBR	674	
Background Threshold	717	

Figure 26

DNA	Sequence	code	count
12 mer	CTGAGACTGTAC	RYGBRG	80845
11 mer	TGAGACTGTAC	GBRYGB	24474
10 mer	GAGACTGTAC	YGBYGR	18339
9 mer	AGACTGTAC	YRBYRG	1667
Mirror of “12 mer”	CATGTCAGAGTC	GRBGYR	1473
Mirror of “11 mer”	CATGTCAGAGT	BGYRBG	1384
8 mer	GACTGTAC	GBRYBG	900
		Background Threshold	622

Figure 27

Full length Reporter

20 pM full length (6 spot) ***33*** counts/fov

Single Spot Reporter

Single spot [fM]	Counts/FOV
20,000	Saturated
2,000	12000
200	308
20	***40***
2	11

20pM NS reporter ≃ 20fM Single spot